DOCUMENT RESUME

ED 404 105 SE 053 768

AUTHOR	Watts, Michelle R.
TITLE	The Implementation of the 1989 Assessment Standards for School Mathematics in Grades K-3.
INSTITUTION	Florida Educational Research and Development Council, Inc., Sanibel.
PUB DATE	93
NOTE	96p.
AVAILABLE FROM	Florida Educational Research Council, Inc., P.O. Box 506, Sanibel, FL 33957 ($3.50 individual copies; $12 for an annual subscription).
PUB TYPE	Collected Works - Serials (022)
JOURNAL CIT	Florida Educational Research Council Research Bulletin; v24 n4 Sum 1993
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	Elementary School Teachers; *Inservice Teacher Education; Mathematics Achievement; Mathematics Education; *Mathematics Tests; Primary Education; *Student Evaluation; Teaching Methods
IDENTIFIERS	*Alternative Assessment; Alternative Grading; *NCTM Curriculum and Evaluation Standards

ABSTRACT

Thirty-three primary teachers were trained in the 1989 Standards for School Mathematics to study the relationship between the implementation of alternative assessment and such variables as: teacher grading orientation, class size, grade level, and student mathematical ability. Teachers were found to use knowledge level items significantly more often than higher level items, and K-1 teachers used items involving manipulative materials significantly more often than grade 2-3 teachers. Patterns of usage by question level, assessment format, manipulative material, and scoring method did not vary according to teacher variables. Mathematical procedures were identified as more appropriately measured by traditional methods; concepts, problem solving, and reasoning by alternative methods. No significant difference in the level of confidence in the information from traditional versus alternative assessment methods was found, but alternative formats were more difficult to use. Results led to recommendations such as: (1) include alternative assessment instruments with mathematics textbooks; (2) provide more manipulative materials and inservice training to teachers; (3) train teachers to use and understand analytic and focused holistic scoring, especially for problem solving; and (4) monitor teachers over time to see if use of alternative assessment increases and if it affects student achievement. (Contains 61 references.) (Author/GW)

FLORIDA EDUCATIONAL
RESEARCH COUNCIL

RESEARCH BULLETIN

The Implementation of the 1989 Assessment Standards for School Mathematics in Grades K-3

by
Michelle R. Watts

Volume 24 | Summer 1993 | Number 4

FLORIDA EDUCATIONAL
RESEARCH COUNCIL, INC.

RESEARCH BULLETIN

The Implementation of the 1989 Assessment Standards for School Mathematics in Grades K-3

by
Michelle R. Watts
Hillsborough CountyPublic Schools

Additional copies of this book may be obtained from:
FERC, Inc.
P.O. Box 506
Sanibel, Florida 33957

Individual Copies $3.50
Annual Subscription $12.00

10% Discount on an order of 5 or more.
Order of less than $20.00 must be accompanied by
a check or money order.

F.E.R.C. NOTES ON THIS BULLETIN

Assessment is certainly a constant in the minds of those persons who believe in measuring, as clearly as possible, what teachers have taught and students have learned as a result of attending school. This study of the implementation of the 1989 Assessments for School Mathematics in Grades K-3 and as such is valuable not only to assessors, evaluators and other measurement specialists, but also to the classroom teachers in the early primary grades in the field of mathematics. F.E.R.C. offers this research for its members and other interested parties.

Charlie T. Council
Executive Director

TABLE OF CONTENTS

	Page
List of Tables	iv
List of Figures	vi
Executive Summary	vii
Introduction	1
Statement of Problem	2
Literature	3
1989 Standards for School Mathematics	3
Mathematics Problem Solving with Manipulative Materials	6
Achievement Assessment	9
Standardized Tests - Strengths and Weaknesses	9
Teacher-made Tests	12
Alternative Testing Formats	13
Methodology	17
Sample	17
Test Item Classification Variables	19
Teacher Classification Variables	20
Data Analysis	22
Results	24

TABLE OF CONTENTS (cont.)

Implementation of Alternative Assessment According to Teacher Grading Orientation, Class Size, Grade Level, and Student Achievement Level 24

Assessment Standards by Level of Question, Assessment Format, Manipulative Materials, and Method of Scoring 38

Evaluation Information Gained by Alternative and Traditional Assessment Strategies 51

Conclusions 63

Implementation of Alternative Assessment According to Teacher Grading Orientation, Class Size, Grade Level, and Student Achievement Level 63

Assessment Standards by Level of Question, Assessment Format, Manipulative Materials, and Method of Scoring 66

Evaluation Information Gained by Alternative and Traditional Assessment Strategies 69

Recommendations 70

References 72

Appendices
1 Grading Orientation Questionnaire 76
2 Assessment Questionnaire 77

LIST OF TABLES

Page

Table 1 Teacher Education Level and Years of Experience

Table 2 Comparison of Item Levels for the K-3 Mathematics Specialist Teachers by Grading Orientation, Class Size, Grade Level, and Student Achievement Level

Table 3 Comparison of Items Using Alternative and Traditional Assessment Formats for the K-3 Mathematics Specialist Teachers by Grading Orientation, Class Size, Grade Level, and Student Achievement Level

Table 4 Mean Proportion of Assessment Formats by Teacher Grading Orientation, Class Size, Grade Level, and Student Mathematics Achievement

Table 5 Comparison of Items Using and Not Using Manipulative Materials for the K-3 Mathematics Specialist Teachers by Grading Orientation, Class Size, Grade Level, and Student Achievement Level

Table 6 Mean Proportion of Items Using Manipulative Materials by Teacher Grading Orientation, Class Size, Grade Level, and Student Mathematics Achievement

Table 7 Comparison of Items Using Alternative and Traditional Scoring for the K-3 Mathematics Specialist Teachers by Grading Orientation, Class Size, Grade Level, and Student Achievement Level

Table 8 Mean Proportion of Scoring Methods by Teacher Grading Orientation, Class Size, Grade Level, and Student Mathematics Achievement

LIST OF TABLES (cont.)

Page

Table 9 Mean Proportion of Content Measured by Teacher Grading Orientation, Class Size, Grade Level, and Student Mathematics Achievement

Table 10 Proportion of Items by Level of Question and Assessment Standard

Table 11 Mean Proportion of Test Items by Assessment Standard and Level of Question

Table 12 Proportion of Alternative and Traditional Assessment Format Items by Assessment Standard

Table 13 Mean Proportion of Test Items by Assessment Standard and Assessment Format

Table 14 Proportion of Items Using and Not Using Manipulative Materials by Assessment Standard

Table 15 Mean Proportion of Test Items by Assessment Standard and Manipulative Materials

Table 16 Proportion of Items Using Alternative and Traditional Scoring Methods by Assessment Standard

Table 17 Mean Proportion of Test Items by Assessment Standard and Method of Scoring

Table 18 Mean Proportion of Test Items by Assessment Standard and Content

Table 19 Content Areas Better Matched to Alternative Assessment Formats

Table 20 Content Areas Better Matched to Unit Tests

LIST OF TABLES (cont.)

Page

Table 21 Frequency and Percent of Teachers: Appropriate Assessment Procedures for each Assessment Standard

Table 22 Confidence in the Evaluation Information Gained through Alternative Assessment Formats and Unit Tests

Table 23 Comparison of the Confidence of Information from Alternative Assessment Formats and Unit Tests

Table 24 Frequency of Using Alternative Assessment Formats and Unit Tests

Table 25 Comparison of the Frequency of Using Alternative Assessment Formats and Unit Tests

Table 26 Difficulty Using Alternative Assessment Formats and Unit Tests

Table 27 Comparison of the Difficulty Using Alternative Assessment Formats and Unit Tests

LIST OF FIGURES

Page

Figure 1 Assessment Matrix

EXECUTIVE SUMMARY

This study was designed to describe the mathematics assessment procedures used by a group of teachers specifically trained in the 1989 Standards for School Mathematics and to look at the relationship between the implementation of alternative assessment and such variables as teacher grading orientation, class size, grade level, and student mathematics ability. Also, the study investigated the differences in evaluative information gained by alternative assessment strategies (e.g., demonstrations) as compared to traditional assessment techniques (e.g., multiple choice).

The mathematics assessment instruments of 33 kindergarten through third grade teachers were analyzed to describe their assessment procedures. The teachers were found to use knowledge level items significantly more frequently than higher level items and to use items with manipulative materials significantly more frequently than items without manipulatives. The kindergarten and first grade teachers used manipulative materials significantly more frequently than did the second and third grade teachers. Significant differences were not found in the use of alternative formats and alternative scoring methods. Patterns of usage by question level, assessment format, manipulative material, and scoring method did not vary according to the teacher variables.

The particular standards the teacher assessed were found to be factors in the assessment practices the teachers chose. Mathematics Procedures were identified by the teachers as more appropriately measured with traditional procedures while Mathematical Power, Concepts, Disposition, Problem Solving. Communications, and Reasoning were identified as more appropriately measured with alternative assessment procedures. There were no statistically significant differences in the level of confidence in the evaluation information from the two assessment approaches, but alternative assessment formats were found to be significantly more difficult to use.

THE IMPLEMENTATION OF THE 1989 ASSESSMENT STANDARDS FOR SCHOOL MATHEMATICS IN GRADES K-3

INTRODUCTION

The results of the National Assessment of Educational Progress have been published since 1969 and indicate that American students are outperformed in mathematics, science, and reading comprehension by students in other industrial societies (Presseisen, 1986). The 1986 National Assessment of Educational Progress results indicate that students do not understand the underlying concepts of mathematics and are unable to apply knowledge in problem solving situations (Kouba, Brown, Carpenter, Lindquist, Silver, & Swafford, 1988). Since 1973, there have been gains, but the improvement has been in lower level skills and basic concepts (Ashworth, 1990). Educators, made aware of lowered educational success when compared with other industrial societies, are proposing intervention activities to counteract this trend (Sternberg & Baron, 1985) such as the instruction and assessment of thinking and problem solving (Presseisen, 1986).

The field of mathematics education has responded to this emphasis on problem solving and the assessment of higher cognitive skills by developing the 1989 Standards for School mathematics (Thompson & Rathmell, 1988). Published by the National Council of Teachers of Mathematics, the standards cover both curriculum and evaluation standards for K-12 school mathematics. The standards encourage the assessment of all aspects of mathematics and suggest that a variety of formats is necessary to fully assess mathematics.

Most student assessment occurring in the classroom with teacher-made tests has been found to be focused at the knowledge level (Mehrens & Lehmann, 1987). With the emphasis on problem solving and reasoning in the 1989 Standards, these current testing practices may not assess the skills taught in the classroom. The suggestion by the 1989 Standards for School Mathematics to use a variety of testing formats have led to discussions over appropriate formats.

Clark, Clark, and Lovitt (1990) reported that the restructuring of the goals and practices of mathematics education must be accompanied by assessment strategies which reflect the new conception of mathematics. Since concepts considered valuable are communicated to students through testing, the assessment must be comprehensive; therefore, Clark, Clark, and Lovitt have suggested that assessment tools must be sensitive to process, as well as product, and that teachers must expand their repertoire of assessment strategies beyond paper and pencil tests. Formats such as observations and checklists, interviews, oral questioning, and portfolios, with alternative scoring approaches and the use of calculators, have received support (Charles, Lester, & O'Daffer, 1988; Clark, Clark, & Lovitt, 1990; Guthrie, 1984; Guthrie & Lissitz, 1985; Haney, 1985; Mehrens & Lehmann, 1987; Peck, Jencks, & Connell, 1989; Robinson, 1987; Silver, 1990; and Webb & Briars, 1990).

In order to implement the 1989 Standards for School Mathematics, Hillsborough County Public Schools developed the K-3 Mathematics Specialist Project. A selected group of teachers received inservice training focused on increasing the teachers' knowledge of mathematics concepts, the integration of manipulative materials into the content areas, teaching methods, and alternative assessment techniques. During the 1990-91 school year, these teachers implemented the 1989 Standards for School Mathematics teaching strategies and assessment methods using the district curriculum.

STATEMENT OF PROBLEM

The primary focus of this study is to describe the manner in which classroom teachers implemented the assessment component of the 1989 Standards for School Mathematics. The assessment methods which were used by K-3 teachers trained in alternative assessment were described. Educational characteristics were studied as related to the assessment method used, and differences in the assessment method employed for each of the seven evaluation standards was examined. A comparison of the application of traditional assessment methods and alternative assessment methods was conducted to determine if differences exist in the evaluative information generated by assessment type. Specifically, the following research questions were addressed:

1. Do K-3 teachers vary after training in the methods of implementing alternative assessment strategies depending on teacher grading orientation, class size, grade level, and student mathematics ability?
2. Do K-3 teachers vary after training in the methods of implementing alternative assessment strategies by assessment standard? The assessment standards of the 1989 Standards for School Mathematics include Mathematical Power, Problem Solving, Communications, Reasoning, Mathematical Concepts, Mathematical Procedures, and Mathematical Disposition.
3. Is there a difference in the evaluative information gained by an application of alternative assessment strategies when compared with traditional assessment techniques?

LITERATURE

The literature that provided the framework for this study is reviewed in three sections and includes theoretical and descriptive literature, as well as empirical studies. The first section focuses on the development and impact of the 1989 Standards for School Mathematics. Mathematics problem solving and the use of manipulatives are covered in the second section while the literature on achievement assessment is addressed in the third section.

1989 STANDARDS FOR SCHOOL MATHEMATICS

The 1989 Standards for School Mathematics, the product of five years of planning and development, are intended to prepare students for the 21st century. These Standards have been welcomed by leading mathematics educators with journal articles describing the changes that will occur as a result of the Standards. The suggested changes cover all areas of mathematics education including curriculum, instructional methods, assessment, and teacher education.

The 1989 Standards for School Mathematics represent the first attempt by a teacher organization to develop national professional standards for school curricula (Crosswhite, Dossey, & Frye, 1989) and have been endorsed by 15 national mathematics associations and by 25 national education associations. The Standards were developed by classroom teachers, supervisors, teacher educators, and grade level experts to be realistic and achievable for the classroom teacher. Representing a consensus of the mathematics education community, the Standards describe what students should be able to do as a result of mathematics education and are expected to have a major influence on curriculum and local, state, and national testing (Thompson & Rathmell, 1988). Described below are issues of particular interest to this study that are currently discussed in the mathematics literature concerning the 1989 Standards for School Mathematics. The pertinent issues include the effect on curriculum, instructional methods, assessment practice, and teacher education.

The curriculum issues of the Standards including decision making, curriculum emphasis, and mathematics concepts have been summarized by Romberg (1988) and Thompson and Rathmell (1988). The purposes of the Standards were to ensure quality, indicate goals, promote change, and reflect current applications of mathematics in the curriculum as a result of the influence of technology (Romberg, 1988). Romberg has argued that the responsibility for curriculum and evaluation has been surrendered to legislators, administrators, textbook publishers, and test publishers. With the development of the Standards curriculum decisions, according to Romberg, would now fall within the responsibility of the National Council of Teachers of Mathematics.

Classroom teachers have always had the ability to select or emphasize content areas of the curriculum; therefore, the development of Standards may not necessarily change the decision maker. The assumption that test publishers control curriculum has frequently been made and was implied by Romberg (1988). Worthen and Spandel (1991) indicate that test publishers developed tests around the content of textbooks and curricular materials, but the manner in which these materials have been used (i.e., controlling curriculum) is not a function of the test but rather one of the user. In other words, the development of the Standards may not have changed the manner in which decisions concerning the mathematics curriculum are made.

An additional curriculum issue included in the Standards addressed the emphasis of mathematics education. This consists of new technology capabilities, such as calculators and computers, a focus on problem solving, the representation (communication) of mathematical ideas, and reasoning. Thompson and Rathmell (1988) discussed the manner in which the above ideas were addressed in the 1989 Standards for School Mathematics. The impact of technology included the use of calculators and computers for classroom mathematics. Problem solving in mathematics education should provide a context for meaningful learning of concepts and skills, which would foster the development of higher order thinking. By focusing on the use of different modes to construct meaning in mathematics, the student would develop mathematics communication. The reasoning area emphasized logical reasoning in mathematics, with students developing either deductive or inductive conclusions.

Additional curriculum issues of the standards included the placement of some concepts at a higher grade level than previously taught and the inclusion of new mathematical topics (Thompson & Rathmell, 1988). For example, place value was moved to second grade and basic addition and subtraction were delayed until grade three. When teachers introduce these concepts before students are developmentally ready, the students rely on rote memory rather than developing an understanding of the concept (Thompson & Rathmell, 1988). New mathematical topics include number and spacial sense, beginning ideas of statistics, and probability. There is also a greater emphasis on measurement, geometry, and estimation (Thompson & Rathmell, 1988).

The effect of the 1989 Standards on instructional methods have been described by DeMana and Waits, 1990; Dossey, 1989; and Thompson and Rathmell, 1988. The Standards outline goals that schools should strive to reach and suggest instructional methods to assist in reaching them Dossey (1989). The Standards suggest changing the student's role from one of passive receptivity to one of active involvement with the support of technology. Classroom activities suggested in the Standards included the extensive and thoughtful use of physical materials to foster the learning of abstract ideas, with students discussing and writing about their results (Dossey, 1989). The standards were developed with the assumption that kindergarten through fourth grade classrooms need a wide variety of physical materials and supplies including mathematics manipulatives and simple household objects (National Council of Teachers of Mathematics, 1989).

DeMana and Waits (1990) have suggested that the instructional focus for mathematics procedures should include setting up problems with the appropriate operations rather than computations. The expectation for paper and pencil computation proficiency has drastically changed with the Standards. Students must still be computationally proficient, but simple computations were recommended, with complex computations undertaken with calculators (Thompson & Rathmell, 1988). Technology has reduced the time needed for drill and practice and has made the completion of computational problems using pencil and paper manipulations obsolete (DeMana & Waits, 1990). As a result of these changes, more classroom time should be available to develop mathematics concepts.

The 1989 Standards for School Mathematics also address student assessment. Mathematics assessment, as described by the 1989 Standards, would change to include a greater emphasis on observations, interviews, student journals, and formats capable of disclosing student misconceptions missed by traditional assessment. Dossey (1989) suggested that success in changing assessment methods would lead to opportunities to strengthen teaching methods. Additional information about assessment is located in a later section.

The 1989 Standards for School Mathematics, as described, represent a change in mathematics education. Not only will curriculum, instruction, and assessment change, but teacher training will also be impacted. In order for mathematics education to change, the reform must include a change in teachers' concepts of mathematics teaching. Teachers need to learn mathematics in the same way they are encouraged to teach by the 1989 Standards for School Mathematics. Cooney (1988) suggested that teacher inservice training should model the teaching strategies to facilitate the implementation of the Standards by classroom teachers.

With the introduction of the 1989 Standards, mathematics educators have described their potential impact on all aspects of mathematics education. All areas of mathematics education may potentially be affected by the 1989 Standards. Change in the classroom, including curriculum, instructional methodology, materials, and assessment would be substantial, if the intent of the Standards is realized. Crosswhite, Dossey, and Frye (1989), have indicated that the Standards must be complemented with instructional methodologies, teacher education, texts, assessment materials, and prototypal instructional materials. Mathematics education research, for the near future, will likely focus on issues pertinent to the 1989 Standards for School Mathematics (National Council of Teachers of Mathematics, 1989).

MATHEMATICS PROBLEM SOLVING WITH MANIPULATIVE MATERIALS

The 1989 Standards for School Mathematics encourage problem solving and the use of manipulative materials as a focus both in the curriculum and assessment. A review of the literature on problem solving and mathematics manipulative materials is provided to assist the reader in understanding the role of problem solving and manipulative materials in the Standards and in alternative assessment.

Problem solving has received more attention from researchers than other topics in mathematics education (Kameenui & Griffin, 1989). The findings of the National Assessment of Educational Progress indicated a lack of problem solving skills in students yet, Kameenui and Griffin (1989) reported that students, upon entering school, had the ability to solve problems without an understanding of the mathematic concepts or operations. This finding was supported by Carpenter (1985); Loif, Carey, Carpenter, and Fennema (1988); and Wearne and Hiebert (1985). They found that children demonstrated mathematics problem solving abilities before formal schooling. Therefore, children entered school with abilities to solve problems, but as formal instruction was provided, this ability diminished.

In addition, children were found by Carpenter (1985) to invent solutions to problems using relatively sophisticated problem-solving strategies prior to formal instruction. Despite their early problem solving skills, Wearne and Hiebert (1985) reported that as students memorized rules which were not understood, the real world and school mathematics were separated in their minds. These authors found that as grade level increased, students did not think about what they were doing when solving problems; they thought of a rule they have memorized and applied it.

A possible explanation for the lack of higher order thinking skills in students has been offered by Grice and Jones (1989) and Quellmalz (1985). Evidence of effective school instruction that focused on thinking has been absent (Grice & Jones, 1989). Similarly, Quellmalz (1985) reported that students were seldom asked to engage in sustained reasoning or to explain their reasoning. Grice and Jones have suggested that the abstractness of reasoning has been a problem for the educational community while the concrete product of the process was manageable.

Stanic and Kilpatric (1988) stated that while problems have been included in the mathematics curriculum, the solving of these problems is a new area of instructional concern. As a result of this emphasis, instructional approaches to problem solving have been offered. Three such approaches have been described by Campbell and Bamberger (1990): (a) teaching about problem solving (instruction in the strategies for problem solving); (b) teaching for problem solving (the application of mathematics); and (c) teaching via problem solving (using problems to teach mathematics concepts and computation). The 1989 Standards for School Mathematics do not suggest any one approach to teaching problem solving. Campbell and Bamberger (1990) feel that the Standards include all three approaches.

The first approach, teaching about problem solving, has been discussed by other mathematics educators. For example, Stiff (1988) suggested teaching students about problem solving through group activities. He maintained that it could not be assumed that students would develop problem solving abilities in mathematics without instruction that focused on the steps of problem solving during the process.

Children have been found to use different strategies to solve problems. Siegler (1989) found that the same child may have used different strategies to solve the same problem. Thus, he proposed that different problem solving strategies should be taught and that educators need to understand the ways students solve problems in order to model alternative approaches.

The second approach, teaching for problem solving, is encouraged by the 1989 Standards for School Mathematics and has been suggested by Campbell and Bamberger (1990) and Wearne and Hiebert (1985). The results of the National Assessment of Educational Progress indicated that solving problems was an area of weakness in students, yet teaching for problem solving has received little attention in the mathematics education literature. Campbell and Bamberger (1990) wrote that problem solving should be integrated into the mathematics program to enable students to apply mathematics to life situations. The connection between school mathematics and the real world, according to Wearne and Hiebert (1985), could be made by incorporating the use of symbols with everyday manipulatives.

The third approach to teaching problem solving, teaching via problem solving, has been discussed by Carpenter (1985), Kameenui and Griffin (1989), Siegler (1989), and Stanic and Kilpatric (1988). All concurred that mathematics instruction must include problems to learn concepts.

One approach to teaching via problem solving was offered by Stanic and Kilpatric (1988). He looked at problem solving as content (teaching about problem solving), problem solving as a skill (teaching for problem solving), and problem solving as an art (teaching via problem solving), and suggested that dealing with problems as an art was the most promising approach for mathematics. This approach required students to discover mathematics, rather than to deal with mathematics in a mechanical computational approach. Similarly, Kameenui and Griffin (1989) recommended using problem solving as a method to teach operations.

More specific recommendations for integrating problem solving have been offered. In examining the development of addition and subtraction problem solving, Carpenter (1985) found that children first used concrete objects, followed by abstract modeling, counting, and then number facts. He maintained that concepts should be taught first by using problem solving in these developmental steps and then by moving to computation.

The use of problem solving approaches in mathematics instruction have been shown to increase student achievement. The effect of teachers' knowledge of student problem solving ability on student performance was demonstrated by Peterson, Carpenter, and Fennema (1989). When teachers had increased knowledge concerning student problem solving strategies, they were able to incorporate the information to positively affect instruction, resulting in increased problem solving and achievement.

Other benefits have been demonstrated by Loif, Carey, Carpenter, and Fennema (1988), who studied the effect of training teachers in the structure and solution strategies to problems, and found that these teachers had more knowledge of their students, used small groups more frequently, and posed more questions. While these teachers spent less class time on factual information, the students had a higher recall of mathematics facts.

The 1989 Standards for School Mathematics have stressed the need for problem solving in mathematics and have encouraged the use of a wide variety of concrete materials as a means to improve the mathematics skills of students. The position of the Standards' writers on problem solving and manipulative materials has been supported by research findings and testing results. For example, the mathematics results of the 1986 National Assessment of Educational Progress indicated a lack of problem solving ability in students and the writers have recommended that use of manipulative materials. These results also demonstrated that students did better when pictures were included. While they could use operations procedurally, students could not explain the procedures and did not understand the underlying concept. The writers of the National Assessment of Educational Progress Report (Kouba, et al, 1988) stressed allowing students time to understand concepts before practicing procedures.

The use of manipulatives in elementary mathematics has been recommended in the literature for several decades (Gilbert & Bush, 1988). For example, theoretical support for manipulatives has been found in the work of Piaget and Gagné (Gilbert & Bush, 1988). Empirical evidence has been provided by Baroody (1989), Sowell (1989), and Suydam and Higgins (1977). Suydam and Higgins (1977) found that student achievement was greater when manipulatives were included in lessons than when manipulatives were omitted. Gilbert and Bush (1988), in their study of manipulative materials, found that primary teachers were generally familiar with a list of eleven mathematics manipulative materials rated by leading educators as the most important. In terms of teacher use, 65 percent of the teachers reported using manipulative materials in their classrooms once or more per week while 19 percent indicated their use of manipulative materials was once or less per month. The most frequently used manipulative devices included counters, bundleables, unifix cubes, and multibase blocks. The study found that inexperienced teachers used manipulatives more frequently than experienced teachers and that use decreased as grade level increased. The study offered no explanation as to why the use of manipulative materials decreased with experience.

Baroody (1989) maintained that meaningful learning in mathematics should begin with concrete experiences and move to the symbolic level. However, he found that manipulatives were not sufficient for learning. Students could use manipulatives mechanically and still not understand the concepts. Baroody found that manipulatives were most effective when the items were familiar to the student.

A meta-analysis of 60 studies to determine the effectiveness of manipulative materials was conducted by Sowell (1989). He reported that manipulative materials were significantly better than abstract instruction when the treatment lasted for a school year or longer, but shorter lengths of use did not produce significant findings. In contrast with the 1986 National Assessment of Educational Progress results, Sowell also found that the use of pictures was no better than abstract instruction. For manipulative instruction to be effective, Sowell stated that the teachers must receive extensive training.

There has been support in the research literature for the integration of problem solving and manipulative materials into the mathematics curriculum. Findings have suggested that students could compute, but they lacked an understanding of the concepts and the ability to solve mathematical problems. Manipulative materials have been recommended as a means to increase the understanding of mathematical concepts, thus aiding student in their ability to apply mathematics and to solve problems.

ACHIEVEMENT ASSESSMENT

With the publication of the 1989 Standards for School Mathematics, assessment that is alternative to traditional paper and pencil tests with forced choice formats have become a major issue in the literature. The *Arithmetic Teacher* now includes a monthly feature on implementing these curriculum and assessment Standards and provides suggestions for classroom teachers. Performance assessment has also received attention in other academic areas such as reading and writing. The following is a review of the literature on current assessment activities including standardized tests, teacher-made tests, and testing formats suggested to match the 1989 Standards for School Mathematics.

Standardized Tests - Strengths and Weaknesses

Standardized testing has been the source of much controversy. The current literature indicates that while standardized testing has been prevalent in education, its limitations have been regularly addressed and argued (Haney, 1985). Worthen and Spandel (1991) summarized the common criticisms of standardized tests including their effect on student learning, the lack of content match, and their effect on curriculum. Since the publication of the 1989 Standards for School Mathematics with their decreased emphasis on standardized tests, the measurement of student achievement has become a critical issue for teachers. The remainder of this section addresses the issues raised concerning the use of standardized tests. One problem that was encountered when reviewing the literature on standardized testing was that the authors did not clearly define the subject of their discussion on standardized testing as norm-referenced testing. However, this review focused on norm-referenced testing.

The effect of standardized tests on student learning and the role of standardized tests has been a major area of criticism. Haney (1985) found that standardized testing was on the increase but challenged its educational role. The primary role, according to Haney, was administrative (program evaluation, selection, and placement) rather than educational. Additional evidence of the administrative role of standardized tests was found by Rudman (1987). Standardized tests, according to Rudman, were used for decisions only remotely related to teaching, such as the rating of schools. Teachers reported that standardized testing took time away from teaching, leading Rudman to conclude that the link between teaching and standardized testing was not a strong one. A possible explanation for the limited effectiveness of standardized testing on education and student learning has been offered by Schalock, Fielding, Schalock, Erickson, and Brott (1985). They found that there was a lack of procedures for reporting test results that addressed the information needs of teachers and policies to guide the inquiry and implications of test information.

Worthen and Spandel (1991) stated that the educational role of a standardized test was to provide general performance information in the content areas. They indicated that these tests provided only a small portion of the assessment information a teacher relied on for decision making. Good classroom assessment begins with teacher assessment, according to Worthen and Spandel, but the standardized test can serve as a supplement to teacher assessment information.

Research findings have not been consistent in regard to teachers' use of standardized test results. Salmon-Cox (1981) found that teachers most frequently used observation, teacher-made tests, and interaction as methods of student assessment. Additionally, Salmon-Cox reported that when standardized tests scores and classroom performance did not agree, the teachers reported using classroom performance measures as opposed to the results from standardized tests. In contrast, Hall, Villeme, and Phillippy (1985) found that beginning teachers weighted standardized statewide minimum skills testing as most important for decisions concerning academic progress, promotion and retention, diagnosis of student weaknesses, and the adequacy of teaching and instructional materials but ranked teacher-prepared tests as most important for student self-evaluation and motivating student learning.

While no test has been perfect, Worthen and Spandel (1991) suggested that standardized tests have been useful. They have allowed for comparability in a manner not available with other types of tests and have allowed educators to get an overall view of student performance. Standardized testing alone has not been harmful but the inappropriate use of their results has the potential to be harmful (Worthen & Spandel, 1991).

The second area of criticism of standardized tests focused on content validity. The mismatch between the content of standardized tests and the school curriculum has been discussed by Haney (1985); Schalock, et al., (1985); Shriner and Salvia (1988); and Worthen and Spandel (1991). The limited curriculum match was described by Shriner and Salvia (1988) and Worthen and Spandel (1991) to be a result of the development of standardized tests for broad use, reflecting most curriculums to some extent, but none precisely. Haney (1985) and Schalock, et al., (1985) concurred that standardized testing may not cover content included in the curriculum, thus, resulting in a limited effect on education.

Several studies have examined the relationship between the content of the mathematics curriculum and commercial standardized tests. Encouragingly, when the relationship between mathematics textbooks and standardized tests was reviewed, Freeman, et al., (1983) found some commonalities between textbooks and tests but found differences as well. However, Shriner and Salvia (1988) reported that mathematics curricula series and standardized tests differed significantly in content and operations tested.

Additional evidence on the validity of standardized testing was provided by Willoughby (1990). When Willoughby focused specifically on the relationship between mathematics problem-solving questions that were congruent with the standards and the mathematics items on standardized testing instruments, he found very low correlations, ranging from -.18 to .11. He expressed concern because educators have assumed that standardized tests measure something important. If teachers used these standardized tests and textbooks were written to these tests, Willoughby maintained that children would not have the opportunity to learn problem solving. Willoughby questioned the appropriateness of standardized tests when mathematics educators were focusing on higher order skills and problem solving.

The third criticism identified by Worthen and Spandel (1991) addressed the influence of standardized tests on what was taught, or teaching to the test. There are claims that standardized tests dictate or restrict what is taught in the classroom to the content measured by the test. The fact that a test may "drive" the curriculum was not the fault of the test, according to Worthen and Spandel. The question that should have been asked was "How were curriculum content decisions made?" Standardized tests have been built around the content of textbooks, teachers, and other tests. Influences between these aspects were difficult to separate.

In summary, the criticisms of standardized testing center on the role of testing and the applicability of results in the classroom, the curriculum match, and the influence on curriculum decisions. The standardized test has generally been adequate when used to differentiate students and make relative judgements about performance, but has been less useful for making instructional decisions or assessing the effect of classroom procedures (Wardrop, et al., 1982).

Teacher-made Tests

Surprisingly, teacher assessment activities have rarely attracted attention in the literature (Mehrens & Lehmann, 1987; Stiggins, 1985). Standardized testing, according to Stiggins (1985), has received much attention despite the fact that standardized tests mean little to classroom teachers. With the primary focus of measurement research on standardized paper-and-pencil testing rather than teacher-made assessment, the assessment areas in which teaches need help remains unknown (Stiggins, 1985). Mehrens and Lehmann (1987) have studied the assessment needs of classroom teachers and found that most teachers were not well-trained in assessment. Stiggins has suggested that the focus of research in the measurement field needs to be expanded to include teacher-made tests. The following presentation of literature is a discussion of the assessment practices of classroom teachers. Unless specified, teachers include those of all grade levels and content areas.

Teachers have been found to regularly use a variety of testing formats and types of tests such as self-developed assessment instruments, observations, and standardized instruments for decision-making (Stiggins, 1985). Similar findings have been reported by Hall, Villeme, and Phillippy (1985). In a study of the types of tests used by beginning teachers, Hall, Villeme, and Phillippy found that all played some role in teacher decision-making but none were judged by beginning teachers as playing a clearly dominant role.

Studies by Hall, Carroll, and Comer (1988); Mehrens and Lehmann (1987); and Stiggins and Bridgeford (1985) have concluded that teacher-made tests were used most frequently for decision-making. Stiggins and Bridgeford (1985) found teacher-made tests (objective formats modeled after standardized tests) were used most frequently for all purposes (diagnostic, grouping, grading, evaluating instruction, and reporting achievement). Hall, Carroll, and Comer (1988) concluded that teachers were able to put tests in their proper perspective using all types of tests. Mehrens and Lehmann (1987) reported that despite the lack of reliability studies on teacher-made tests, 75 percent of teachers used their own tests for decision-making, including grouping and grading.

The testing formats used by classroom teachers have received little research attention. Stiggins and Bridgeford (1985) reported that teachers most frequently mentioned observation as a method to obtain information for decision-making. Although observations have been found to be used most frequently by classroom teachers, Stiggins, Conklin, and Bridgeford (1986) found teachers of all subjects and grade levels used matching items more frequently than multiple choice or true-false.

In addition, Stiggins and Bridgeford (1985) found that 88 percent of the teachers used performance tests to some extent. Performance tests were described by Stiggins and Bridgeford (1985) as assessing several important student characteristics. These included the application of a skill, the completion of a task in a real or simulated environment, and the production of an observable task.

To further describe classroom assessment practices, Stiggins and Bridgeford (1985) found that grade level contributed to the differences in the type of classroom assessment used. As grade level increased, teachers were more likely to use teacher-made tests over published tests. The study did not investigate the availability of published tests for all grade levels. Academic subject area was also related to testing format. Mathematics and science teachers tended to rely more heavily on objective format paper and pencil tests, while writing teachers tended to use performance (process) assessment more frequently.

Additional insight into classroom testing practices has been offered by Mehrens and Lehmann (1987) and Stiggins, Conklin, and Bridgeford (1986). Mehrens and Lehmann found that 80 percent of teacher-made test questions were at the knowledge level (assessing student recall of factual information). Stiggins, Conklin, and Bridgeford (1986) also found classroom testing was predominately at the knowledge level.

To summarize the literature on teacher-made tests, classroom teachers have relied primarily on self-constructed tests consisting of objective format items. However, the work of Stiggins and Bridgeford (1985) suggest that both observation and performance tests are integral parts of teacher assessment practices. Since the preponderance of items could be classified as knowledge items, the issue of curriculum match may once again be addressed. With standardized testing, the curriculum match issue centered on content. With teacher-made tests, the curriculum match question may center on the match between item levels, knowledge items as opposed to higher level items, and the curriculum.

Alternative Testing Formats

With the emphasis on problem solving and reasoning in the 1989 Standards, current teacher-made testing practices that focus on the knowledge level may not assess the skills taught in the classroom. In response to the need for a better match between testing format and assessment objectives, the mathematics education community has recommended alternative assessment formats. The following review will discuss the strengths and weaknesses of the multiple choice format and those testing formats that have been presented as alternatives to the multiple choice item.

Multiple choice test items have appeared on classroom tests in all grade levels and subjects (Carey, 1988; Guthrie, 1984). Multiple choice items generally included on standardized tests were found by Gutrhrie (1984) to be seldom written to assess achievement other than factual recall. Conversely, Mehrens and Lehmann (1984) and Sax (1980) reported wide use of the multiple choice format due to its versatility. One advantage of the multiple choice format identified by Sax was the ability to measure objectives from the knowledge level to the most complex level. Mehrens and Lehmann reported that multiple choice questions could measure student ability in both factual recall and reasoning. Therefore, the question to be addressed may not center on the item format, but rather on the skill of the item writer.

Studies of the multiple choice format have been conducted by Frary (1985); Kolstad, Briggs, and Kolstad (1990); Norris (1989 & 1990); and Schoen, Blume, and Hoover (1990). The results of these studies indicated both weaknesses and strengths of the multiple choice format. Kolstad, Briggs, and Kolstad (1990) found that when assessing student achievement, the multiple choice format may not provide information on false ideas and misinformation. While indicating that research on multiple-choice versus free response format (completion) was limited, Frary (1985) found, in a simulation study, that reliability and validity were somewhat lower on the multiple choice format than on a free response format. In contrast, Schoen, Blume, and Hoover (1990) reported that the multiple choice format, with well-developed distractors, was able to assess mathematics estimation procedures used by student as well as the open-ended format could.

Norris (1989 & 1990) has studied the use of the multiple choice format to assess critical thinking. Skeptical of the multiple choice format on the grounds that only weak evidence of the thinking process can be generated by multiple choice tests, Norris (1990) compared verbal reports of the thinking process and the multiple choice format He found that the type of item had no effect on the thinking process and suggested that interviews could be used to validate a multiple choice thinking test. In his earlier research, Norris (1989) indicated that the breadth of critical thinking may not be adequately assessed using the multiple choice format. Developing foils that take the various aspects of critical thinking into consideration may be impractical; therefore Norris suggested student interviews to assess thinking.

Despite their popularity, due to their versatility and ease of scoring, multiple choice items have limitations and are not appropriate for all testing purposes (Guthrie, 1984). Measurement theorists have warned against inappropriate testing formats by suggesting that the item format must be congruent with the conditions, behavior, content, and behavioral objectives of the assessment (Carey, 1988). Guthrie and Lissitz (1985) and Robinson (1987) concurred that testing formats should vary with the testing purpose and educational decisions to be made. Berlak (1985) suggested examining a variety of testing formats including portfolios and profiles, described as documentary evidence of student performance, and observations of student performance as ways to gain information about student performance.

Stiggins (1982) furthered the idea that assessment format and decision-making were related. He studied direct and indirect assessment formats in writing. The direct assessment format involved evaluating students' knowledge of writing rules and procedures through writing samples while the indirect assessment format typified the paper and pencil testing approach using a multiple choice format. Stiggins found that while the two formats did assess some of the same performance factors, each format also assessed some unique aspects of writing. This finding led to the conclusion that format selection should be dependent upon the educational decision to be made and the type of information needed.

The controversy over the multiple choice format and the suggestion by the 1989 Standards for School Mathematics to use a variety of testing formats, have led to discussions over appropriate formats. Clark, Clark, and Lovitt (1990) reported that the restructuring of the goals and practices of mathematics education must be accompanied by assessment strategies which reflect the new conception of mathematics. Since concepts considered valuable were communicated to students through testing, the assessment must be comprehensive; therefore, Clark, Clark, and Lovitt have suggested that assessment tools must be sensitive to process, as well as product, and that teachers must expand their repertoire of assessment strategies beyond paper and pencil tests. Formats such as observations and checklists, interviews, oral questioning, and portfolios, with alternative scoring approaches and the use of calculators, have received support (Charles, Lester, & O'Daffer, 1988; Clark, Clark, & Lovitt, 1990; Guthrie, 1984; Guthrie & Lissitz, 1985; Haney, 1985; Mehrens & Lehmann, 1987; Peck, Jencks, & Connell, 1989; Robinson, 1987; Silver, 1990; and Webb & Briars, 1990).

The observation and checklist method of student assessment has been suggested by Clark, Clark, and Lovitt (1990); Mehrens and Lehmann (1987); and Webb and Briars (1990). Clark, Clark, and Lovitt, believing that a wealth of assessment information was available in the classroom, suggested that teachers observe student behaviors during informal assessment activities in the classroom by way of a checklist. This information could then serve all assessment purposes. Webb and Briars (1990) concurred that informal assessment (observation) could be recorded thereby reducing the need to assess the same concept in a formal procedure. Mehrens and Lehmann (1987) suggested that observational data could give teachers information not available in other formats. These studies did not investigate the resources available for classroom teachers to maintain observational records nor the appropriateness of using observation to assess the formative activities that occurred in the learning process.

Peck, Jencks, and Connell (1989) and Silver (1990) studied the benefit of student interviews as a testing format. Peck, Jencks, and Connell reported that brief interviews, focused on student reasoning and the justification of procedures used to solve problems, combined with paper and pencil tests yielded more student information concerning concept understanding than the written test alone. Silver (1990) concluded that using interviews and think-aloud probes would allow the teacher to gain information on the student's thinking process not available with other formats. Using only paper and pencil tests, Peck, Jencks, and Connell found that teachers classified students incorrectly, in terms of concept understanding, 52 percent of the time. Concluding that paper and pencil tests alone may not have correctly evaluated students' conceptual understanding, Peck, Jencks, and Connell suggested that conducting student interviews, at the point of concept introduction and completion, would result in improved assessment of student understanding of mathematical ideas.

Additional classroom assessment formats have been suggested. Oral questioning and answering techniques have been suggested by Robinson (1987) while Guthrie and Lissitz (1985) advised using teacher judgement and process records, which describe the student in terms of cognition as well as how and where they were making errors. Guthrie (1984) proposed that free response essay formats were necessary to measure interpretation, problem solving, and the application of principles.

The use of student portfolios has been recommended by Guthrie (1984) and Haney (1985). Haney (1985) surveyed assessment procedures in the United States and located a small school district which successfully implemented alternative assessment using portfolios. Student records included narrative descriptions of students' abilities, observations, and examples of written work. He concluded that using portfolios as an alternative assessment was a realistic possibility.

In addition to issues related to testing format, scoring procedures have also received attention in the literature. Procedures for classroom teachers, using holistic scoring to evaluate the problem solving process, have been developed by Charles, Lester and O'Daffer (1988). They divided holistic scoring into three methods which included analytic scoring, focused holistic scoring, and general impression scoring. Analytic scoring required the evaluator to assign points, based on established criteria, to certain phases of the problem solving process. The result is a score for each phase. Focused holistic scoring occurred when a numerical score, based on specific criteria relevant to the thinking process, was assigned to the total solution of a problem. General impression scoring, unlike focused holistic scoring which required the development of specific written criteria, involved rating the total solution numerically based on the general impression.

With the curriculum focus on problem solving, Otis and Offerman (1988) suggested that a focused holistic scoring method could be easily modified for use by individual teachers to score problem solving activities. To assess problem solving, according to Otis and Offerman, both the thinking process and the final product must be evaluated using holistic scoring. The product of mathematics problem solving has not been a difficult area to assess, but the process has been frequently ignored. Holistic and analytic scoring were suggested by Webb and Briars (1990) as alternative methods to right/wrong scoring.

A final assessment issue is centered on the use of calculators. The 1989 Standards for

School Mathematics included the use of technology not only for instruction but also for assessment. Heid (1988), for example, stated that students should have access to calculators in all testing situations. DeMana and Waits (1990) also agreed that calculators should be included in the classroom. The suggestion for including computing devises for routine procedures is based on the widespread availability of more powerful and less expensive calculators. With access to calculators outside the classroom, the focus inside the classroom could include concepts and principles, rather than product. Heid (1988) expanded on this to suggest that calculators would be included in the classroom when they were included in testing.

In summary, a variety of testing formats have been suggested as alternatives to current classroom assessment procedures. There is no consensus as to the optimal format, but rather a smorgasbord of choices have been offered including demonstration or performance, interviews, process with holistic scoring, and observation. Research on the effectiveness of these formats is limited, but the trend toward alternative assessment is progressing in the field of mathematics education with the development of the 1989 Standards for School Mathematics.

METHODOLOGY

In order to study the manner in which classroom teachers implemented the assessment component of the 1989 Standards for School Mathematics and to compare the differences in the evaluative information generated by traditional assessment methods and alternative assessment methods the following methodology was used.

SAMPLE

Schools

All district elementary schools were invited to participate in the K-3 Mathematics Specialist Project. Ten schools were selected from a volunteer pool that represented the district elementary schools on characteristics including the size of the school, the socio-economic status of the school (the proportion of students on free and reduced lunch), the geographic location, and the ethnic characteristics of the student population.

Teachers

Once school locations were selected, the principal recommended teachers to participate in the project based on the following criteria: interest in the program, skill in the instruction of mathematics, grade level taught, and potential ability to function as a trainer. Thirty-four teachers in grades kindergarten to three were selected to participate in the K-3 Mathematics Specialist Project. At nine school sites, from two to three teachers were included in the project. There was an attempt to distribute the teachers across grade levels. One of the ten elementary schools, designated as a model school, included twelve of its teachers in the project (four teachers at each of the project grade levels).

The distribution of teachers by grade level included seven kindergarten teachers, ten grade one teachers, seven grade two teachers, and nine grade three teachers. One teacher had a combined second/third grade class. Table 1 contains the educational level and teaching experience of the teachers by grade level.

TABLE 1

TEACHER EDUCATION LEVEL AND YEARS OF EXPERIENCE

	Degree		Experience			
Grade	BA	MA	1-2 Years	3-5 Years	6-10 Years	>11 Years
Kindergarten	5	2	0	0	2	5
Grade 1	6	4	4	0	0	6
Grade 2	5	3	1	1	0	6
Grade 3	8	1	1	1	2	5
Total	24	10	6	2	4	22
Percent of Total	71%	29%	18%	6%	12%	65%

The following discussion describes the procedures and data analysis that were employed to answer research questions one and two:

1. Do K-3 teachers vary after training in the methods of implementing alternative assessment strategies depending on teacher grading orientation, class size, grade level, and student mathematics ability?

2. Do K-3 teachers vary after training in the methods of implementing alternative assessment strategies by assessment standard? The assessment standards of the 1989

Standards for School Mathematics include Mathematical Power, Problem Solving, Communications, Reasoning, Mathematical Concepts, Mathematical Procedures, and Mathematical Disposition.

TEST ITEM CLASSIFICATION VARIABLES

Each K-3 Mathematics Specialist teacher maintained a portfolio containing the assessment instruments used in their classroom when measuring students for summative purposes. The portfolio was analyzed to classify every test item on the variables of level of question, assessment format, use of manipulative materials, scoring method, and content. The variables selected on which to classify the assessment items were based on the foci of the 1989 Standards for School Mathematics. In addition, each item was classified according the standard measured. The item classification variables and standards are described below:

Level of Questions. Items were classified as either knowledge level items or higher order questions. This classification was made using Bloom's Taxonomy (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956).

Assessment Format. Items were classified according to the assessment format used by the teacher. The formats included were forced choice (multiple choice, true-false, and matching), oral questioning, student demonstration or performance assessment, journals, and free response.

Use of Manipulative Materials. The items were also classified according to the use of concrete objects in conjunction with the assessment process. Each specific manipulative material was recorded.

Method of Scoring. The test items were classified according to the method used to score student responses. The scoring methods included right/wrong, focused holistic scoring, analytic scoring, and general impression scoring.

Content Area. Items were also classified according to the mathematics content or concept that was assessed.

Mathematical Power. Test items were classified as Mathematical Power when the assessment measured the extent to which students' have integrated all aspects of mathematical knowledge.

Problem Solving. The test items measuring Problem Solving assessed the students' ability to use problem solving techniques, verify and interpret results, ask questions, and use given information.

Communications. Test items were classified as Communications when the assessment measured the students' ability to attach meaning to concepts and procedures of mathematics; fluency in talking about, understanding and evaluating mathematical ideas; and use of vocabulary, notation, and structure to express and understand ideas and relationships.

Reasoning. When the assessment measured the students' use of different types of reasoning, the test items were classified as Reasoning.

Mathematical Concepts. These items measured the students' understanding of definitions, the ability to discriminate between attributes of a concept, to represent concepts in various ways, and to recognize the various meanings of concepts.

Mathematical Procedures. Test items were classified as Mathematical Procedures when the assessment measured the execution of procedures including when to apply procedures, why procedures work, how to verify correct answers, and to differentiate correct procedures and incorrect procedures.

Mathematical Disposition. Test items were classified as Mathematical Disposition when the assessment measured the students' attitude toward mathematics as well as the tendency to think and to act in positive ways toward mathematics.

For each individual test, the proportion of items classified according to the item classification variables and the standards was recorded on the Assessment Matrix shown in Figure 1. In order to ensure agreement of the classifications, interrater agreement was estimated. Each instrument was rated and a random sample of 37.5 percent was rated by a second rater. The interrater agreement on the ratings between the first and second raters was 90.0 percent. A third rater was used when disagreement existed between the first and second rater. In the case of the later, the two ratings in agreement were used.

TEACHER CLASSIFICATION VARIABLES

The K-3 Mathematics Specialist teachers were then classified according to their grading orientation, class size, grade level taught, and student mathematics achievement.

Grading Orientation. In order to determine the grading orientation of the K-3 Mathematics Specialist teachers, the Grading Orientation Questionnaire, located in Appendix 1, was completed by 30 of the 34 project teacher (88% return rate).

The teachers were classified as either an achievement oriented grader or a non-achievement oriented grader based on their responses to the Grading Orientation Questionnaire. The mean weight of the two achievement factors (post-tests and seatwork/homework) was compared as a category with the mean weight of the non-achievement factors (extra credit, attitude, effort, motivation, participation, and behavior). The grading orientation of the teacher was determined by the category of factors with the highest mean weight.

Class Size. The median size (23.5) of the project teachers' mathematics class was used to classify the class size as a small class (16 to 23 students) or a large class (24 to 31 students).

ASSESSMENT MATRIX

Grade Level: Grading Philosophy: Class Size: Student Achievement:

	Mathematics Power	Problem Solving	Communi-cations	Reasoning	Mathematics Concepts	Mathematics Procedures	Mathematics Disposition
Level of Questions Knowledge Level Higher Level							
Assessment Format Forced Choice Free Response Oral Demonstration Journal							
Manipulative Materials							
Scoring Method Right/Wrong Holistic Analytic Impression							
Content							

Figure 1. K-3 Mathematics Specialist Project Assessment Matrix

Student Mathematics Ability. The median mathematics achievement score from the Spring 1990 administration of the Stanford-8 Achievement Mathematics test was used to classify the academic achievement of students in the class for each K-3 mathematics Specialist teacher. The classes of the kindergarten and first grade teachers were not classified as test results were not available. Median class achievement less than the 40th percentile was classified as low achievement. Class achievement greater than or equal to the 40th percentile and less than or equal to the 60th percentile was classified as average achievement. Class achievement above the 60th percentile was classified as high achievement. Due to the small number of classes with low and average student achievement, these two groups were combined into low to average student achievement for the analyses based on student achievement.

DATA ANALYSIS

Research Questions 1 and 2

Following the item classification, the mean proportion of test items for each item classification variable and standard was computed for each teacher using the following procedures:

1. For each assessment instrument, the proportion of items in each category was computed. For example, an instrument with ten items, five of which were higher level questions and five were knowledge level questions was recorded as .50 higher level questions and .50 knowledge level questions. The same procedure was used to determine the proportion of items using each assessment format, each manipulative material, each scoring method, and each content area.
2. The proportion of items was also determined for each assessment standard. For example, if an instrument included ten items, five of which measured Mathematical Concepts with higher level questions, .50 was recorded under Mathematical Concepts - higher level questions. For each assessment instrument, the total proportion of items for item level, assessment format, manipulative materials, scoring method, and content would equal 1.00 across the seven assessment standards.
3. Once each assessment instrument was rated and the proportions determined, the mean proportion for the teacher was then calculated. The total proportion of items within each cell of the Assessment Matrix was divided by the number of assessment instruments to determine the mean proportion of test items. The mean proportion for each cell was computed and then analyzed to answer research question one.

Inferential statistics were then used to determine reliable differences in the implementation of alternative assessment according to the teacher stratification variables and the standards. For the first research question, contrasting the use of alternative assessment according to the teacher stratification variables, one sample t-tests were used to compare:

1. the proportion of higher level questions with the proportion of knowledge level questions;
2. the proportion of items using alternative assessment formats (demonstration items, oral items, and journal items) with that of traditional assessment formats (forced choice items and free response items);
3. the proportion of items using manipulative materials with the proportion of items not using manipulative materials; and
4. the proportion of items scored using the alternative scoring methods (analytic scoring, focused holistic scoring, and general impression scoring) with the proportion of traditionally scored items (right/wrong scoring).

In addition, separate independent means t-tests were used to compare the mean proportion of higher level questions, alternative format items, items using manipulative materials, and alternatively scored items with the following teacher groupings:

1. Teachers with an achievement grading orientation were compared to teachers with a non-achievement grading orientation.
2. Teachers with small classes were compared to teachers with large classes.
3. Teachers of kindergarten and first grade were compared to teachers of second and third grade.
4. Teachers of students with low to average mathematics achievement were compared to teachers of students with high mathematics achievement.

The second research question addressed the use of alternative assessment methods according to each standard. One sample t-tests were used to compare the item classification variables described above for each standard. The standards of Problem Solving, Communications, Mathematical Concepts, and Mathematical Procedures were included in the analyses. Mathematical Power, Reasoning, and Mathematical Disposition were not included because the frequency with which the K-3 Mathematics Specialist teachers assessed these standards was small.

The content data were not analyzed using inferential statistics. This data could not be collapsed into any logical categories with which to perform a t-test.

In addition to the inferential analyses, the mean proportion of items, when stratified by teacher grading orientation, class size, grade level, student achievement, and each standard were used to describe patterns in the teachers' classroom assessment practices.

Research Question 3

In order to determine the evaluative information gained by alternative assessment procedures when compared with traditional assessment procedures, an Assessment Questionnaire was developed. The Assessment Questionnaire is located in Appendix

2. The issues explored included content areas where alternative or traditional assessments was more appropriate, the type of evaluative information available through alternative and traditional assessment, and differences in teacher confidence in the evaluative information from alternative and traditional assessment. In addition, the frequency which alternative assessment formats were used and the relative difficulty of implementing alternative assessment was explored.

The teacher questionnaires were tabulated and comments were summarized to include common and unique comments. The results were analyzed using descriptive statistics and *t*-tests to determine if differences existed in the teachers' frequency of use, difficulty of use, and confidence of information between alternative assessment formats and unit tests.

RESULTS

The purpose of this study was to examine the manner in which classroom teachers implemented the assessment component of the 1989 Standards for School Mathematics and to determine the evaluative information gained by alternative and traditional assessment procedures. The results are described separately for each research question.

IMPLEMENTATION OF ALTERNATIVE ASSESSMENT ACCORDING TO TEACHER GRADING ORIENTATION, CLASS SIZE, GRADE LEVEL AND STUDENT ACHIEVEMENT LEVEL

Research Question 1: Do K-3 teachers vary after training in the methods of implementing alternative assessment strategies depending on teacher grading orientation, class size, grade level, and student mathematics ability?

Level of Questions

When the proportion of higher level questions was compared with the proportion of knowledge level questions, the difference in proportions was significant (t=-2.29, p=.029). The teachers were found to use knowledge level questions significantly more frequently than higher level questions. The results of the analyses on the item levels are displayed in Table 2.

When the proportion of higher level questions was compared according to the teacher stratification variables, no significant differences were found. Therefore, teacher grading orientation, class size, grade level, and student achievement level did not relate significantly with the proportion of higher level questions used by the teachers.

Assessment Format

There was not a statistically significant difference in the teachers' use of alternative and traditional assessment format items ($t=-.63$, $p=.532$). Also, there were no significant differences in the proportion of items using alternative assessment formats according to teacher grading orientation, class size, grade level, and student achievement levels. The results of the analyses on the assessment formats are included in Table 3.

When the use of assessment formats were reviewed according to the teacher variables, the teachers were found to vary in the frequency of using each assessment format but the general pattern of use was the same regardless of teacher grading orientation, class size, grade level, and student achievement level. The teachers used free response items most frequently followed by demonstration items. The formats of journal, oral, and forced choice were used infrequently by the teachers. Therefore, patterns in the teachers' use of assessment format did not vary according to the teacher stratification variables. The results are contained in Table 4.

Manipulative Materials

A significant difference was found in the proportion of items using and not using manipulative materials ($t=3.68$, $p=.001$). The teachers used items with manipulative materials significantly more frequently than items without manipulative materials.

A significant difference also existed in the use of manipulative materials according to grade level. The kindergarten and first grade teachers used manipulative materials significantly more frequently than did the second and third grade teachers ($t=2.25$, $p=.032$). There were no significant differences in the use of manipulative materials according to teacher grading orientation, class size, and student achievement level. The results of the analyses on the use of manipulative materials are presented in Table 5.

The teachers used items with manipulative materials extensively when assessing student performance. Unifix cubes, base-ten blocks, and counters tended to be used with the greatest frequency. The other materials were used infrequently and were combined into the category of other manipulatives. The teachers used such a variety of manipulative materials that there was not any pattern of use according to teacher grading orientation, class size, grade level, and student achievement. These results are displayed in Table 6.

Scoring Method

There were no significant differences found in the proportion of items using alternative scoring methods when compared with the proportion of items using traditional scoring methods ($t=-1.82$, $p=.077$). Grading orientation, class size, grade level, and student mathematics achievement were not found to have significant association with the proportion of items using alternative scoring methods. Table 7 includes the results of the analyses on the scoring methods used by the teachers.

When the mean proportion of items using each scoring method was reviewed, traditional scoring or right/wrong scoring was found to be the most commonly used scoring procedure. These results are displayed in Table 8. Of the scoring procedures that are reflective of alternative assessment (analytic scoring, focused holistic scoring, and general impression scoring), general impression scoring was used with the greatest frequency. The use of analytic and focused holistic scoring was infrequent by the teachers. The pattern of using each scoring method was similar regardless of teacher grading orientation, class size, grade level, and student achievement.

Content

The specific mathematics concept or content measured by the instruments contained in the teachers' assessment portfolios were combined into the categories of algorithms, number concepts, geometry, and other concepts. The concepts included in the other concepts category, such as measurement, volume, time, and money, were measured infrequently by the project teachers. These results are presented in Table 9. The teachers were found to most frequently measure the algorithms, number concepts, and geometry regardless of grading orientation, class size, grade level, and student mathematics achievement. The teachers did include assessment items in their portfolios which assessed a wide range of mathematics content.

TABLE 2

COMPARISON OF ITEM LEVELS FOR THE K-3 MATHEMATICS SPECIALIST TEACHERS BY GRADING ORIENTATION, CLASS SIZE, GRADE LEVEL, AND STUDENT ACHIEVEMENT LEVEL

	N	Mean Proportion	df	t	p
All K-3 Mathematics Specialist Teachers					
Higher Level Questions	33	.401	32	-2.29	.029*

TABLE 2 (cont.)

Knowledge Level Questions	33	.598			
Grading Orientation - Higher Level Questions					
Achievement Grading Orientation	22	.399	26	-.42	.676
Non-Achievement Grading Orientation	6	.443			
Class Size - Higher Level Questions					
Small Classes	15	.436	31	.73	.468
Large Classes	18	.372			
Grade Level - Higher Level Questions					
Kindergarten and First Grade	16	.396	30	-.30	.769
Second and Third Grade	16	.422			
Student Achievement - Higher Level Questions					
Low to Average Student Achievement	6	.250	15	-2.09	.054
High Student Achievement	11	.492			

*Significant at the .05 alpha level.

TABLE 3

COMPARISON OF ITEMS USING ALTERNATIVE AND TRADITIONAL ASSESSMENT FORMATS FOR THE K-3 MATHEMATICS SPECIALIST TEACHERS BY GRADING ORIENTATION, CLASS SIZE, GRADE LEVEL, AND STUDENT ACHIEVEMENT LEVEL

	N	Mean Proportion	df	t	p
All K-3 Mathematics Specialist Teachers					
Alternative Assessment Formats	33	.468	32	-.63	.532
Traditional Assessment Formats	33	.531			
Grading Orientation - Alternative Assessment Formats					
Achievement Grading Orientation	22	.427	26	-.74	.468
Non-Achievement Grading Orientation	6	.529			
Class Size - Alternative Assessment Formats					
Small Classes	15	.421	31	-.87	.391
Large Classes	18	.508			
Grade Level - Alternative Assessment Formats					

TABLE 3 (cont.)

Kindergarten and First Grade	16	.486	30	.50	.621
Second and Third Grade	16	.436			
Student Achievement - Alternative Assessment Formats					
Low to Average Achievement	6	.450	15	-.01	.989
High Achievement	11	.451			

TABLE 4

MEAN PROPORTION OF ASSESSMENT FORMATS BY TEACHER GRADING ORIENTATION, CLASS SIZE, GRADE LEVEL, AND STUDENT MATHEMATICS ACHIEVEMENT

	N	Demonstration	Journal	Oral	Forced Choice	Response	Free Total[1]
Grading Orientation[2]							
Achievement	22	.259	.055	.113	.113	.459	.999
Non-Achievement	6	.310	.057	.162	.048	.423	1.000
Both	2	.415	.135	.050	.095	.310	1.005
Class Size[3]							
Small	15	.235	.094	.092	.088	.492	1.001
Large	18	.336	.049	.123	.091	.399	.998
Grade Level[4]							
K-1	16	.343	.052	.091	.104	.409	.999
2-3	16	.216	.091	.129	.076	.489	1.001
Student Achievement[5]							
Low to Average	6	.245	.045	.160	.135	.417	1.002
High	11	.236	.107	.108	.045	.503	.999

[1]Total may not sum to 1.000 due to rounding.

[2]Three teachers with unclassified grading orientations are not included in the total group.

[3]Small Classes = 16 to 23 students
Large Classes = 24 to 31 students

[4]The multiple grade level teacher was not included in the total.

[5]Low to Average Student Achievement = 1-60 National Percentile
High Student Achievement = 61-99 National Percentile
The kindergarten and first grade teachers are not included.

TABLE 5

COMPARISON OF ITEMS USING AND NOT USING MANIPULATIVE MATERIALS FOR THE K-3 MATHEMATICS SPECILAIST TEACHERS BY GRADING ORIENTATION, CLASS SIZE, GRADE LEVEL, AND STUDENT ACHIEVEMENT LEVEL

	N	Mean Proportion	df	t	p
All K-3 Mathematics Specialist Teachers					
Use of Manipulative Materials	33	.662	32	3.68	.001*
Non-Use of Manipulative Materials	33	.334			
Grading Orientation - Use of Manipulative Materials					
Achievement Grading Orientation	22	.616	26	-1.51	.142
Non-Achievement Grading Orientation	6	.810			
Class Size - Use of Manipulative Materials					
Small Classes	15	.583	31	-1.68	.104
Large Classes	18	.729			
Grade Level - Use of Manipulative Materials					
Kindergarten and First Grade	16	.759	30	2.25	.032*

TABLE 5 (cont.)

Second and Third Grade	16	.564			
Student Achievement - Use of Manipulative Materials					
Low to Average Student Achievement	6	.610	15	.48	.638
High Achievement	11	.551			

*Significant at the .05 alpha level.

42

TABLE 6

MEAN PROPORTION OF ITEMS USING MANIPULATIVE MATERIALS BY TEACHER GRADING ORIENTATION, CLASS SIZE, GRADE LEVEL, AND STUDENT MATHEMATICS ACHIEVEMENT

	N	Base-Ten Blocks	Unifix Cubes	Counters	Other Manipulatives	NonUse
Grading Orientation[1]						
Achievement	22	.098	.103	.051	.369	.379
Non-Achievement	6	.012	.127	.000	.658	.203
Both	2	.000	.065	.035	.600	.300
Class Size[2]						
Small	15	.084	.087	.047	.364	.418
Large	18	.069	.141	.031	.495	.264
Grade Level[3]						
K-1	16	.034	.149	.038	.542	.237
2-3	16	.106	.092	.032	.336	.434
Student Achievement[4]						
Low to Average	6	.033	.112	.028	.439	.388
High	11	.161	.073	.045	.276	.445

[1]Three teachers with unclassified grading orientations are not included in the total group.

[2]Small Classes = 16 to 23 students
Large Classes = 24 to 31 students

[3]The multiple grade level teacher was not included in the total.

[4]Low to Average Student Achievement = 1-60 National Percentile
High Student Achievement = 61-99 National Percentile
The kindergarten and first grade teachers are not included.

TABLE 7

COMPARISON OF ITEMS USING ALTERNATIVE AND TRADITIONAL SCORING FOR THE K-3 MATHEMATICS SPECILAIST TEACHERS BY GRADING ORIENTATION, CLASS SIZE, GRADE LEVEL, AND STUDENT ACHIEVEMENT LEVEL

	N	Mean Proportion	df	t	p
All K-3 Mathematics Specialist Teachers					
Alternative Scoring	33	.409	32	-1.82	.077
Traditional Scoring	33	.590			
Grading Orientation - Alternative Scoring					
Achievement Grading Orientation	22	.383	26	.19	.849
Non-Achievement Grading Orientation	6	.356			
Class Size - Alternative Scoring					
Small Classes	15	.367	31	-.78	.411
Large Classes	18	.445			
Grade Level - Alternative Scoring					
Kindergarten and First Grade	16	.466	30	1.03	.311

TABLE 7 (cont.)

Second and Third Grade	16	.361			
Student Achievement - Alternative Scoring					
Low to Average Achievement	6	.372	15	.17	.868
High Achievement	11	.349			

45

TABLE 8

MEAN PROPORTION OF SCORING METHODS BY TEACHER GRADING ORIENTATION, CLASS SIZE, GRADE LEVEL, AND STUDENT MATHEMATICS ACHIEVEMENT

	N	Holistic Scoring	Analytic Scoring	General Impression Scoring	Right/Wrong Scoring	Total[1]
Grading Orientation[2]						
Achievement	22	.037	.043	.303	.617	1.000
Non-Achievement	6	.023	.023	.310	.643	.999
Both	2	.100	.000	.600	.300	1.000
Class Size[3]						
Small	15	.033	.042	.292	.633	1.000
Large	18	.040	.026	.379	.554	.999
Grade Level[4]						
K-1	16	.021	.024	.421	.532	.998
2-3	16	.037	.044	.280	.639	1.000
Student Achievement[5]						
Low to Average	6	.067	.033	.272	.628	1.000
High	11	.044	.046	.259	.652	1.000

[1]Total may not sum to 1.000 due to rounding.

[2]Three teachers with unclassified grading orientations are not included in the total group.

[3]Small Classes = 16 to 23 students
Large Classes = 24 to 31 students

[4]The multiple grade level teacher was not included in the total.

[5]Low to Average Student Achievement = 1-60 National Percentile
High Student Achievement = 61-99 National Percentile

The kindergarten and first grade teachers are not included.

TABLE 9

MEAN PROPORTION OF CONTENT MEASURED BY TEACHER GRADING ORIENTATION, CLASS SIZE, GRADE LEVEL, AND STUDENT MATHEMATICS ACHIEVEMENT

	N	Algorithms	Number Concepts	Geometry	Other Concepts	Total[1]
Grading Orientation[2]						
Achievement	22	.434	.160	.069	.334	.997
Non-Achievement	6	.185	.270	.123	.422	1.000
Both	2	.430	.000	.140	.425	.995
Class Size[3]						
Small	15	.305	.125	.107	.463	1.000
Large	18	.309	.222	.079	.389	.999
Grade Level[4]						
K-1	16	.225	.333	.075	.365	.998
2-3	16	.523	.034	.110	.330	.997
Student Achievement[5]						
Low to Average	6	.415	.033	.200	.354	1.002
High	11	.572	.031	.056	.335	.994

[1]Total may not sum to 1.000 due to rounding.

[2]Three teachers with unclassified grading orientations are not included in the total group.

[3]Small Classes = 16 to 23 students
Large Classes = 24 to 31 students

[4]The multiple grade level teacher was not included in the total.

[5]Low to Average Student Achievement = 1-60 National Percentile
High Student Achievement = 61-99 National Percentile
The kindergarten and first grade teachers are not included.

ASSESSMENT STANDARDS BY LEVEL OF QUESTION, ASSESSMENT FORMAT, MANIPULATIVE MATERIALS, AND METHOD OF SCORING

The results presented below are related to the second research question. Research Question 2: Do K-3 teachers vary after training in the methods of implementing alternative assessment strategies by assessment standard? The assessment standards of the 1989 Standards for School Mathematics include Mathematical Power, Problem Solving, Communications, Reasoning, Mathematical Concepts, Mathematical Procedures, and Mathematical Concepts, Mathematical Procedures, and Mathematical Disposition.

Level of Questions

When the proportion of higher level questions was compared with the proportion of knowledge level questions, there were significant differences for each assessment standard analyzed. As shown in Table 10, the project teachers used higher level items significantly more frequently when assessing the standard of Problem Solving ($t=3.82$, $p=.001$) and the standard of Communications ($t=2.73$, $p=.010$) when compared with the use of knowledge level items. The reverse occurred when assessing the standards of Mathematical Concepts and Mathematical Procedures. The proportion of knowledge level items was significantly greater when assessing Mathematical Concepts ($t=4.11$, $p=.001$) and Mathematical Procedures ($t=-5.11$, $p=.001$) than was the proportion of higher level items. Therefore, the standard measured was an important factor in the item levels.

Descriptive statistics were used to identify patterns in the teachers' classroom assessment practices according to the assessment standards of the 1989 Standards for School Mathematics. As shown in Table 11, the most frequently measured standards included Mathematical Concepts (standard 5; .56) and Mathematical Procedures (standard 6; .20). Mathematical Power (standard 1) and Mathematical Disposition (standard 7) were not assessed at all by the teachers. When the project teachers measured Problem Solving (standard 2) and Reasoning (standard 4), the items were all higher level questions.

Assessment Format

The results of the analyses of the teachers' assessment instruments, by assessment standard and assessment format, are displayed in Table 12. The results indicate that there was a significant difference when the proportion of alternative assessment mat items was compared with the proportion of traditional assessment format items for the assessment standard of Mathematical Procedures (t=-3.80, p=.001). When assessing Mathematical Procedures, the teachers used items representing traditional assessment formats significantly more frequently than when alternative assessment format items were used. There were no significant differences in the proportion of items using alternative and traditional assessment formats for the standards of Problem Solving (t=-1.16, p=.255), Communications (t=-.50, p=.617), and Mathematical Concepts (t=1.56, p=.129). Thus, the project teachers did vary significantly in the proportion of items using alternative and traditional assessment formats for the standard of Mathematical Procedures but not for the standards of Problem Solving, Communications, and Mathematical Concepts.

The proportion of items using each specific assessment format by assessment standard is presented in Table 13. The teachers used traditional formats on 53 percent of the items while alternative formats were used on 48 percent of the items. Overall, free response items were used with the greatest frequency and when assessing Problem Solving (standard 2), Communications (standard 3), and Mathematical Procedures (standard 6). The teachers used the demonstration format most frequently when assessing Mathematical Concepts (standard 5). Thus, the pattern of using the assessment formats varied by the assessment standards.

Manipulative Materials

The results of the analyses on the proportion of items using manipulative materials, stratified by the assessment standards, are presented in Table 14. A significant difference was found in the proportion of items using manipulative materials when compared to the proportion of items not using manipulative materials for the standard of Mathematical Concepts (t=-4.13, p=.001). The project teachers used items with manipulative materials significantly more frequently when assessing Mathematical Concepts. The standards of Problem Solving, Communications, and Mathematical Procedures did not differ significantly in the proportion of items using and not using manipulative materials.

The proportion of items using each manipulative material according to the assessment standards is summarized in Table 15. The teachers used such a variety of manipulative materials that no clear pattern of using specific manipulative materials by standard appeared. Overall, the teachers used unifix cubes (.12), base-ten blocks (.08), and counters (.06) with the greatest frequency. The other specific manipulatives were used infrequently and were grouped together as other manipulatives.

Scoring Method

The teachers' assessment items were analyzed in relation to the assessment standards and the method of scoring. The results are displayed in Table 16. Scoring method differed significantly when the project teachers were assessing the Mathematical Procedures standard (t=-5.58, p=.001). The project teachers used items employing traditional scoring methods significantly more frequently as compared to items employing alternative scoring methods when assessing Mathematical Procedures. There were no significant differences in item usage across scoring methods for the standards of Problem Solving (t=.22, p=.828), Communications (t=.99, p=.327), and Mathematical Concepts (t=-1.32, p=.196).

Table 17 includes the mean proportion of items using traditional scoring, alternative scoring, and each scoring method. Traditional scoring was used on 59 percent of the items and alternative scoring was used on 41 percent of the items. The pattern of scoring method use varied according to the standard assessed. Items using right/wrong scoring were used most frequently when measuring the standards of Mathematical Concepts (standard 5; .32) and Mathematical Procedures (standard 6; .17). When measuring Communications (standard 3), the project teachers used items with general impression scoring most frequently. Analytic and focused holistic scoring were used infrequently for each standard.

Content

Table 18 includes the mean proportion of items according to the mathematics content assessed and by the assessment standards. Overall, the algorithms were assessed most frequently for each standards. The Mathematical Concepts standard was measured with the greatest variety of content areas.

TABLE 10

PROPORTION OF ITEMS BY LEVEL OF QUESTION AND ASSESSMENT STANDARD
N=33

Standard	Item Level	Mean Proportion	df	t	p
Problem Solving	Higher Level Items	.094	32	3.82	.001*
	Knowledge Level Items	.000			
Communication	Higher Level Items	.111	32	2.73	.010*
	Knowledge Level Items	.019			
Mathematical Concepts	Higher Level Items	.174	32	-4.11	.001*
	Knowledge Level Items	.388			

TABLE 10 (cont.)

Mathematical Procedures	Higher Level Items	.011	32	-5.11	.001*
	Knowledge Level Items	.190			

*Significant at the .05 alpha level.

TABLE 11

MEAN PROPORTION OF TEST ITEMS BY ASSESSMENT STANDARD AND LEVEL OF QUESTION
N=33

Item Level	Standard							
	1	2	3	4	5	6	7	Total
Higher Level Items	.00	.09	.11	.01	.17	.01	.00	.40
Knowledge Level Items	.00	.00	.02	.00	.39	.19	.00	.60
Total	.00	.09	.13	.01	.56	.20	.00	1.00

Standard 1 Mathematical Power
Standard 2 Problem Solving
Standard 3 Communications
Standard 4 Reasoning
Standard 5 Mathematical Concepts
Standard 6 Mathematical Procedures
Standard 7 Mathematical Disposition

TABLE 12

PROPORTION OF ALTERNATIVE AND TRADITIONAL ASSESSMENT FORMAT ITEMS BY ASSESSMENT STANDARD
N=33

Standard	Assess. Format	Mean Proportion	df	t	p
Problem Solving	Altern. Assess. Formats	.032	32	-1.16	.255
	Trad. Assess. Formats	.064			
Communication	Altern. Assess. Formats	.056	32	-.50	.617
	Trad. Assess. Formats	.074			
Mathematical Concepts	Altern. Assess. Formats	.325	32	1.56	.129
	Trad. Assess. Formats	.236			
Mathematical Procedures	Altern. Assess. Formats	.046	32	-3.80	.001*

TABLE 12 (cont.)

Trad. Assess. Formats	.156

TABLE 13

MEAN PROPORTION OF TEST ITEMS BY ASSESSMENT STANDARD AND ASSESSMENT FORMAT
N=33

Assessment Format	Standard							
	1	2	3	4	5	6	7	Total
Alternative Assessment Formats								
Demonstration	.00	.02	.01	.00	.24	.02	.00	.29
Journal	.00	.01	.04	.00	.02	.00	.00	.07
Oral	.00	.00	.01	.01	.07	.03	.00	.11
Alternative Total	.00	.03	.06	.01	.33	.05	.00	.48
Traditional Assessment Formats								
Forced Choice	.00	.00	.01	.00	.06	.02	.00	.09
Free Response	.00	.06	.07	.00	.18	.13	.00	.44
Traditional Total	.00	.06	.08	.00	.24	.15	.00	.53
Total	.00	.09	.14	.01	.57	.20	.00	1.00

Standard 1 Mathematical Power
Standard 2 Problem Solving
Standard 3 Communications
Standard 4 Reasoning
Standard 5 Mathematical Concepts
Standard 6 Mathematical Procedures
Standard 7 Mathematical Disposition

TABLE 13

MEAN PROPORTION OF TEST ITEMS BY ASSESSMENT STANDARD AND ASSESSMENT FORMAT

N=33

Assessment Format	Standard							
	1	2	3	4	5	6	7	Total
Alternative Assessment Formats								
Demonstration	.00	.02	.01	.00	.24	.02	.00	.29
Journal	.00	.01	.04	.00	.02	.00	.00	.07
Oral	.00	.00	.01	.01	.07	.03	.00	.11
Alternative Total	.00	.03	.06	.01	.33	.05	.00	.48
Traditional Assessment Formats								
Forced Choice	.00	.00	.01	.00	.06	.02	.00	.09
Free Response	.00	.06	.07	.00	.18	.13	.00	.44
Traditional Total	.00	.06	.08	.00	.24	.15	.00	.53
Total	.00	.09	.14	.01	.57	.20	.00	1.00

Standard 1 Mathematical Power
Standard 2 Problem Solving
Standard 3 Communications
Standard 4 Reasoning
Standard 5 Mathematical Concepts
Standard 6 Mathematical Procedures
Standard 7 Mathematical Disposition

TABLE 14

PROPORTION OF ITEMS USING AND NOT USING MANIPULATIVE MATERIALS BY ASSESSMENT STANDARD N=33

Standard	Use of Manip.	Mean Proportion	df	t	p
Problem Solving	Use of Manip.	.056	32	-.83	.412
	Non-Use of Manip.	.040			
Communication	Use of Manip.	.060	32	.27	.789
	Non-Use of Manip.	.070			
Mathematical Concepts	Use of Manip.	.417	32	-4.13	.001*
	Non-Use of Manip.	.142			
Mathematical Procedures	Use of Manip.	.079	32	-1.37	.181
	Non-Use of Manip.	.122			

icant at the .05 alpha level.

TABLE 15

MEAN PROPORTION OF TEST ITEMS BY ASSESSMENT STANDARD AND MANIPULATIVE MATERIALS
N=33

Manipulative Materials	Standard							
	1	2	3	4	5	6	7	Total
Base-Ten Blocks	.00	.02	.00	.00	.03	.03	.00	.08
Counters	.00	.01	.00	.00	.03	.02	.00	.06
Unifix Cubes	.00	.01	.00	.00	.08	.03	.00	.12
Other Manipulatives	.00	.01	.05	.00	.29	.04	.00	.39
Use of Manipulatives	.00	.05	.05	.00	.43	.12	.00	.65
Non-Use of Manipulatives	.00	.04	.07	.00	.14	.08	.00	.33

Standard
1 Mathematical Power
2 Problem Solving
3 Communications
4 Reasoning
5 Mathematical Concepts
6 Mathematical Procedures
7 Mathematical Disposition

TABLE 16

PROPORTION OF ITEMS USING ALTERNATIVE AND TRADITIONAL SCORING METHODS BY ASSESSMENT STANDARD N=33

Standard	Scoring Method	Mean Proportion	df	t	p
Problem Solving	Altern. Scoring Methods	.050	32	.22	.828
	Trad. Scoring Methods	.046			
Communication	Altern. Scoring Methods	.082	32	.99	.327
	Trad. Scoring Methods	.048			
Mathematical Concepts	Altern. Scoring Methods	.239	32	-1.32	.196
	Trad. Scoring Methods	.322			
Mathematical Procedures	Altern. Scoring Methods	.028	32	-5.58	.001*

TABLE 16 (cont.)

Trad. Scoring Methods	.173

*Significant at the .05 alpha level.

TABLE 17

MEAN PROPORTION OF TEST ITEMS BY ASSESSMENT STANDARD AND METHOD OF SCORING
N=33

Scoring Method	Standard							
	1	2	3	4	5	6	7	Total
Alternative Scoring Methods								
Analytic	.00	.02	.01	.00	.00	.01	.00	.04
Focused Holistic	.00	.00	.00	.00	.03	.00	.00	.03
General Impression	.00	.03	.07	.01	.21	.02	.00	.34
Alternative Total	.00	.05	.08	.01	.24	.03	.00	.41
Traditional Scoring Methods								
Right/Wrong	.00	.05	.05	.00	.32	.17	.00	.59
Traditional Total	.00	.05	.05	.00	.32	.17	.00	.59
Total	.00	.10	.13	.01	.56	.20	.00	1.00

Standard 1 Mathematical Power
Standard 2 Problem Solving
Standard 3 Communications
Standard 4 Reasoning
Standard 5 Mathematical Concepts
Standard 6 Mathematical Procedures
Standard 7 Mathematical Disposition

TABLE 18

MEAN PROPORTION OF TEST ITEMS BY ASSESSMENT STANDARD AND CONTENT
N=33

Content	Standard							
	1	2	3	4	5	6	7	Total
Algorithms	.00	.06	.06	.00	.09	.18	.00	.39
Geometry	.00	.01	.00	.00	.08	.00	.00	.09
Number Concepts	.00	.02	.01	.00	.15	.00	.00	.18
Other Concepts	.00	.02	.07	.01	.25	.01	.00	.36
Total[1]	.00	.11	.14	.01	.57	.19	.00	1.02

[1]Total may not sum to 1.00 due to rounding.

Standard
1 Mathematical Power
2 Problem Solving
3 Communications
4 Reasoning
5 Mathematical Concepts
6 Mathematical Procedures
7 Mathematical Disposition

EVALUATION INFORMATION GAINED BY ALTERNATIVE AND TRADITIONAL ASSESSMENT STRATEGIES

The third research question examined the difference in the evaluative information gained by an application of alternative assessment strategies when compared with traditional assessment techniques. The results are presented below.

Research Question 3: Is there a difference in the evaluative information gained by an application of alternative assessment strategies when compared with traditional assessment techniques?

The Assessment Questionnaire (Appendix 2) was administered to the teachers in May, 1991. The questionnaire was returned by 30 teachers (88% return rate).

Content

It was anticipated that the teachers would identify some of the primary mathematics content areas as better matched with alternative assessment formats and some content as better matched with unit tests from the textbook. These tests represent a traditional method of assessment. Questions 7 and 8 from the Assessment Questionnaire addressed these issues.

The content identified by the project teachers as better matched with alternative assessment procedures is listed in Table 19. The most frequent responses of the teachers included measurement and time (18 teachers), geometry (11 teachers), problem solving (7 teachers), counting money and change (7 teachers), graphing (7 teachers), fractional parts (7 teachers), and place value (6 teachers). The focus of the content areas identified by the teachers were mathematical concepts rather than mathematical computation and could, therefore, best be classified as non-procedural.

TABLE 19

CONTENT AREAS BETTER MATCHED TO ALTERNATIVE ASSESSMENT FORMATS
N=30

Content Area	Number of Teachers*	Percent of Teachers
Measurement and time	18	60%
Geometry	11	37%
Problem Solving	7	23%
Counting money and change	7	23%
Graphing	7	23%
Fractional parts	7	23%
Place Value	6	20%
Reasoning	4	13%
Communications	3	10%
Word problems	3	10%
Addition and subtraction with trading	3	10%
Symmetry	2	7%
Patterns	2	7%
Counting	2	7%
Sorting	2	7%
Perimeter	1	3%
Decimals	1	3%
Spacial sense	1	3%
Area	1	3%
Concepts	1	3%
All content in the kindergarten curriculum	1	3%

*Subjects may have identified more than one content area.

The teachers identified the content listed in Table 20 as better matched to the unit tests. In contrast with the content areas matched with alternative formats, the teachers identified mathematical procedures as better matched with the unit tests. The most frequently listed content areas included basic facts (8 teachers), basic algorithms (8 teachers), addition (6 teachers), and subtraction (6 teachers). Therefore, the teachers matched different content areas with the unit tests and alternative assessment formats.

TABLE 20

CONTENT AREAS BETTER MATCHED TO UNIT TESTS
N=30

Content Area	Number of Teachers*	Percent of Teachers
Basic facts	8	27%
Basic Algorithms	8	27%
Addition	6	20%
Subtraction	6	20%
Computation	4	13%
Multiplication	2	7%
Time	1	3%
Numeration	1	3%
Division	1	3%
Money	1	3%
Estimation	1	3%
Rounding	1	3%

*Subjects may have identified more than one content area.

Evaluative Information

Question 11 from the Assessment Questionnaire was used to determine the evaluative information gained by alternative assessment procedures and the unit tests. The teachers were asked to indicate which assessment procedure (alternative assessment or unit tests) was the more appropriate format for each of the 1989 Assessment Standards. The teachers' responses are reported in Table 21.

The teachers responded that alternative assessment or both alternative assessment and unit tests could appropriately be used to assess Mathematical Power, Problem Solving, and Mathematical Disposition. The Communications standard and the Reasoning standard, according to the project teachers (83% and 63% respectively), were felt to be more appropriately measured by alternative assessment formats.

The Mathematics Procedures and Mathematical Concepts standards were reported by a majority of the teachers (60% and 53% respectively) to be appropriately measured by both format types. Alternative assessment formats were selected by a larger percentage of teachers when compared to the unit tests for each standard with the exception of Mathematical Procedures. As indicated on Table 21, the teachers feel that alternative assessment formats and the unit tests are both generally appropriate to assess the standards with neither playing a clearly dominate role.

TABLE 21

FREQUENCY AND PERCENT OF TEACHERS: APPROPRIATE ASSESSMENT PROCEDURES FOR EACH ASSESSMENT STANDARD N=30

Standard	Altn. Assmt. Formats	Unit Tests	Both Altn. Assmt. Formats and Unit Tests	Blank	Total
Mathematical Power	16 53%	0 0%	12 40%	2 7%	30 100%
Problem Solving	12 40%	1 3%	15 50%	2 7%	30 100%
Communications	25 83%	0 0%	3 10%	2 7%	30 100%
Reasoning	19 63%	0 0%	9 30%	2 7%	30 100%
Mathematical Concepts	8 27%	2 7%	18 60%	2 7%	30 100%
Mathematical Procedures	2 7%	10 33%	16 53%	2 7%	30 100%
Mathematical Disposition	12 40%	0 0%	16 53%	2 7%	30 100%

Confidence in Evaluation Information

Question 9 and question 10 from the Assessment Questionnaire were used to address the teachers' confidence in the evaluation information gained through alternative assessment procedures and unit tests.

The amount of confidence the teachers had in the evaluation information gained through alternative assessment formats and the unit tests are presented in Table 22. A majority of the teachers (53%) responded that they had moderate confidence in the evaluation information from alternative assessment, with 20% indicating total confidence. None of the project teachers reported no confidence in the evaluation information gained from alternative assessment formats.

The confidence in the evaluation information gained through the units tests was most frequently reported by the teachers as high (40%) with 3% of the teachers indicating total confidence and 0% reporting no confidence.

TABLE 22

CONFIDENCE IN THE EVALUATION INFORMATION GAINED THROUGH ALTERNATIVE ASSESSMENT FORMATS AND UNIT TESTS
N=30

	1 None	2 Little	3 Moderate	4 High	5 Total	Blank
Alternative Assessment Formats	0 0%	0 0%	16 53%	8 27%	6 20%	0 0%
Unit Tests	0 0%	4 13%	8 27%	12 40%	1 3%	5[1] 17%

[1]The kindergarten teachers do not have unit tests and left the question blank.

The level of confidence in the information gained through alternative assessment and unit tests was analyzed using a correlated means *t*-test. The results, presented in Table 23, indicate no significant difference in the level of teacher confidence in the information gained through alternative assessment formats in comparison to that gained through unit tests (t=0.00, p=1.00). Overall, the teachers were equally confident the information gained through alternative assessment tests and unit tests.

TABLE 23

COMPARISON OF THE CONFIDENCE OF INFORMATION FROM ALTERNATIVE ASSESSMENT FORMATS AND UNIT TESTS

	N	Mean	*df*	*t*	*p*
Alternative Assessment Format	25	3.40	24	0.00	1.00
Unit Tests	25	3.40			

Scale
1 = No Confidence
2 = Little Confidence
3 = Moderate Confidence
4 = High Confidence
5 = Total Confidence

Frequency of Use of Alternative Formats and Unit Tests

Question 1 and question 2 from the Assessment Questionnaire were used to determine the frequency which alternative formats and unit tests were used. The frequency with which the teachers used alternative assessment formats and the unit tests available in the mathematics text book series is reported in Table 24. A majority of the teachers (53%) indicated that they used the unit tests regularly with the remaining teachers reporting less frequent use. In contrast, only 17% of the teachers reported regularly using alternative assessment formats.

TABLE 24

FREQUENCY OF USING ALTERNATIVE ASSESSMENT FORMATS AND UNIT TESTS
N=30

	1 Rarely or not at all	2 Some	3 A lot	4 Regularly	Total[1]
Alternative Assessment Formats	5 17%	14 47%	6 20%	5 17%	30 101%
Unit Tests	5 17%	5 17%	4 13%	16 53%	30 100%

[1]Total may not sum to 100 due to rounding.

Scale

1	Rarely or Not at all	0% - 20%
2	Some	21% - 50%
3	A lot	51% - 80%
4	Regularly	81% - 100%

A correlated means t-test was performed on the teachers' frequency of using alternative assessment formats when compared to their frequency of using the unit tests. The results are presented in Table 25. There was a significant difference in the frequency of using alternative assessment formats and unit tests by the teachers (t=-4.82, p=.001). The unit tests were found to be used significantly more frequently than were the alternative assessment formats.

TABLE 25

COMPARISON OF THE FREQUENCY OF USING ALTERNATIVE ASSESSMENT FORMATS AND UNIT TESTS

	N	Mean	df	t	p
Alternative Assessment Formats	30	2.367	29	-4.82	.001*
Unit Tests	30	3.033			

*Significant at the .05 alpha level.

Scale

1	Rarely or Not at all	0% - 20%
2	Some	21% - 50%
3	A lot	51% - 80%
4	Regularly	81% - 100%

Difficulty Using Alternative Formats and Unit Tests

Question 3 and question 5 from the Assessment Questionnaire were used to address the difficulty using alternative formats and unit tests. The degree of difficulty experienced by the teachers in using alternative assessment formats and the unit tests is summarized in Table 26. The most frequent (53%) response of the teachers indicated that they perceived some difficulty in using alternative assessment formats. Alternative assessment was reported as very easy to use by 7 percent of the teachers and very difficult to use by 20 percent of the teachers.

In contrast, a majority of the project teachers (67%) found the unit tests to be very easy to use. The remaining teachers indicated that the unit tests were somewhat easy to use (13%) or very difficult to use (7%). The kindergarten textbook series does not have unit tests, therefore, the kindergarten teachers did not respond to the difficulty in using the unit tests.

TABLE 26

DIFFICULTY USING ALTERNATIVE ASSESSMENT FORMATS AND UNIT TESTS
N=30

	1 Very Easy	2 Somewhat Easy	3 Somewhat Difficult	4 Very Difficult	Blank	Total
Alternative Assessment Formats	2 7%	5 17%	16 53%	6 20%	1 3%	30 100%
Unit Tests	20 67%	4 13%	0 0%	2 7%	4[1] 13%	30 100%

[1]The kindergarten teachers do not have unit tests and left the question blank.

The difficulty of using alternative assessment formats was compared with the difficulty of using the unit tests using a correlated means t-test. The results of the t-test, as shown in Table 27, indicate a significant difference in the perceived difficulty of using alternative formats and unit tests ($t=9.38$, $p=.001$). The project teachers found the alternative assessment formats to be significantly more difficult to use than the unit tests.

TABLE 27

COMPARISON OF THE DIFFICULTY USING ALTERNATIVE ASSESSMENT FORMATS AND UNIT TESTS

	N	Mean	df	t	p
Alternative Assessment Formats	26	2.769	25	9.38	.001*
Unit Tests	26	1.384			

*Significant at the .05 alpha level.

Scale
1 = Very Easy
2 = Somewhat Easy
3 = Somewhat Difficult
4 = Very Difficult

The Assessment Questionnaire, items 4 and 6, also addressed the type of problems the teachers experienced when using alternative assessment formats and the unit tests. The problems with alternative assessment, identified by the teachers, are summarized in Table 28. The most frequently identified concern centered on the time (15 teachers) it took to assess students using alternative formats. Additional problems included grading (12 teachers) and the process to develop alternative formats (8 teachers).

TABLE 28

PROBLEMS USING ALTERNATIVE ASSESSMENT FORMATS
N=30

Problem	Number of Teachers*	Percent of Teachers
Time	15	50%
Grading	12	40%
Creating the Assessment Instruments	8	27%
None	5	17%

*Subjects may have identified more than one problem.

The teachers also identified problems using the unit tests. As shown in Table 29, the most common response was that they did not experience any problems in using the unit tests (15 teachers). Ten teachers identified problems with validity. The issues related to validity concerned the lack of assessment of the thought process, the difficulty in assessing an understanding of concepts, and the lack of consistency between the way skills were taught and tested.

TABLE 29

PROBLEMS USING UNIT TESTS
N=30

Problem	Number of Teachers*	Percent of Teachers
None	15	50%
Validity Issues	10	30%

*Subjects may have identified more than one problem.

CONCLUSIONS

IMPLEMENTATION OF ALTERNATIVE ASSESSMENT ACCORDING TO TEACHER GRADING ORIENTATION, CLASS SIZE, GRADE LEVEL, AND STUDENT ACHIEVEMENT LEVEL

Grading orientation, class size, and student achievement did not associate with the implementation of alternative assessment. There was a significant difference in the use of items with manipulative materials according to grade level. The kindergarten and first grade teachers used items with manipulative materials significantly more frequently than did the second and third grade teachers. The pattern of classroom assessment practices did not vary according to the teacher stratification variables. The following patterns were found:

1. Knowledge level questions were used with the greatest frequency.
2. The most frequently used formats were free response items followed by demonstration items. Journal, oral, and forced choice formats were used infrequently.
3. The most frequently used manipulative materials included unifix cubes, base ten blocks, and counters. The project teachers used a great variety of manipulative materials.
4. Right/wrong scoring followed by general impression scoring were the most frequently used methods. Analytic and focused holistic scoring were used infrequently.
5. The algorithms, number concepts, and geometry were assessed most frequently by the K-3 Mathematics Specialist teachers.

The degree to which knowledge level questions were used by the teachers can be compared with the findings of other studies. Mehrens and Lehmann (1987) and Stiggins, Conklin, and Bridgeford (1986) found 80% of teacher made questions to be at the knowledge level. The K-3 Mathematics Specialist teachers' rate of 60 percent is less frequent than found in these other studies (Mehrens & Lehmann, 1987; Stiggins, Conklin, & Bridgeford, 1986).

The 1989 Standards for School Mathematics advocate an increased use of higher level items but do not define a standard with which to establish a goal (National Council of Teachers of Mathematics, 1989). If past research findings can serve as a comparison, the project teachers demonstrated a higher rate of using higher level test items. In addition, the students in the primary grades are mastering the concepts and knowledge of mathematics. Therefore, it may be expected that the primary grades would include a greater focus on knowledge level skills. An issue that remains unclear includes the degree to which these teachers increased their use of higher level questions as a result of participation in the K-3 Mathematics Specialist Project. Without preproject data, this cannot be explored.

The use of manipulative materials in conjunction with assessment was a common practice. The K-3 Mathematics Specialist teachers used manipulative materials on an average of 66 percent of the test items. The 1989 Standards for School Mathematics encouraged an increased use of manipulative materials (National Council of Teachers or Mathematics, 1989) thus, the fact that two-thirds of the assessment items, on the average, included manipulative materials would seem to indicate a substantial incorporation of manipulative materials into the assessment process.

The significantly greater use of manipulative materials with classroom assessment in kindergarten and first grade concurred with the findings of Gilbert and Bush (1988) which indicated that the use of manipulative materials decreased as grade level increased. The 1989 Standards for School Mathematics encouraged the use of manipulatives in all grade levels (National Council of Teachers of Mathematics, 1989). They suggested that as new concepts are presented, a progression should be made from the concrete to the pictorial and then the abstract. New concepts are presented in each grade level; therefore, the use of manipulative materials is appropriate in the higher grade levels as well as the lower grade levels. Also, empirical evidence, provided by Baroody (1989), Sowell (1989), and Suydam and Higgins (1977) showed that student achievement is greater when manipulative materials are included in the lesson.

The significant difference by grade level in the use of manipulative materials indicates that the second and third grade teachers have not implemented this aspect of alternative assessment to the same degree as the kindergarten and first grade teachers. A possible explanation for this difference may be that the kindergarten and first grade teachers used manipulative materials extensively prior to participating in the K-3 Mathematics Specialist Project. Although the second and third grade teachers did not reach the same rate of use as did the kindergarten and first grade teachers, their current rate of use may represent an increase over their prior use. Without preproject data, this cannot be explored.

The use of alternative formats (demonstration, oral, and journal) averaged 47 percent of the test items, which is a substantial proportion of the test items. The 1989 Standards for School Mathematics suggest using a variety of assessment formats (National Council of Teachers of Mathematics, 1989) but the selection of the format should be congruent with the content, subjects, and information needs of the teacher (Carey, 1988). The degree to which the teachers used the alternative assessment formats of demonstration, oral, or journal (47%) indicates wide use of a variety of alternative assessment formats.

The teachers used the alternative scoring methods of analytic scoring, focused holistic scoring, and general impression scoring on an average of 41 percent of the test items. This rate of use indicates that the project teachers did follow the intent of the standards which encouraged the use of alternative scoring.

Therefore, the K-3 Mathematics Specialist teachers did implement alternative assessment techniques, as encouraged by the 1989 Standards for School Mathematics, in the areas of item level, use of manipulative materials, assessment format, and scoring method.

The lack of significant differences in the assessment practices of the teachers according to teacher grading orientation, class size, and achievement level of their students seems to indicate that these factors, which should not affect classroom assessment, did not. The assessment practices of teachers should be based on an appropriate match between the content and students rather than teacher characteristics.

ASSESSMENT STANDARDS BY LEVEL OF QUESTIONS, ASSESSMENT FORMAT, MANIPULATIVE MATERIALS, AND METHODS OF SCORING

The implementation of alternative assessment was found to vary significantly for some of the alternative assessment variables according to the assessment standards and not with other variables and standards. The only consistent relationship between alternative assessment and standard was with the item level. The following relationships were identified.

1. Higher level items were included in the assessment instruments significantly more frequently when assessing Problem Solving and Communications. Knowledge levels items were used significantly more frequently when assessing Mathematical Concepts and Mathematical Procedures.
2. Items representing traditional assessment formats were used significantly more frequently than were items representing alternative assessment formats when measuring Mathematical Procedures.
3. When assessing Mathematical Concepts, items using manipulatives were used significantly more frequently than when manipulatives were not used.
4. Traditionally scored items were used significantly more frequently than were alternatively scored items when assessing Mathematical Procedures.
5. The algorithms were the most frequently measured content for each standard.

The standard of Mathematical Concepts was measured, on the average, by more than one-half of the items contained in the assessment portfolios of the project teachers. DeMana and Waits (1990) suggest that more classroom time should be devoted to the development of mathematical concepts rather than computation. The assessment of mathematical concepts occurred so frequently by the project teachers, possibly indicating that their focus was not on computation but rather on concepts as DeMana and Waits and the 1989 Standards for School Mathematics have encouraged.

There were significant differences in the use of higher level questions when compared with the use of knowledge level questions on each standard that was assessed with a sufficient frequency to conduct a *t*-test. The standards of Problem Solving and Communications, where higher level questions were used significantly more frequently, appear to be more aligned with higher level than with lower level items. When assessing Problem Solving the teachers were generally interested in assessing the ability of the students to use mathematics in a practical way. The definition of Problem Solving includes such terms as "formulate," "apply," "verify," and "generalize." These would include assessment items at the higher level rather than the lower level. Thus, the significantly greater use of higher level items in the Problem Solving standard appears to be an indication that the project teachers matched the assessed concepts and item level.

The Communications standard requires the student to write about mathematics, using its symbols and mathematical terms; thus, the students must understand the language of mathematics to use it for communication. The definition of the Communications standard includes the terms "understand," "interpret," "evaluate," "use," and "model" which would require higher level items. As with Problem Solving, the Communications standard definition is focused at the higher level; therefore, the significantly greater use of higher level items may indicate that the teachers are appropriately measuring the standard.

The definition of the Mathematical Concepts standard is focused at both the knowledge level and higher level. Despite the fact that knowledge level items were used significantly more frequently than were higher level items, 30 percent of the Mathematical Concepts items were higher level items.

The Mathematical Procedures standard is also focused at both knowledge and higher level items, but many of the higher level procedure items could possibly be classified as measuring a different standard. For example, an item where the student explained why to use addition could be classified as a Communications item rather than Mathematical Procedures and an item where the student used addition to solve a problem could be classified as a Problem Solving item. Therefore, the significantly greater use of knowledge level items in Mathematical procedures may have occurred because the standards do not represent discrete mathematic skills. There is an overlap in the skills included in many of the standards and may account for the significantly greater use of knowledge level items when assessing Mathematical Procedures.

The significantly greater use of traditional assessment formats when assessing Mathematical Procedures may be related to the nature of the standard. The definition of the Mathematical Procedures standard includes the recognition of appropriate procedures and the execution of procedures. Although alternative assessment formats can be used to assess these student outcomes, the traditional formats of forced choice and free response are a very efficient and effective assessment format to measure these skills. With the limited classroom time available for mathematics instruction and assessment, there may not be a need to assess these Mathematical Procedures outcomes using the more time consuming alternative formats. The teachers did use alternative assessment formats. An average of 25 percent of the Mathematical Procedures items were alternative formats compared with 75 percent of the items that were traditional formats. Therefore, these results may indicate that the teachers were judicious in their use of alternative assessment formats, selecting the most appropriate format to assess the targeted student outcome.

The most frequently used format when assessing Problem Solving and Communications was free response. Norris (1989) and Stiggins (1982) have suggested that alternative formats are appropriate when measuring problem solving (Norris, 1989) and communications (Stiggins, 1982). The assessment of communications using alternative and traditional formats was found by Stiggins (1982) to assess different aspects of communication leading to the conclusion that alternative formats are necessary to fully assess communications. The limited use of alternative formats by the project teachers may indicate that the teachers did not have sufficient training to develop alternative assessment items for Problem Solving and Communications.

Although alternative assessment formats were used for each standard assessed, the distribution of alternative and traditional assessment formats for the Problem Solving and Communications standards may need further time for item development. The use of alternative and traditional assessment formats for the Mathematical Concepts and Mathematical Procedures standards may be appropriate.

The 1989 Standards for School Mathematics encouraged the use of manipulative materials for each standard. The significantly greater use of manipulatives with the Mathematical Concepts standard indicates that the project teachers did implement this aspect of alternative assessment. The Mathematical Concepts standard includes models and concept properties. This standard definition appears to incorporate the use of manipulative materials while the other standards, by definition, do not include such a dependence on manipulative materials. These results seem to illustrate that the K-3 Mathematics teachers used manipulative materials when the standard suggested their use. Thus, the Mathematical Concepts standard definition may explain the significantly greater use of manipulative materials.

To further demonstrate that the project teachers used manipulative materials when indicated, manipulatives had limited use when assessing Communications. This standard includes demonstrating mathematical ideas which could include manipulative materials but the primary focus is on speaking, writing, and visually depicting mathematical ideas. Therefore, the standards where manipulative materials are appropriate, Mathematical Concepts, Problem Solving, and Mathematical Procedures, the use of manipulative materials was extensive. Overall, the use of manipulative materials was a common practice.

The Mathematical Procedures standard incorporates the computational aspects of mathematics. The significantly greater use of traditional scoring (right/wrong) for this standard appears to be an appropriate application of the scoring methods to student outcome. The more time consuming alternative scoring methods do not lend themselves readily to the scoring of computation problems at these early grade levels.

When assessing Problem Solving and Mathematical Concepts the teachers used right/wrong scoring with the greatest frequency. These standards are those where the alternative scoring approaches of focused holistic scoring and analytic scoring are most appropriate due to their focus on the thought process (Charles, Lester, O'Daffer, 1988). The infrequent use of alternative scoring approaches may possibly indicate that the teachers do not fully understand the application of alternative scoring to Problem Solving and Mathematical Concepts. Overall, the teachers used right/wrong scoring most frequently.

In summary, there was no clear pattern of using alternative assessment for any one standard or group of standards. Consequently, no conclusions concerning the implementation of alternative assessment can be drawn in relation to the 1989 Standards for School Mathematics. Traditional assessment procedures (knowledge level items, traditional formats, and traditional scoring) were used significantly greater when assessing Mathematical Procedures than were alternative assessment procedures.

It is noteworthy that the standards of Mathematical Power and Mathematical Disposition were not assessed by any items included in the project teachers' assessment portfolios. Also, the Reasoning standard was assessed infrequently. Possibly, the teachers did not have a clear understanding of these standards or of procedures for assessing these standards. The definitions and parameters for Mathematical Power and Mathematical Disposition are less well defined. In addition, these traits are less observable, possibly accounting for their reduced emphasis in the project teachers' assessment portfolios.

EVALUATION INFORMATION GAINED BY ALTERNATIVE AND TRADITIONAL ASSESSMENT STRATEGIES

The third research question examined the difference in the evaluative information gained by an application of alternative assessment strategies when compared with traditional assessment techniques. The issues that were explored include content areas where alternative or traditional assessment was more appropriate, the type of evaluative information available through alternative and traditional assessment, and the difference in teacher confidence in the evaluative information from alternative and traditional assessment. In addition, the frequency which alternative assessment formats were used and the relative difficulty of implementing alternative assessment was examined.

The content areas identified by the teachers as more related to alternative assessment included mathematics concepts and problem solving while the content areas identified as better matched with the unit tests (traditional assessment) were mathematical procedures. These identified contents appear to indicate that the project teachers have an understanding of when alternative assessment procedures and traditional assessment procedures are appropriate. The forced choice formats are very effective and efficient methods to measure mathematical procedures while the alternative assessment procedures may give the teachers more in-depth information concerning concept development and problem solving (Norris, 1989). Therefore, the teachers identified different content areas where alternative assessment and the unit tests are more appropriate.

confidence level of the teachers in the evaluation information gained by alternative assessment was more frequently reported as moderate for the alternative assessment formats and high for the unit tests. Although the mean confidence ratings were not different at a level of statistical significance, the lower teacher confidence in the evaluation information gained from alternative assessment than from traditional assessment may reflect their lack of refined skills in developing and using alternative assessment procedures.

Alternative assessment formats were found to be more difficult to use than the unit tests. The project teachers indicated problems in using alternative assessment as time, grading, and creating the instruments. The teachers found the unit tests easier to use but identified concerns related to test validity. The unit tests are prepared for the teachers and included in the textbook series. Grading is simply a matter of determining the percent correct while implementing alternative assessment frequently requires the teacher to develop the instrument and to administer the instrument individually or in a small group. These alternative procedures all require time, of which teachers have little. Without well-developed criteria for scoring, defending the assigned grades was a concern for teachers, possibly resulting in the reduced use of alternative assessment. Published instruments including alternative formats with developed holistic scoring criteria, as readily available as the unit tests, may solve some of the difficulty of use issues.

RECOMMENDATIONS

1. Due to the time required to develop alternative assessment procedures for the classroom, there is a need for published instruments, as readily available as forced choice and free response instruments, to be included in the mathematics textbook series for teachers to use and modify for their specific classroom setting.
2. Due to the widespread use of manipulative materials by the project teachers, efforts should be made to provide a greater number of classroom teachers with appropriate manipulative materials and inservice training on their use.
3. The relative lack of using analytic and focused holistic scoring may indicate that the teachers do not have a sufficient understanding of the assessment and scoring of students' cognitive processes. The training model may benefit by modification in this area to strengthen the teachers' understanding of the assessment and scoring of problem solving.
4. The project teachers should be monitored over a period of time to determine if their use of alternative assessment strategies increases and to determine if student achievement is positively impacted when the teachers reach a higher degree of alternative assessment usage.
5. The standards of Mathematical Disposition and Mathematical Power were not assessed by any project teacher. The vagueness of the definitions of these standards

may have contributed to a limited understanding of the standards and of methods to measure these standards. In addition, these standards reflect a more affective component of mathematics and the role of these attributes in classroom assessment is controversial. The National Council of Teachers of Mathematics may improve the implementation of these assessment standards by developing materials to increase the teachers' understanding and role of these standards.

REFERENCES

Ashworth, K.P. (1990). Math and science: A nation still at risk. *Principal, 69 (3),* 15-17.

Baroody, A.J. (1989). One point of view; Manipulatives don't come with guarantees. *Arithmetic Teacher, 37* (2), 4-5.

Berlak, H. (1985). Testing in a democracy. *Educational Leadership, 43* (2), 16-17.

Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). *Taxonomy of Educational Objectives, The Classification of Educational Goals. Handbook 1. Cognitive Domain.* New York: David McKay Company, Inc.

Campbell, P.F. & Bamberger, H.J. (1990). Implementing the standards: The vision of problem solving in the standards. *Arithmetic Teacher, 37* (9), 14-17.

Carey, L.M. (1988). *Measuring and Evaluating School Learning.* Boston: Allyn and Bacon, Inc.

Carpenter, T.P. (1985). How children solve simple word problems. *Education & Urban Society, 17* (4), 417-425.

Charles R., Lester, F., & O'Daffer, P. (1988). *How to evaluate process in problem solving.* Reston, Va: National Council of Teachers of Mathematics.

Clark, D.J., Clark, D.M., & Lovitt, C.J. (1990). Changes in mathematics teaching call for assessment alternatives. *1990 National Council of Teachers of Mathematics Yearbook* (pp. 118-129). Reston, Va: The Council.

Cooney, T.J. (1988). The issue of reform: What have we learned from yesteryear? *Mathematics Teacher, 81* (5), 352-363.

Crosswhite, F.J., Dossey, J.A., & Frye, S.M. (1989). NCTM standards for school mathematics: Visions for implementation. *Mathematics Teacher, 82* (8), 664-671.

DeMana, F. & Waits, B.K. (1990). The role of technology in teaching mathematics. *Mathematics Teacher, 83* (1), 27-31.

Dossey, J.A. (1989). Transforming mathematics education. *Educational Leadership, 47* (3), 22-24.

Frary, R.B. (1985). Multiple-choice versus free-response: A simulation study. *Journal of Educational Measurement, 22* (1), 21-31.

Freeman, D.J., Kuh, T.M., Porter, A.C., Floden, R.E., Schmidt, W.H., & Schwille, J.R. (1983). Do textbooks and tests define a national curriculum in elementary school mathematics? *The Elementary School Journal, 83* (5), 501-513.

Gilbert, R.K. & Bush, W.S. (1988). Familiarity, availability & use of manipulative devises in mathematics at the primary level. *School Science and Mathematics, 88* (6), 459-469.

Grice, G.L. & Jones, M.A. (1989). Teaching thinking skills: State mandate and the K-12 curriculum. *The Clearing House, 62* (8), 337-341.

Guthrie, J.T. (1984). Testing higher level skills. *Journal of Reading, 28* (2), 188-190.

Guthrie, J.T. & Lissitz, R.W. (1985). A framework for assessment-based decision making in reading education. *Educational Measurement: Issues and Practice, 4* (2), 26-30.

Hall, B.W., Carroll, D., & Comer, C.B. (1988). Test use among classroom teachers and its relationship to teaching level and teaching practices. *Applied Measurement in Education, 1* (2), 145-156.

Hall, B.W., Villeme, M.G., & Phillippy, S.W. (1985). How beginning teachers use test results in critical education decisions. *Educational Research Quarterly, 9* (3), 12-18.

Haney, W. (1985). Making testing more educational. *Educational Leadership, 43* (2), 4-13.

Heid, M.K. (1988). Calculators on tests-One giant step for mathematics education. *Mathematics Teacher, 81* (9), 710-713.

Kameenui, E.J. & Griffin, C. (1989). National crisis in verbal problem solving in mathematics: A proposal for examining the role of basal mathematics programs. *The Elementary School Journal, 89* (5), 575-593.

Kolstad, R.K., Briggs, L.D., & Kolstad, R.A. (1990). Multiple-choice rules prevent the selection of wrong options in examinations. *Education, 110* (3), 360-363.

Kouba, V.L., Brown, C.A., Carpenter, T.P., Lindquist, M.M., Silver, E.A. & Swafford, J.O. (1988). Results of the fourth NAEP assessment of mathematics: Measurement, geometry, data interpretation, attitudes, and other topics. *Arithmetic Teacher, 35* (9), 10-16.

Loif, M.M., Carey, D.A., Carpenter, T.P., & Fennema, E. (1988). Integrating assessment and instruction. *Arithmetic Teacher, 36* (3), 53-55.

Mehrens, W.A. & Lehmann, I.J. (1987). Using teacher-made measurement devises. *NASSP Bulletin, 71* (496), 36-44.

National Council of Teachers of Mathematics, Commission on Standards for School Mathematics. (1989). *Curriculum and Evaluation Standards for School Mathematics.* Reston VA.: Author.

Norris, S.P. (1989). Can we test validly for critical thinking? *Educational Researcher, 18* (9), 21-26.

Norris, S.P. (1990). Effect of eliciting verbal reports of thinking on critical thinking test performance. *Journal of Educational Measurement*, 21(1), 41-58.

Otis, M.J. & Offerman, T.R. (1988). Assessing for learning: How do you evaluate problem solving? *Arithmetic Teacher, 35* (8), 49-51.

Peck, D.M., Jencks, S.M., & Connell, M.L. (1989). Improving instruction through brief interviews. *Arithmetic Teacher, 37* (3), 15-17.

Peterson, P.L., Carpenter, T. & Fennema, E. (1989). Teacher's knowledge of student's knowledge in mathematics problem solving: Correlational and case analysis. *Journal of Educational Psychology, 81* (4), 558-569.

Presseisen, B.Z. (1986). *Thinking skills: Research and practice.* Washington: NEA Professional Library, National Education Association.

Quellmalz, E. (1985). Needed: Better methods for testing higher-order thinking skills.

Educational Leadership, 43 (2), 29-35.
Robinson, G.E. (1987). Assessing for learning. *Arithmetic Teacher, 35* (1), 33.
Romberg, T.A. (1988). NCTM's curriculum and evaluation standards: What they are and why they are needed. *Arithmetic Teache r,* 35 (9), 2-3.
Rudman, H.C. (1987). Classroom instruction and tests: What do we really know about the link? *NASSP Bulletin, 71* (496), 3-22.
Salmon-Cox, L. (1981). Teachers and standardized achievement tests: What's really happening? *Phi Delta Kappan, 62* (9), 631-634.
Sax, G. (1980). *Principles of Educational and Psychological Measurement and Evaluation.* Belmont, CA: Wadsworth, Inc.
Schalock, D., Fielding, G., Schalock, M., Erickson, J., & Brott, M. (1985). Integrating teaching and testing with program management. *Educational Leadership, 43* (2), 55-58.
Schoen, H.L., Blume, G., & Hoover, H.D. (1990). Outcomes and processes on estimation test items in different formats. *Journal for Research in Mathematics Education, 21* (1), 61-73.
Shriner, J. & Salvia, J. (1988). Chronic non-correspondence between math curricula and arithmetic tests. *Exceptional Child, 55* (3), 240-248.
Siegler, R.S. (1989). Strategy diversity and cognitive assessment. *Educational Researcher, 18* (9), 15-20.
Silver, E.A. (1990). Contributions of Research to Practice: Applying Findings, Methods, and Perspectives. *1990 Yearbook, National Council of Teachers of Mathematics* (pp. 1-11). Reston, VA: National Council of Teachers of Mathematics.
Sowell, E.J. (1989). Effects of manipulative materials in mathematics instruction. *Journal for Research in Mathematics Education, 20* (5), 498-505.
Stanic G.M. & Kilpatric, J. (1988). Historical perspectives on Problem Solving in the Mathematics Curriculum. *The Teaching and Assessing of Mathematics Problem Solving* (pp. 1-22). Reston, VA: The National Council of Teachers of Mathematics.
Sternberg, R. & Baron, J.B. (1985). A statewide approach to measuring critical thinking skills. *Educational Leadership, 43* (2), 40-43.
Stiff, L.V. (1988). Problem solving by example. *School Science and Mathematics, 88* (8), 666-675.
Stiggins, R.J. (1982). A comparison of direct and indirect writing assessment methods. *Research in the Teaching of English, 16* (2), 101-114.
Stiggins, R.J. (1985). Improving assessment where it means the most: In the classroom. *Educational Leadership, 43* (2), 69-74.
Stiggins, R.J. & Bridgeford, N.J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement, 22* (4), 271-286.
Stiggins, R.J., Conklin, N.F., & Bridgeford, N.J. (1986). Classroom assessment: A key to effective education. *Educational Measurement: Issues and Practice, 15* (2), 5-17.

Suydam, M.N. & Higgins, J.L. (1977). *Activity-based learning in elementary school mathematics: Recommendations for research.* Columbus, OH: Center for Science, Mathematics & Environment Education, College of Education, Ohio State University. (ERIC Document Reproduction Service).

Thompson, C.S. & Rathmell, E.C. (1988). NCTM's Standards for School Mathematics, K-12. *Arithmetic Teacher, 35* (9), 17-19.

Wardrop, J.L., Anderson, T.H. Hively, W., Hasting, C.N., Anderson, R.I., & Muller, K.E. (1982). A framework for analyzing the reference structure of educational achievement tests. *Journal of Educational Measurement, 19* (1), 1-18.

Wearne, D. & Hiebert, J. (1985). Teaching for thinking in mathematics. *Curriculum Review, 25* (1), 65-68.

Webb, N. & Briars, D. (1990). Assessment in Mathematics Classrooms K-8. *1990 Yearbook, National Council of Teachers of Mathematics* (pp. 108-117). Reston, VA: National Council of Teachers of Mathematics.

Willoughby, S.S. (1990). Three promising developments in testing. *Educational Leadership, 47* (7), 91-91.

Worthen, B.R. & Spandel, V. (1991). Putting the standardized test debate in perspective. *Educational Leadership, 48* (5), 65-69.

APPENDIX 1

GRADING ORIENTATION QUESTIONNAIRE

Name ____

Please indicate how much weight you typically give to each of the following elements in determining the academic grade for the students in your mathematics class.

Use this key:

1 - No weight; not considered at all

2 - Minimal weight

3 - Moderate weight

4 - Moderately heavy weight

5 - Heavy weight; considered one of the primary factors

Mathematics academic grades based on:

______ Students' achievement on post-tests (summative assessment using any measurement method)

______ Students' achievement on seatwork or homework (formative assessment using any measurement method)

______ Students' completion of extra credit work

______ Students' attitude

______ Students' effort

______ Students' motivation

______ Students' classroom participation

______ Students' classroom behavior (adherence to classroom rules)

APPENDIX 2

ASSESSMENT QUESTIONNAIRE

Name ____

1. Overall, how frequently did you use alternative assessment procedures for grading in your mathematics classroom?

1	2	3	4
Used Rarely or Not at All (0%-20%)	Used Some (21%-50%)	Used a Lot (52%-80%)	Used Regularly (81%-100%)

2. Overall, how frequently did you use the unit tests for grading in your mathematics classroom?

1	2	3	4
Used Rarely or Not at All (0%-20%)	Used Some (21%-50%)	Used a Lot (52%-80%)	Used Regularly (81%-100%)

3. Please rate the degree of difficulty you experienced using alternative assessment procedures in your mathematics classroom?

1	2	3	4
Very Easy	Somewhat Easy	Somewhat Difficult	Very Difficult

4. What type of problems, if any, did you experience when using alternative assessment procedures?

5. Please rate the degree of difficulty you experienced using the unit tests in your mathematics classroom?

1	2	3	4
Very Easy	Somewhat Easy	Somewhat Difficult	Very Difficult

6. What type of problems, if any, did you experience when using the unit tests?

7. Were there content areas (i.e., specific concepts, procedures, skills) better suited to alternative assessment procedures than to unit tests?

If so, which content areas?

8. Were there content areas (i.e., specific concepts, procedures, skills) better suited to the unit tests than to alternative assessment procedures?

If so, which content areas?

9. Please rate the amount of confidence that you have in the evaluation information gained through alternative assessment procedures.

1	2	3	4	5
No Confidence	Little Confidence	Moderate Confidence	High Confidence	Total Confidence
I don't believe the information describes what the student knows				I believe the information describes what the student knows

APPENDIX 2 (cont.)

10. Please rate the amount of confidence that you have in the evaluation information gained through unit tests.

1	2	3	4	5
No Confidence	Little Confidence	Moderate Confidence	High Confidence	Total Confidence
I don't believe the information describes what the student knows				I believe the information describes what the student knows

89

APPENDIX 2 (cont.)

11. Please put a check in the box that indicates the type of assessment procedure that you feel is the more appropriate for each assessment standard.

Assessment Standard	Alternative Assessment Procedures	Unit Tests	Both Alternative Assessment Procedures and the Unit Tests are appropriate
Mathematical Power - the integration of mathematical knowledge.			
Problem Solving - applying strategies for problem solving.			
Communications - expressing mathematical ideas by speaking, writing, demonstrating, or visually.			
Reasoning - inductive and deductive reasoning.			
Mathematical Concepts - understanding mathematical concepts.			
Mathematical Procedures - executing mathematical computations.			
Mathematical Disposition - mathematics confidence.			

BOARD OF DIRECTORS

Mary Kay Habgood - 1990-93
PRESIDENT
215 Manatee Avenue West
Bradenton, FL 34205

Mel Lucas - 1990-93
PAST PRESIDENT
620 E. University Avenue
Gainesville, FL 32601

Betty Hurlbut - 1992-95
TREASURER
426 School Street
Sebring, FL 33870

Cecil Carlton - 1992-95
P.O. Box 1470
Pensacola, FL 32597

Jack McAfee - 1991-94
P.O. Box 2648
Vero Beach, FL 32960

Lee Rowell - 1990-93
445 W. Amelia Street
Orlando, FL 32802

Tom Conner - 1991-94
P.O. Box 787
LaBelle, FL 33935

Esther Orteiza - 1990-93
P.O. Box 670
Ocala, FL 32670

John Headlee - 1991-94
1507 West Main Street
Inverness, FL 32650

Sandra McDonald - 1992-95
40 Orange Street
St. Augustine, FL 32084

Jayne Hartman - 1992-95
2909 Delaware Avenue
Ft. Pierce, FL 34947

Adrian Cline - 1991-94
530 La Solona Avenue
Arcadia, FL 33821

ADVISORS

Jake Beard
College of Education
Florida State University
Tallahassee, FL 32306

William Castine
College of Education
Florida A & M University
Tallahassee, FL 32307

Charles Hayes
Central Florida Community College
P.O. Box 1388
Ocala, FL 32678

Theresa Vernetsen
College of Education
University of Florida
Gainesville, FL 32611

Lee A. Shiver
Florida Education Center
Department of Education
325 W. Gaines Street, Suite 414
Tallahassee, FL 32399

Allen Fisher
College of Education
Florida International University
Tamiami Trail
Miami, FL 33199

Dan Morris
College of Education
Florida Atlantic University
Boca Raton, FL 33431

Carl Balado
College of Education
University of Central Florida
Orlando, FL 32816

Bette Soldwedel
College of Education
University of North Florida
Jacksonville, FL 32216

Phyllis NeSmith
Florida School Board
Association
P.O. Box 446
Nocatee, FL 33864

John Follman
College of Education
University of South Florida
4204 Fowler Avenue
Tampa, FL 33620

Sam Mathews
Educational Research &
Development Center
University of West Florida
Pensacola, FL 32514

Charlie T. Council
EXECUTIVE DIRECTOR
P.O. Box 506
Sanibel, FL 33957

COUNCIL MEMBERS

County	Address	Contact Person
Alachua	620 E. University Ave. Gainesville, FL 32601	Mel Lucas
Charlotte	1445 Piatti Drive Punta Gorda, FL 33948	John Wiegman
Citrus	1507 West Main Street Inverness, FL 32650	John Headlee
Collier	3710 Estey Avenue Naples, FL 33942	Roger Otten
DeSoto	530 LaSolona Avenue Arcadia, FL 33821	Adrian Cline
Dixie	Post Office Box 4-V Cross City, FL 32628	Lloyd Jones
Escambia	Post Office Box 1470 Pensacola, FL 32597	Cecil Carlton
Glades	Post Office Box 459 Moore Haven, FL 33471	Gary Clark
Hamilton	Post Office Box 1059 Jasper, FL 32952	Carol Blair
Hardee	Post Office Drawer 678 Wauchula, FL 33873	Derrell Bryan
Hendry	Post Office Box 787 LaBelle, FL 33953	Tom Conner
Hernando	919 North Broad Street Brooksville, FL 34601	Phyllis McIntyre

COUNCIL MEMBERS

County	Address	Contact Person
Highlands	426 School Street Sebring, FL 33870	Betty Hurlbut
Hillsborough	Post Office Box 3408 Tampa, FL 33601	John Hilderbrand
Indian River	Post Office Box 2648 Vero Beach, FL 32960	Jack McAfee
Lee	2055 Central Avenue Ft. Myers, FL 33901	Mike Jones
Madison	Post Office Box 449 Madison, FL 32340	Roy W. Scott
Manatee	215 Manatee Avenue W. Bradenton, FL 34205	Mary K. Habgood
Marion	Post Office Box 670 Ocala, FL 32670	Esther Oteiza
Martin	Post Office Box 1049 Stuart, FL 33494	Deana Hughes
Nassau	1201 Atlantic Avenue Fernandina Beach, FL 32034	James T. Northey
Okeechobee	100 S.W. 5th Avenue Okeechobee, FL 33472	Danny Mullins
Orange	445 West Amelia Street Orlando, FL 32802	Lee Rowell
Palm Beach	3910 RCA Blvd., Suite 1011 Palm Beach Gardens, FL 33410	Marc Baron

COUNCIL MEMBERS

County	Address	Contact Person
Pasco	7227 U.S. Highway 41 Land O'Lakes, FL 33537	Madhabi Banerji
St. Johns	40 Orange Street St. Augustine, FL 32048	Sandra McDonald
St. Lucie	2909 Delaware Avenue Ft. Pierce, FL 34947	Jayne Hartman
Santa Rosa	603 Canal Street Milton, FL 32570	Frank Green
Sarasota	2418 Hatton Street Sarasota, FL 33577	Rick Nations
Suwannee	224 Parshly Street Live Oak, FL 32060	Marvin Johns
Taylor	Post Office Box 509 Perry, FL 32347	Lester Padgett
Union	55 S.W. 6th Street Lake Butler, FL 32054	Ray Dukes
Wakulla	Post Office Box 100 Crawfordville, FL 32327	Jan Putnal

NOTES