DOCUMENT RESUME

ED 302 020 EC 211 817

AUTHOR	Fuchs, Lynn S.; And Others
TITLE	Effects of Alternative Goal Structures within Curriculum-Based Assessment.
INSTITUTION	George Peabody Coll. for Teachers, Nashville, Tenn.
SPONS AGENCY	Special Education Programs (ED/OSERS), Washington, DC.; Spencer Foundation, Chicago, Ill.
PUB DATE	1 Jan 88
GRANT	G008530198
NOTE	32p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Diagnostic Teaching; *Disabilities; Elementary Education; *Goal Orientation; *Instructional Effectiveness; *Student Educational Objectives; Student Evaluation; *Teaching Methods
IDENTIFIERS	*Curriculum Based Assessment

ABSTRACT

The study assessed the effects of alternative goal structures within curriculum based assessment (CBA) in the area of math. Subjects were 30 elementary level special education teachers, assigned randomly to a dynamic goal CBA, static goal CBA, or control group for 15 weeks. Two pupils in each class were identified to evaluate the effects of the instructional intervention. In the dynamic goal condition, teachers employed CBA, and (1) modified instructional programs when student progress fell below expectations and (2) increased goals when student progress exceeded expectations. In the static goal condition, teachers employed CBA and modified programs when progress was below expectations, but did not systematically increase goals in response to progress that exceeded anticipated improvement rates. Multivariate analyses of variance conducted on fidelity of treatment measures indicated that dynamic goal teachers increased goals more frequently and, by the study's completion, employed more ambitious goals. Multivariate analyses of covariance indicated that students in the dynamic goal group had better content mastery than control students, whereas students in the static goal group did not. Content coverage for the three groups was comparable. Implications for instructional goal-setting practice are discussed. (Author/DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED302020

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Effects of Alternative Goal Structures within Curriculum-Based Assessment

Lynn S. Fuchs, Douglas Fuchs, and Carol L. Hamlett
Peabody College of Vanderbilt University

January 1, 1988

Requests for reprints should be sent to Lynn S. Fuchs, Box 328, Peabody College, Vanderbilt University, Nashville, TN 37203.

This research was supported by a Spencer Foundation Seed Grant and Grant #G008530198 from the Office of Special Education Programs, U.S. Department of Education, to Vanderbilt University. Points of view or opinions stated in this report do not necessarily represent official agency positions.

Running Head: GOALS WITHIN CURRICULUM-BASED ASSESSMENT

EC 211817

Abstract

This study assessed the effects of alternative goal structures within curriculum-based assessment (CBA) in the area of math. Subjects were 30 special education teachers, assigned randomly to a dynamic goal CBA, static goal CBA, or control group for 15 weeks. Teachers implemented the procedures with their classes; within each class, two pupils were targeted as research participants for evaluating the effects of the instructional intervention. In the dynamic goal condition, teachers employed CBA, and (a) modified instructional programs when student progress fell below expectations and (b) increased goals when student progress exceeded expectations. In the static goal condition, teachers employed CBA and modified programs when progress was below expectations, but did not systematically increase goals in response to progress that exceeded anticipated improvement rates. Control teachers monitored student progress using conventional practice. Multivariate analyses of variance conducted on fidelity of treatment measures indicated that dynamic goal teachers increased goals more frequently and, by the study's completion, employed more ambitious goals. Multivariate analyses of covariance indicated that students in the dynamic goal group had better content mastery than control students, whereas students in the static goal group did not. Content coverage for the three groups was comparable. Implications for instructional goal-setting practice are discussed.

Effects Alternative Goal Structures within Curriculum-Based Assessment

Curriculum-based assessment is a generic term that most commonly refers to the evaluation of instructional needs by measuring pupil performance within the local school curriculum (Gickling, 1981; Tucker, 1985). Defined within the framework of measurement theory, where curricular validity represents the correspondence between tests and programmatic goals (Yalow & Popham, 1983), curriculum-based assessment links the parameters of testing to the goal statements that constitute school curricula.

Consequently, within curriculum-based assessment, meaningful goal specification is critical. Specification of a goal precedes and defines the curriculum-based assessment process. The goal dictates (a) the material on which measurement will occur, (b) the behavior to be observed, and (c) the criteria for judging attainment. For example, if a goal specified that a student would read third grade material with fluency and comprehension by year's end, the curriculum-based assessment would be operationalized in the following way. Stimulus materials for the assessment would comprise third grade passages from the basal text; the behaviors observed and scored during measurement would be oral reading fluency (e.g., words read correctly per minute) and comprehension (e.g., content words retold from the passage); and the criteria for judging proficiency would be standards for fluency (e.g., 70 words correct per minute) and comprehension (e.g., 40 content words retold).

Given the close connection between goals and measurement in this process, it would appear that selection of appropriate goals, which are both realistic and ambitious, is essential to effective curriculum-based assessment.

Despite the apparent importance of goals to curriculum-based assessment, the literature on appropriate goal specification is extremely limited (see for example Fuchs, 1986a). In one long-term, large-scale project, curriculum-based assessment improved pedagogy and enhanced student achievement (Fuchs, Deno, & Mirkin, 1984). However, post-hoc analyses indicated that the ambitiousness of the goal employed within the ongoing curriculum-based assessment may have mediated student achievement outcomes (Fuchs, Fuchs, & Deno, 1985). Although this finding supports use of ambitious goals, it fails to provide useful guidelines for determining ambitious goals for individual students. An absense of such guidelines has serious implications for practice. When a practitioner formulates a goal, especially for a poorly achieving pupil, the broad outline and direction of its mission are clear: The child's rate of progress must be accelerated. But the extent to which the student's learning rate can be enhanced remains unclear, and corresponding procedures for identifying appropriately ambitiousness, but realistic, goals are unknown.

Consequently, the first purpose of this study was to explore alternative goal structures within a particular variant of curriculum-based assessment, known as "curriculum-based measurement" (CBM) (Deno, 1987). CBM has been developed empirically during the past decade and provides clear guidelines for many dimensions of the assessment process, including measurement methods,

graphing procedures, and aspects of data-utilization (see Deno, 1985 1986; Deno & Fuchs, 1987; Fuchs, 1986b). However, CBM development has not adequately addressed procedures for goal specification.

The current study contrasted a conventional CBM process, wherein teachers measure and evaluate student progress with respect to a fixed annual goal, to a contrasting CBM procedure, requiring teachers to measure and evaluate student progress in relation to a dynamic goal. In this dynamic goal condition, goals are adjusted upward whenever actual student progress indicates that a more ambitious rate of progress, or a more difficult goal, may be attainable. Within this innovative CBM process, the notion of goal use is reconceptualized: Instead of viewing it as a static, terminal component that fixes dimensions of assessment, the goal is related dynamically to evaluation and instruction. This study contrasted static and dynamic CBM goal structures with a control group, which did not employ systematic student monitoring procedures. Results should clarify (a) how goals affect the curriculum-based assessment process and (b) how teachers might develop ambitious but realistic goal statements.

The second purpose of this study was to investigate the scope of skills which CBM affects. Some have proposed (see Slavin, 1987, for a related discussion) that, although closely linked measurement and instruction may improve student mastery of the skills specified for assessment, it may reduce the extent to which teachers cover a broader curricular scope. Within this framework, the hypothesis favors the performance of CBM groups on content mastery measures, which tap the extent to which students achieve skills

closely related to CBM systems. However, the hypothesis predicts that performance on content coverage indices, which assess the extent to which students achieve a broader set of related skills, would favor control groups. The current study investigated this hypothesis by including content mastery and content coverage measures. Additional outcomes were teachers' fidelity of treatment, including indices of their goal-setting behavior.

Method

Subjects

Teachers. Participants were 30 teachers in 16 schools in a southeastern metropolitan area. Teachers were assigned randomly to three treatment groups: dynamic goal CBM, static goal CBM, and control. One-way analyses of variance revealed no significant differences among teacher groups on the following variables: age level, total years teaching, years computer use, years in current school, and the self-reported number of times per year teachers had evaluated students' math performance prior to the study. Descriptive and inferential statistics on these variables are displayed in Table 1. Additionally, a chi-square test applied to teachers' highest educational degree revealed no statistically significant difference among groups, $\chi^2(4, N = 30) = 2.54$, ns. In the dynamic goal CBM, static goal CBM, and control groups, respectively, there were 4, 5, and 6 teachers with bachelor's degrees; 5, 5, and 4 with master's degrees; and 1, 0, and 0 with specialist's certificates.

Insert Table 1 about here

Information also was collected on two aspects of teacher efficacy, personal teaching efficacy and general teaching efficacy. These factors of the Teacher Efficacy Scale (Gibson & Dembo, 1984), which correspond to Bandura's (1977, 1978, 1982) two factor theoretical model of self-efficacy, (a) demonstrate adequate convergent and discriminant validity and (b) are associated with effective teaching variables including size of instructional group, use of criticism, and persistence in failure situations (Gibson & Dembo, 1984). A multivariate one-way analysis of variance was conducted on these two scales, using Wilks' lambda to test for equality of group centroids. This analysis revealed no significant differences among groups, F (4,52) = .33, ns (Wilks' lambda = .951). Univariate F values and descriptive statistics for these efficacy factors are shown in Table 1.

Students. From each teacher's class, two students were randomly selected to serve as the research participants on which study procedures would be assessed. One-way analyses of variance conducted on students' chronological age, grade placement, and level of math performance as estimated by their teachers revealed no significant differences. Descriptive and inferential statistics on these variables are shown in Table 1. Additionally, chi-square tests applied to students' race and sex indicated the experimental groups were comparable. In each group, 19 boys and one girl participated. In the dynamic goal CBM, static goal CBM, and control groups, respectively, there were 9, 9, and 8 minority students.

Measures

Two types of math achievement tests were employed: (a) a curriculum-based Math Computation Test, which served as the content mastery measure, and (b) a math concepts test, the Concepts of Number subtest of the Stanford Achievement Test, which served as the content coverage measure.

Curriculum-based test. The Math Computation Test (MCT; Fuchs, 1987a) samples math problems across grades 1 through 6 from the computation objectives of the Tennessee Basic Skills First Math Curriculum. This curriculum encompasses a state-wide set of competencies that are (a) expected for promotion across grades and (b) tested annually in a state-wide criterion-referenced testing program. Pupils are provided directions in standard format and have 10 minutes to complete 36 problems. Performance is scored in terms of number of correct digits written. Digits allow credit for partially correct problems and, as an index of achievement, appear to be more sensitive than correct problems (see Deno, 1985 for related discussion).

Criterion validity of the MCT with respect to the Concepts of Number and Math Computation subtests of the Stanford Achievement Test was .82 and .83, as calculated on this sample. Internal consistency reliability (Cronbach's alpha) obtained in this study was .93. Interscorer agreement, calculated on 30% of the protocols in this study, was 98%. (Percentage of agreement = [agreement between Rater A and Rater B/(agreements between A and B, disagreements between A and B, and omissions)] X 100, see Coulter cited in Thompson, White, & Morgan, 1982.)

This measure was deemed a useful and appropriate measure of content mastery because of the following reasons. Since it was derived from a

state-wide curriculum equally applicable to the experimental and control groups, it should have assessed achievement on curriculum targeted for all groups, including experimental and control students (see Slavin, 1987, for related discussion). Moreover, the MCT represents more than a simple index of mastery of each individual's CBM curriculum: It encompasses the entire grades 1 through 6 computation curriculum, not just the subset of the curriculum incorporated into any particular child's goal.

Commercial standardized test. Form E of the Primary 2, Primary 3, and Intermediate 1 levels of the Concepts of Number (CN) subtest of the Stanford Achievement Test (Gardner, Rudman, Karlsen, & Merwin, 1982) was employed. The level administered to each pupil was determined by matching the teacher's estimated math level as closely as possible to the grade levels corresponding to a form of the test. The CN requires pupils to respond to multiple-choice questions concerning number concepts by marking their selected answers. At the Primary 2 and 3 levels, examiners read the question numbers and stems aloud, while pupils follow along silently. Examiners pause for 10 seconds between questions to allow pupils to mark answers. At the Intermediate 1 level, pupils read questions to themselves and complete the questions at their own pace, but within a maximum of 20 minutes.

For the Primary 2 level of CN, Kuder-Richardon Formula #20 and alternate-form reliability coefficients were both .86. Intercorrelations among the Stanford Tests and correlations with the Otis-Lenon School Ability Test were between .62 and .78. For the Primary 3 level, Kuder-Richardson and alternate-form reliability coefficients also were .86; intercorrelations and

coefficients with the Otis-Lennon ranged from .62 to .72. Reliability figures for the Intermediate 1 form were .88 and, .82, with validity coefficients between .59 and .81. Standard scores, which are comparable across levels of the test, were employed in analysis. Interscorer agreement (see formula above), calculated on 30% of the protocols in this study, was 98%.

In this investigation, this outcome was conceptualized as an index of content coverage, rather than content mastery. (See Slavin, 1987 for related discussion.) As an index of content coverage, it assessed the extent to which teachers covered material beyond the scope of the CBM computation curriculum.

Treatments

Curriculum-based measurement. For 15 weeks, each experimental teacher (including dynamic and static goal teachers) employed CBM to track their target pupils' progress toward math goals. The CBM system was rooted in the Tennessee Basic Skills First Math Program (BSF). The math computation objectives tested at each grade level within the BSF were listed. Teachers inspected these lists and determined an appropriate grade level on which to establish each student's goal. This level included the pool of math objectives the teacher hoped the student would master by the year's end.

Using a standard measurement task, teachers assessed each pupil's math performance at least twice weekly, for 2 minutes, each time on a different probe representing the type and proportion of problems from the BSF goal level they had designated. So, if the teacher set a student's goal as third grade, the teacher was provided with 50 alternate test forms, each of which sampled the BSF third grade computation objectives in the proportion tested on the BSF

third grade criterion-referenced end-of-year test. Each CBM test comprised 36 problems, displayed in random order, encompassing randomly generated numerals. Therefore, each CBM test could be conceptualized as a short form of the BSF third grade computation test. So, as teachers monitored pupil growth on these tests, they could estimate progress toward mastery of the corresponding grade level of the BSF computation curriculum. Performance was scored in terms of number of correct digits per 2 minutes.

Teachers administered and scored these bi-weekly CBM tests in one of two ways. Half of the teachers administered and scored the assessments themselves; the other half employed software that automatically administers tests to pupils and scores performance (see Fuchs, Hamlett, & Fuchs, 1987 for description of this software). With automatic data collection, children are assessed at a computer and scores are saved to a teacher's data-management disk for decisionmaking. Teachers administering the tests themselves were required to enter pupil performance scores onto the data-management disk by hand. However, the administration and scoring procedures for the teacher- and computer-administrations were completely analogous. Additionally, within the two experimental groups (dynamic and static goal), half of the teachers conducted testing themselves while the other half used computerized testing. A two-way analysis of covariance (computer vs. teacher administrations and dynamic vs. static goal, with pretreatment scores as the covariate) indicated no significant achievement effects for the administration factor or for the administration X goal interaction. Therefore, this administration factor is not discussed further in this paper.

Teachers collected 12 initial measurements (see teacher training below) to master the measurement procedures and to provide a training and acclimation period to students in the test-taking procedure, before using the assessment information for instructional decisionmaking. Teachers employed the last three measurement points to calculate a median baseline performance. Teachers next set a performance criterion, representing their best estimate of what the student might accomplish by year's end.

Then, each week, teachers employed data-management software (Fuchs et al., 1987) that automatically (a) stores and graphs the math assessment data, (b) applies a set of data-evaluation decision rules to the graphed performance, and (c) provides feedback to teachers to communicate those decisions. This data-management software displays a graph on the computer screen, showing (a) the pupil's performance over time, (b) an aimline reflecting the desired slope of improvement from baseline to goal, and (c) a regression line superimposed over the data points that have been collected since the last instructional change and extrapolated to the goal date. Figure 1 shows a sample computerized graph. Depending on the group to which teachers had been randomly assigned, the software applied either dynamic or static goal decision rules to the graphed data.

Insert Figure 1 about here

Dynamic goal decision rules. The dynamic goal decision rules guided the development of the student's instructional program in the following way.

First, 7 to 10 measurements were collected. Then, if the regression line was less steep than the aimline, the decision was that the teacher (a) introduce an instructional change to the pupil's program in an attempt to improve the rate of progress, and (b) collect 7 to 10 additional data points before reapplying the decision rules. If the regression line was as steep as the aimline, the decision was that the teacher (a) collect additional data points and (b) within the next 2 weeks, reapply the decision rules to the data that had been collected since the last intervention change. If the regression line was steeper than the aimline, the decision was for the teacher to (a) raise the goal to a higher level (this new goal criterion was predicted on the basis of the student's current rate of progress) and (b) collect 7 to 10 additional data points before reapplying the decision rules. Additionally, once the student's performance approached a criterion reflecting nearly maximum performance on the curriculum level, the decision was that the teacher conduct CBM on the next higher grade level of the math curriculum.

Static goal decision rules. The static goal decision rules guided instructional programming in the same manner as did the dynamic goal decision rules, with one exception. When the regression line projected that the student would surpass the goal criterion performance, the decision rule was to collect additional data points and, within the next 2 weeks, reapply the decision rules to the data that had been collected since the last intervention change. Therefore, although teachers were free to increase their goals whenever they deemed it appropriate and wished to do so, teachers were not directed to raise their goals when the slope of the regression line exceeded

that of the aimline.

The data-management software displayed the graph with regression line and aimline. Below the graph, on the same screen, a statement communicated the appropriate decision: "Oh-oh. Make a teaching change"; "OK! Collect more data"; "OK! Raise the goal to X" (only for dynamic goal group; X = projected goal); "Move to a higher grade level for measurement"; or "Insufficient data for analysis." (See example in Figure 1.)

Control. Control group teachers stated annual goals for their pupils on a standard form. Self-reports provided on an open-ended posttreatment questionnaire (Fuchs, 1987b) indicated these teachers monitored pupil progress toward goals using end-of-unit math test results, unsystematic observation of performance, and workbook performance for their database. Such descriptions match those reported by larger samples of practitioners (Fuchs, Fuchs, & Warren, 1982; Mirkin & Potter, 1982). Consequently, the control group served as a benchmark representing conventional practice in monitoring pupil progress toward goals.

Training

Teachers were trained to implement the CBM procedures in the following sequence.

1. Project staff visited each teacher for approximately 20 minutes. During these visits, teachers inspected math objectives lists and designated a BSF goal level for each participating pupil.

2. Teachers attended a 2-hour after-school workshop, at which time they received packets of materials for each pupil, including seven measurement

probes from the math level they had designated and measurement directions. Topics addressed at the workshop included a rationale for CBM and for computer applications to CBM, and methods of administering and scoring the bi-weekly progress monitoring probes.

3. Within the next month, teachers administered and scored these seven assessments for each pupil. Staff visited teachers to observe administrations and scorings. During the observation, project staff completed a checklist, indicating whether the teacher had implemented each component of the procedure correctly, and provided corrective feedback to the teacher as required. Staff continued to visit teachers, observe, and complete checklists until they had observed at least one completely correct administration and scoring.

4. Teachers attended a second 2-hour after-school workshop. At this session, they received information about and practiced using the data-management software.

5. In the next 3 weeks, teachers administered 5 additional probes and used the data-management software to inspect graphs at least twice.

6. Staff met individually with teachers to train them in their decision rules, and procedures for setting goals and completing forms on which they described their instructional programs. At this time, the 15-week study began.

7. Staff subsequently met with teachers in their classrooms, once every 2 weeks for 20 minutes. During these visits, staff inspected graphs with teachers, discussed pupil performance patterns, and assisted teachers in problem solving concerning treatment implementation. On average, during the

15-week study, teachers received 7 visits (range = 5 to 13).

Fidelity of Treatment

The accuracy with which teachers implemented the treatment was assessed using the Math Modified Accuracy of Implementation Rating Scale (M-MAIRS; Fuchs, 1987), which comprises three subscales: Initial Set Up (taking baseline, graphing data, writing goals, and drawing aimlines), Measurement (task administration, reliability of scoring, and frequency of measurement), and Data Utilization (entering information into the data-management system, describing instructional procedures, delineating instructional modifications, and timing instructional changes). (See Fuchs, 1987a for description of each item.) Each item is rated on a 5-point Likert-type scale (0=low; 4=high), in accordance with detailed scoring guidelines (see Fuchs 1987a for complete description of scoring guidelines). Staff were trained in scoring the MAIRS during one 3-hour session. Percentage of agreement (see formula above), calculated on 15% of the protocols, was 96.

To explore complementary aspects of treatment implementation, three other indices were summarized. The number of goal changes and number of intervention changes introduced by the teacher, following the specification of an initial goal and program, were counted by inspecting the data editor of the software, which recorded these changes. Finally, the final goal ratio (final goal level divided by baseline median) was calculated for each pupil. Interrater agreement on these indices was calculated for 20% of the students, with percentage of agreement (see formula above) of 98 and 100, respectively.

Data Collection

M-MAIRS data were collected for the two experimental groups (dynamic and static goal) 10 weeks into the study. One student per teacher was selected randomly as the target on whom M-MAIRS was assessed. The additional indices of treatment implementation were collected after treatment implementation. Achievement data on the MCT were collected for the the experimental and control groups immediately before and after the study: These data-collection periods are labeled below as "Time 1" and "Time 2." Project staff administered the MCT individually to students using a standard format. CN data were collected for the three groups immediately following treatment implementation, in groups of one to six pupils, in standard format.

Data Analysis

Since teachers, rather than students, had been assigned randomly to treatment groups, "teacher" was designated the statistical unit of analysis. On the M-MAIRS, since one student per teacher had been selected randomly for measurement, no data transformation was necessary. For number of goal changes, number of intervention changes, final goal ratio, and MCT and CN scores, measurements were aggregated across pupils for each teacher before data were analyzed.

The three subscales of the M-MAIRS were entered into a one-way multivariate analysis of variance (MANOVA). The number of goal changes, number of intervention changes, and final goal ratio were subjected to separate ANOVAs.

On the Time 1 MCT, mean performances for the dynamic goal CBM, static goal CBM, and control groups, respectively, were 33.20 (SD = 15.12), 41.75 (SD

= 20.90), and 47.55 (13.07). A one-way ANOVA applied to these pretreatment scores indicated that initial math achievement among the experimental groups was not reliably different, F (2,27) = 1.87, p = .17. However, given the raw score differences among groups, analyses of covariance (ANCOVAs) were employed for subsequent analysis of achievement data.

The Time 2 MCT and the CN achievement were subjected to a multivariate one-way ANCOVA, which controlled for the Time 1 MCT achievement levels. Before applying ANCOVAs, the assumptions of homogeneous regression coefficients were tested and scattergrams were inspected; the assumptions of equal regression slopes and linearity of Y on X appeared tenable. Follow-up univariate ANCOVAs and Scheffe tests were employed, as required.

Results

Fidelity of Treatment

The one-way MANOVA was conducted on the M-MAIRS subscales. Wilks' lambda criterion was used to test for equality of group centroids, and the value calculated with the Wilks' lambda procedure were transformed into an F value through Rao's approximation (Cooley & Lohnes, 1962). The test for lambda revealed no significant differences associated with the treatment factor, F (3,16) = 1.02, ns (Wilks' lambda = .840). Univariate F (1,18) values for the Measurement Set Up, Measurement, and Data Evaluation subscales, respectively, were 1.00, .23, and 1.75. Descriptive statistics for the experimental groups are shown in Table 2.

Insert Table 2 about here

Descriptive statistics for the additional indices of treatment implementation also are displayed in Table 2. ANOVAs yielded the following statistical values: F (1,18) = 12.81, $p < .01$ for number of goal changes, F (1,18) = 2.40, $p = .14$ for number of intervention changes, and F (1,18) = 9.42, $p < .01$ for final goal ratio. Inspection of the means indicates the dynamic goal group introduced a greater number of goal changes and utilized a higher final goal ratio than the static goal group.

Achievement

The MANCOVA on the MCT and CN revealed a significant effect, F (4,48) = 2.71, $p < .05$ (Wilks' lambda = .675). Scores on the math achievement tests are shown in Table 2. The ANCOVA on the MCT produced a significant effect, F (2,26) = 4.67, $p < .05$. Follow-up Scheffe tests indicated the dyanamic goal CBM group's adjusted achievement level was greater than that of the controls; no other differences were statistically significant. The statistically significant effect for the dynamic goal vs. control groups on the MCT was associated with an effect size magnitude of .52 (where effect size magnitude derived from ANCOVA = adjusted experimental mean minus adjusted control mean, divided by the quantity $MS_W [df_W - 1]/[1 - r^2_{xy}][df - 2]$, see Glass, McGaw, & Smith, 1983). Other effect size magnitudes were: .25 for the contrast between the static goal CBM and control groups and .28 for the contrast between the dynamic and static goal CBM groups.

The ANCOVA on the CN indicated no reliable differences among groups, F (2,26) = 1.63, $p = .21$. However, contrary to the hypothesis that predicted

that achievement for the control group on this content coverage measure would exceed that of the experimental groups (see Slavin, 1987 for discussion), the direction of the means favored CBM . The effect size magnitudes (see formula above) for the dynamic goal vs. the control groups was .39; for the static goal vs. the control groups, .35.

Discussion

As indicated on the M-MAIRS measure, the CBM teachers complied with the overall CBM procedures with comparable accuracy. Nevertheless, findings on two additional fidelity of treatment indices revealed that their methods of using goals to structure their measurement and instructional planning activities did differ. The dynamic goal CBM teachers, whose decision rules required them to increase goals whenever actual progress exceeded the anticipated rate of improvement, increased goals more frequently than did the static goal CBM teachers, who were always free, but never required, to increase goals. Moreover, by the study's completion, the goals of the dynamic goal group were more ambitious than those of the static goal group. These accuracy of implementation findings document the integrity of the experimental treatment: CBM teachers employed the monitoring procedures as required, and the goal-setting behaviors of teachers in the contrasting CBM groups differed and conformed to the specified treatments.

Concurrent with teachers' use of alternative goal-setting processes, the CBM procedures resulted in differential student achievement. Students in the dynamic goal CBM group achieved better than the controls on the MCT, the measure of content mastery, while the achievement of the static goal CBM group

did not exceed that of the controls. The effect size magnitude associated with the dynamic goal CBM procedures was .52, or approximately one-half standard deviation. This indicates that, in terms of the standard normal curve and an achievement test scale with a population mean of 100 and standard deviation of 15, one might expect the use CBM with dynamic goals to increase the typical achievement outcome score from 100 to approximately 107.5. This finding supports previous research in psychology indicating that adults in work settings perform better with difficult goals (Locke, Shaw, Saari, & Latham, 1981). Additionally, the current study corroborates a post-hoc special education analysis (Fuchs et al., 1985), where teachers who employed more difficult CBM goals effected better student achievement outcomes.

The current study contributes to this previous literature in at least five important ways. First, within a controlled investigation, it documents the importance of ambitious goals for use by teachers in their assessment and instructional planning with youngsters. Second, this study suggests that, despite the importance of ambitious goals, educators' typical goal-setting standards may underestimate many students' potential: The study procedures allowed teachers to establish their initial goals freely, in line with the progress rates they anticipated. However, with these initial goals, teachers in the dynamic goal group were required to increase goals for an average of more than one out of every two pupils. This goal increasing behavior was prompted by students' exceeding teachers' initial goal expectations. Third, in addition to demonstrating teachers' goals may underestimate potential progress rates, the current study indicates that, without systematic prompting

to raise goals, practitioners cannot be expected to do so: Among the 20 student participating in the static goal group, there was only one instance of a teacher raising a goal. Fourth, with more ambitious goals, the sample of dynamic goal teachers in this study introduced more intervention changes. Seemingly, their more difficult goals required intervention changes to stimulate better growth rates. Although this difference, favoring the number of interventions for dynamic goals, was not reliable (p = .14), this trend suggests that additional research exploring the relation between ambitious goals and teachers' instructional behavior and programmatic development may be warranted.

A fifth and important way in which this study contributes to the related literature is by providing an example of a workable methodology the education community might employ for empirically deriving ambitious, but realistic, goals. A persistent problem for education has been that it is difficult, if not impossible, to anticipate the scope of attainable, but ambitious, goals. The current study provides a process by which goals can be developed dynamically, so that progress toward mastery is monitored closely and goals are adjusted upward whenever possible. Given (a) the finding that such goal adjustment, specifically, and goal ambitiousness, generally, enhances student achievement, along with (b) the relative ease and efficacy of CBM with automatic data collection and management systems (see Fuchs, Hamlett, Fuchs, Stecker, & Ferguson, in press), teachers might consider adoption of CBM systems that incorporate dynamic goal-setting procedures.

A second purpose of the current study was to assess the extent to which

CBM affected not only a measure of content mastery, but also an index of content coverage. As proposed by critics of the mastery learning field, a close link between measurement and instruction may result in better content mastery of the measured domain, but may decrease content coverage, or the extent to which teachers address skills related to, but not synomymous with, the measured domain (Slavin, 1987). To investigate this hypothesis, the current investigation employed a computation test as a measure of content mastery and a number concepts test as a measure of content coverage. Although findings favored the growth of the dynamic goal CBM over the control group on the content computation mastery measure, no significant differences were found on the number concepts mastery coverage index. Nevertheless, the hypothesis suggests that the control students' achievement on the mastery coverage test should have exceeded that of the CBM groups. In fact, even though number concepts were not the focus of the CBM measurement domain, the effect size magnitude on the number concepts tests favored students in the CBM groups, with a mean magnitude of .37. This suggests that that (a) CBM teachers did cover a curriculum broader than that encompassed in the CBM procedures, and/or (b) students who grew in computation concurrently progressed or generalized their newly acquired skills to a broader curricular scope. Either way, results do not support the contention that a close connection between measurement and instruction leads to more limited achievement on content coverage measures.

References

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review, 84, 191-215.

Bandura, A (1978). Reflections on self-efficacy. Advances in Behavioral Research and Theory, 1, 237-269.

Bandura, A. (1982). Self-efficacy mechanism in human agency. American Psychologist, 37, 122-147.

Cooley, W.W., & Lohnes, R.R. (1962). Multivariate procedures for behavioral sciences. New York: John Wiley & Sons.

Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. Exceptional Children, 52, 219-232.

Deno, S.L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. School Psychology Review, 15, 358-374.

Deno, S.L. (1987). Curriculum-based measurement. Teaching Exceptional Children, 20(1), 41-42.

Deno, S.L., & Fuchs, L.S. (1987). Developing curriculum-based measurement systems for data-based special education problem solving. Focus on Exceptional Children, 19(8), 1-16.

Fuchs, L.S. (1986a). Use of goals. [Spencer Foundation Seed Grant]. (Available from L.S. Fuchs, Box 328, Peabody College, Vanderbilt

University, Nashville, TN 37203.)

Fuchs, L.S. (1986b). Monitoring the performance of mildly handicapped students: Review of current practice and research. Remedial and Special Education, 7, 5-12.

Fuchs, L.S. (1987a). Improving data-based instruction through computer technology: Continuation application. (Available from L.S. Fuchs, Peabody College, Vanderbilt University, Nashville, TN.)

Fuchs, L.S. (1987b). Instructional planning with curriculum-based measurement. Unpublished manuscript. (Available from L.S. Fuchs, Peabody College of Vanderbilt University, Nashville, TN 37203.)

Fuchs, L.S., Deno, S.L., & Mirkin, P.K. (1984). Effects of frequent curriculum-based measurement of evaluation on pedagogy, student achievement, and student awareness of learning. American Educational Research Journal, 21, 449-460.

Fuchs, L.S., Fuchs, D., & Deno, S.L. (1985). The importance of goal ambitiousness and goal mastery to student achievement. Exceptional Children, 52, 63-71.

Fuchs, L.S., Fuchs, D., & Warren, L.M. (1982). Teacher practice in evaluating student progress toward goals (Research Report No. 81). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities. (ERIC Document Reproduction Service No. ED 224 197)

Fuchs, L.S., Hamlett, C., Fuchs, D. (1987). Computer applications to curriculum-based measurement. Unpublished computer program, Nashville, TN.

Fuchs, L.S., Hamlett, C., Fuchs, D., Stecker, P.M., & Ferguson, C. (in press).

Conducting curriculum-based measurement with computerized data collection: Effects on efficiency and teacher satisfaction. Journal of Special Education Technology.

Gardner, E.F., Rudman, H.C., Karlsen, B., & Merwin, J.C. (1982). Stanford achievement test. Psychological Corporation.

Gibson, S. & Dembo, M.H. (1984). Teacher efficacy: A construct validation. Journal of Educational Psychology, 76, 569-582.

Gickling, E.E. (1981). The forgotten learner. Nevada Public Affairs Review, 1, 19-22.

Glass, G.V., McGaw, B., & Smith, M.L. (1981). Meta-analysis in social research. Beverly Hills: Sage.

Locke, E.A., Shaw, K.N., Saari, L.M., & Latham, G.P (1981). Goal-setting and task performance: 1969-1980. Psychological Bulletin, 90, 125-152.

Mirkin, P.K., & Potter, M.L. (1982). A survey of program planning and implementation practices of teachers (Research Report No. 80). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.

Office of the Federal Register (1986). Code of federal regulations, Parts 300 - 399. Washington, DC: Author.

Slavin, R.E. (1987). Mastery learning reconsidered. Review of Educational Research, 57, 175-213.

Thompson, R.H., White, K.R., & Morgan, D.P. (1982). Teacher-student interaction patterns in classrooms with mainstreamed mildly handicapped students. American Educational Research Journal, 19, 220-236.

Tucker, J. (1985). Curriculum-based assessment: An introduction. Exceptional Children, 52, 199-204.

Yalow, E., & Popham, J. (1983). Content validity at the crossroads. Educational Researcher, 12(8), 10-14, 21.

Figure Captions

Figure 1. Example of a computerized graph.

Table 1

Teacher and Student Demographic Variables

Variable	Dynamic Goal $\bar{X}$	Dynamic Goal SD	Static Goal $\bar{X}$	Static Goal SD	Control $\bar{X}$	Control SD	F[a]
Teacher							
Age[b]	2.7	.82	2.5	.53	2.6	1.17	.13
Years Teaching	8.1	4.38	9.1	4.70	9.8	5.88	.29
Years Computer Use	1.4	1.51	1.2	1.23	1.1	1.60	.11
Years in School	3.8	3.05	5.6	2.99	4.6	2.41	1.01
Evaluation	20.2	15.95	24.1	23.78	20.7	16.45	.12
Personal Efficacy	4.8	.65	4.6	.50	4.8	.37	.55
General Efficacy	4.8	.56	4.7	.29	4.7	.47	.39
Student							
Age	10.9	1.51	11.7	2.13	11.9	1.55	1.39
Grade	5.0	1.05	4.4	1.94	5.3	1.21	1.72
Math Grade	3.2	1.36	3.2	1.13	3.7	1.10	1.96

[a]None of these F values was statistically significant. For teacher variables, dfs = 2,27; for student variables, dfs = 2,57.

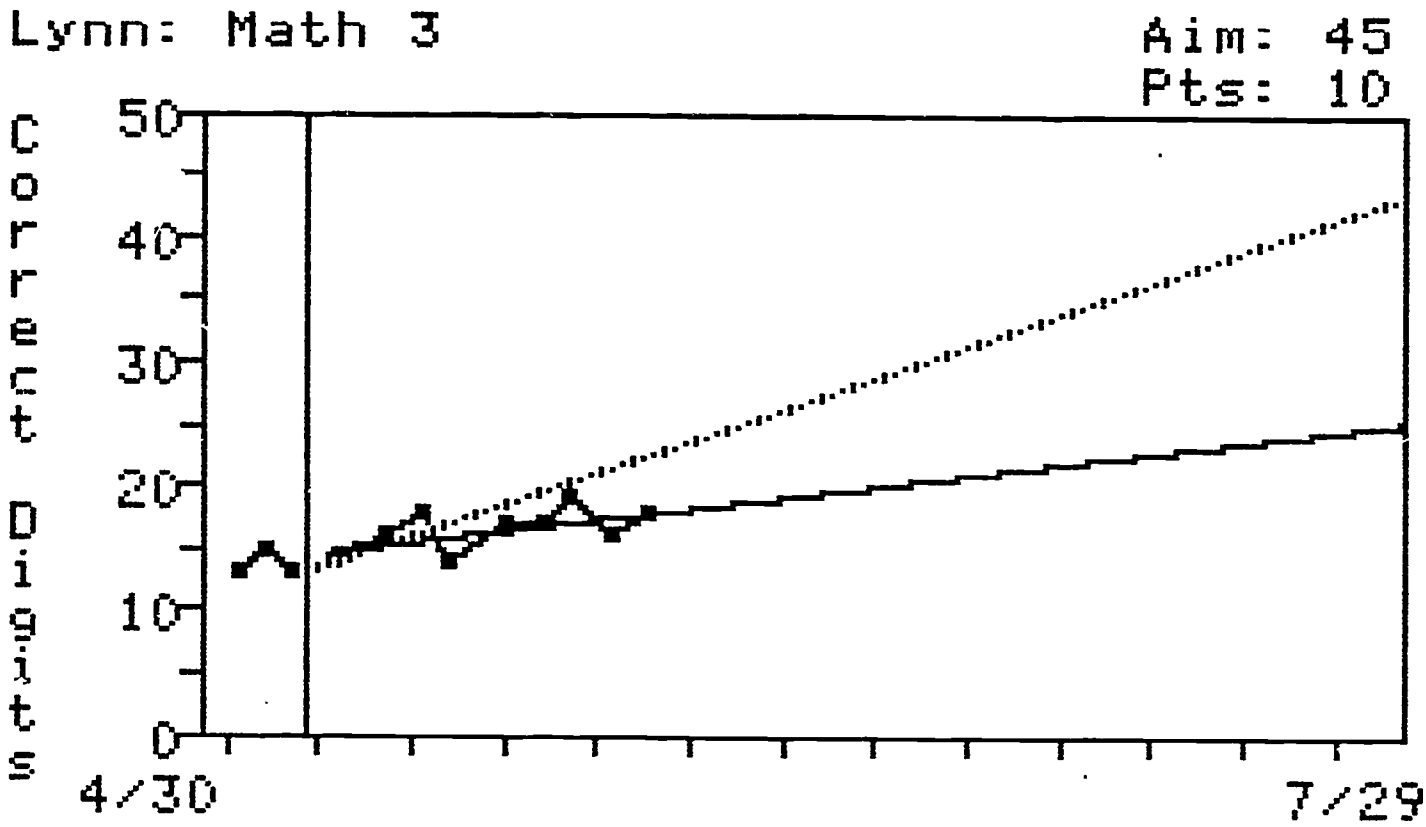
Table 2

Fidelity of Treatment and Achievement Scores by Groups

	Dynamic Goal		Static Goal		Control	
Variable	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Fidelity of Treatment						
MAIRS - Set Up	4.00	.00	3.98	.06	--	--
MAIRS - Measurement	3.67	.36	3.59	.39	--	--
MAIRS - Evaluation	3.52	.34	3.22	.63	--	--
Number Goal Changes	.60	.46	.05	.16	--	--
Number Intervention Changes	1.80	.68	1.40	.46	--	--
Final Goal Ratio	3.56	1.31	2.20	.49	--	--
Achievement						
Time 1 MCT	33.20	15.12	41.75	20.78	47.55	13.06
Time 2 MCT	43.05	15.23	46.00	20.24	50.00	12.74
Adjusted MCT[a]	50.82	--	46.21	--	42.02	--
Adjust CN[b]	576.83	--	575.16	--	556.96	--

[a] MCT is Math Computation Test Time 2 scores, adjusted using the Time 1 MCT covariate.

[b] CN is Number Concept score, adjusted using the Time 1 MCT as a covariate.



Uh-oh! Make a teaching change.