

DOCUMENT RESUME

ED 431 012

TM 029 844

AUTHOR Belle, Dana
TITLE Traditional Assessment versus Alternative Assessment.
PUB DATE 1999-05-00
NOTE 58p.; Master's Thesis, Kean University of New Jersey.
PUB TYPE Dissertations/Theses - Masters Theses (042)
EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Educational Assessment; *Elementary School Students;
Elementary School Teachers; Grade 5; Intermediate Grades;
Mathematics; Reading; *Student Evaluation; *Test Format;
Test Use
IDENTIFIERS *Alternative Assessment; Scoring Rubrics

ABSTRACT

The purpose of this study was to determine whether a teacher can use one type of assessment to evaluate students' abilities fairly. The question is whether or not alternative assessment strategies are necessary to meet students' individual needs. The research, conducted with 28 fifth-grade students, compared their traditional and alternative reading and mathematics scores. Surveys were also distributed to 20 teachers and 100 students. The results indicate that the two types of testing cannot be compared a majority of the time, indicating a need for both types of assessment. The survey findings suggest that teachers and students are individuals who all need various types of assessment. The findings support the hypothesis. Six appendixes contain the teacher and student surveys, traditional scoring grade books, and the mathematics and reading rubrics. (Contains 7 tables and 25 references.) (Author/SLD)

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

Traditional Assessment versus Alternative Assessment

By:
Dana Belle

Presented in Partial Fulfillment of the Requirements for the Degree of Master of Arts

Kean University of New Jersey

May 1999

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Dana Belle

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

TM029844

ABSTRACT

The purpose of this study was to determine whether a teacher can utilize one type of assessment to fairly evaluate students' abilities. The question is whether or not alternative assessment strategies are necessary to meet students' individual needs. The research was conducted with 28 fifth grade students. Their traditional and alternative reading and math scores were compared. Surveys were also distributed to 20 teachers and 100 students. The results indicated that the two types of testing cannot be compared a majority of the time, indicating a need for both types of assessments. The surveys suggested the idea that teachers and students are individuals who all need various types of assessing. The hypothesis was supported.

ACKNOWLEDGMENTS

I would like to thank Dr. Albert J. Mazurkiewicz for his guidance and support throughout this study.

I would also like to thank my husband, Christopher Belle, for his expertise of computer knowledge and husbandry.

And lastly, I would like to thank my family for taking care of Connor when I was working on this paper.

DEDICATION

This paper is dedicated to my family: Christopher, Connor, Duncan, Logan, Barbara, Frank, Dawn, and Chipmunk who encouraged and supported me, tried to keep quiet around me, and were patient when I was stressed!

Without my family, this experience would not have been completed. For this reason and a million others, I dedicate this paper and all of my love to you.

TABLE OF CONTENTS

	Page Numbers:
I. Abstract	i
II. Acknowledgments	ii
III. Dedication	iii
IV. List of Tables	iv
V. Traditional Assessment vs. Alternative Assessment	
Introduction	1-5
Hypothesis	5
Procedure	5-10
Results	10-13
Discussion	13-14
VI. Assessment Procedures: Related Research	15-34
VII. References	35-37
VIII. Appendices	
A: Teacher Survey	
B: Student Survey	
C: Math Tradition	
D: Math Rubric	
E: Reading Traditional	
F: Reading Rubric	

LIST OF TABLES

	Page Numbers:
1. Grades: Standard and Rubric	10
2. Percentage of Teachers Utilizing Specific Testing Methods	11
3. Teachers' Beliefs	11
4. Results of 100 Student Surveys on Frequency of Test Types	12
5. 100 Student's Test Type Preferences	12
6. Mean, Standard Deviation, and Correlation of Mathematic Scores	13
7. Mean, Standard Deviation and Correlation of Reading Scores	13

Research has told us that students learn in various ways. Therefore, we can assume that students will need to be tested in a variety of ways in order to be fair to the individual. Traditional ways of assessing have been widely accepted in the past. But today, many teachers feel that alternative testing is the most accurate way of evaluating students' higher order thinking skills. The question is whether or not we need alternate methods to evaluate a student's learning and if we do, which methods are most effectively being used today.

Traditional assessment can be defined as evaluations that include standardized and classroom achievement tests with mostly closed-ended items, such as true/false, multiple choice, and fill-in-the blanks. In the past, these types of tests were mostly used. But recently, there has been a growing trend towards alternative assessment. This type of evaluation includes any critical thinking or higher order skills. The techniques utilized are performance based assessments, observation techniques, student self assessment, and portfolios. (Bol, 1998)

In all classrooms around the world, assessment is being used. In the United States today, the focus seems to be on authentic assessment rather than traditional. As described by Rick Gordon, in order for authentic learning to be utilized successfully in a classroom, a number of elements

must be understood. Students must actively be involved in solving problems, which should involve a variety of decisions and concepts. Students must be able to work cooperatively and the activity needs to be meaningful to the students. The problem must be “real” or relevant to their lives. And the activities must be linked to the entire curriculum. Gordon emphasizes that unless a student sees why the work needs to be done, he will not retain and therefore, cannot apply the skill to his life. (Gordon, 1998)

In agreement with Gordon was research done by Steele and Arth in 1998. These researchers also stated that activities must be meaningful to the students in order for knowledge to be retained. One of the conclusions reached was that memorization is not showing the students why they need to know the skill and, therefore, will not maintain the information after the test. An example for alternative assessment that was suggested is the mastery learning approach . These are goals that the teacher and individual student prepare, then strive to fulfill. When the content is learned, alternative assessments are performed with the teacher to test the student’s knowledge. These can include informal dialogue, such as asking oral questions, observing the student, and having him demonstrate, and journal writing. (Steele, 1998)

A 1995-1996 study in a Southern urban school district supported the idea that teachers use a variety of methods to evaluate their students. Research has shown that all teachers, with or without professional training, do use alternative methods when grading their students. The report also concluded that the personality of the teacher, experience, and grade level is an altering factor in how a teacher evaluates. The more confidence a teacher has, the more willing they are to try new ideas. This confidence comes with experience. Depending on how a teacher feels about alternative assessment will play a large part in predicting what she will use to assess performance. The study also found that the most frequent types of assessment utilized were observation and performance tasks. It was also found that elementary teachers use alternative types of testing more frequently than the high school teachers. One reason for this is the element of time. A high school teacher has only fifty minutes a day with her students compared to an elementary teacher's all day. This time allotment makes it easier to observe and complete various projects. In conclusion, this research study stated that alternative assessment has been used in most classrooms though it was not always called this. (Bol, 1998)

It was also hypothesized that students' study habits are influenced by the type of tests used. For example, in a study completed by Crooks in 1988, he concluded that a student's study habits were influenced by the types of assessments used. If the class had to memorize specific facts, that is what was done. And if essay type questions were frequently given, students prepared for that type of testing. The teachers felt that the essay or short answer tests were more difficult to grade and therefore given less to the students. Higher order questions were used less frequently regardless of achievement or motivation levels of the students. Teachers failed to accentuate higher order thinking skills to their students for a variety of reasons. The reasons include teacher's own misunderstanding of a higher order skill, lack of materials, and perceptions that only the smart children will understand. It was concluded that students are not being challenged to think on a higher level as much as they should. (Fleming, 1998)

Although alternative assessment seems to have many beneficial advantages for the students, there are some disadvantages as well. In a perfect world, class sizes would be small with no behavior problems, and a community would be enthusiastic and support financially. Unfortunately this is not always so. And sometimes, even when it is, there are still problems when applying this type of assessment.

For example, in 1991 Vermont became the first state to use portfolios in a statewide assessment program. This required grades four through eight to have portfolios in mathematics and composition. There were also “uniform tests” given to fourth and eighth graders which consisted of a writing sample and various mathematical problems. Conclusions reported were positive and negative. Critics did “acknowledge that measurable improvement has occurred in classroom instruction, especially in the teaching of problem-solving strategies.” But, statewide results were inconsistent with each other. The results of the individuals were each unique to their school and therefore, unreliable to compare with others intellectually. In the second year of implementation, more training and support was given throughout the year to the teachers. The portfolios were more useful to the teachers and students when used for individual instruction and improvements. In conclusion, Vermont decided that portfolios are more useful when there are no specific state educational standards. (Nidds, 1997)

Alternative assessments definitely have advantages for the students and teachers in any classroom. This type of assessment allows a deeper look into a student’s intellect and then, can help improve the way he or she is taught. Unfortunately, teachers are required to work with state standards

and students are responsible for knowing the information taught. The only way to test their knowledge is in some form of assessment. Which types of assessment is open to question.

Hypothesis:

For the purpose of this study, it was hypothesized that teachers cannot use one type of assessment to fairly evaluate a student's abilities but must utilize alternate assessment strategies to meet student's individual needs.

Procedure

The researcher conducting this study was a fifth grade teacher in the building where the data was to be collected. She was able to gather data quite easily due to her school position. Within the course of the study, data was collected through surveys from students and teachers. The researcher then observed, tested, and altered her assessing methods without halting the learning process of her students.

The research took place in an urban area at a fifth and sixth grade school. There was 11 fifth grade and 11 sixth grade classes with an average of 28 students per class. The assessment strategies in the building were non uniform and therefore all teachers can utilize different methods. Report cards are distributed four times a year with grades of A through F. Training

has begun this year for alternative assessment using rubrics. The GOALS performance based measure of achievement tests were given to the students for the first time in September. This test, by Harcourt Brace Jovanowich Inc., is supposed to replace the Metropolitan Achievement tests by next year. The GOALS test is graded using rubrics and consists of ten free response questions in reading and mathematics to test higher order thinking skills. During the 1998- 1999 school year, the teachers in this school are to be trained to teach towards the GOALS test. Basically, the teacher and students at this school are being immersed with alternative assessing in the everyday classroom along with the traditional methods used in the past.

The teachers were given a survey (Appendix A) to fill out on the types of assessment they use, which they enjoy the most, and which they feel is the most beneficial to the students. The students were also given a survey (Appendix B) asking which kinds of tests they receive the most of, which ones they enjoy or find the easiest, and the most beneficial. A random sample of student surveys distributed between the fifth and sixth grade classes to a total of 100 was collected. No names were placed on any of the surveys.

The researcher teaches all subjects - mathematics, social studies, science, reading, spelling, and english. The classes switch for reading and

mathematics homogeneously. The researcher is a four year veteran at this school teaching fifth grade. She has altered her assessment methods through the years to suit her and her students needs due to academic and behavioral reasons. The teachers in the school can use traditional methods, such as textbook tests and dittos and/or alternative homemade assessments, such as science experiments or crafts. During the past for years, she has used both types of methods equally. But this year, the administration has been pushing the rubric scales and therefore, all teachers need to attempt alternative methods more than usual.

Scores of her students were recorded using the traditional 100 point scale and rubric 3 to 0 scale. The researcher has learned the behavior of her students and knows what they are academically capable of. The question is whether poor students can increase their scores and knowledge using alternative methods. And how do the honor roll students compare ? Scores will be recorded and at the end of the marking period traditional and alternative grades will be averaged separately to see if extreme differences are seen.

The researcher has assumed that students and teachers are more familiar with the traditional types of testing, and therefore will rely and utilize them more often than alternative methods. Assumptions can be made

that all students in this building have had similar assessment experiences. And we can assume that there are students who are extremely bright while there are those who are academically challenged.

There are a few limitations in this in this research as well. The surveys will reflect the differences in the testing types for each class. Teachers and their personalities alter the way they teach and test. There will be contrast in teaching experience and training of the teachers. The school in which the study was taking place requires that all teachers use all types of methods. Although it is never enforced by the administrator, the teachers know to complete both. And because of the GOALS testing being put in place, assumptions can be made that all teachers are being trained. But this is not so in this school. Only a five teachers have been formally trained for the first five training sessions on alternative assessment and then another five teachers will be chosen for the second half of the sessions. The ten total sessions are different and probably should be shared with all teachers in the school. Due to this difference, all teachers are trained differently - if at all.

Data was collected through the surveys and the grading of the researcher's students. The surveys were collected and the information will be ranked according to the highest numbers. The grading from the

researcher was averaged and then compared using a point system which is mandated by the school. The following table illustrates the traditional and rubric scoring.

Table 1

GRADES: STANDARD AND RUBRIC

ACADEMIC GRADE	STANDARD	RUBRIC
A	100 - 90	3.00 - 2.60
B	89 - 80	2.50 - 1.60
C	79 - 70	1.50 - .60
D	69 - 0	.50 - 0

Results

The results of the questionnaires were organized by the major research question. The first section addressed the kinds of assessments teachers reported using in their classrooms and how often they were used. The results appear in Table 2.

Table 2

PERCENTAGE OF TEACHERS UTILIZING SPECIFIC TESTING METHODS

METHOD	ALWAYS	SOMETIMES	SELDOM	NEVER
MULTIPLE CHOICE	2	15	3	0
TRUE / FALSE	3	12	4	1
FILL IN THE BLANK	4	16	0	0
SHORT ESSAY	6	14	0	0
ESSAY	4	11	5	0
MATCHING	0	10	10	0
PROJECT	2	14	4	0
TEXT BOOK MADE TESTS	2	18	0	0
YOUR OWN MADE TESTS	2	18	0	0
RUBRIC	0	8	12	0
TRADITIONAL GRADES	8	12	0	0

Many of the items were used “Sometimes”, allowing the children the benefit of all types of testing. And due to the unfamiliarity with rubric scoring, it was used the most seldom.

Table 3 examines the teachers ideas of what a student would find easy, most beneficial, or enjoyable in their class.

Table 3

TEACHER'S BELIEFS

BELIEF	RANK		
	1	2	3
EASY	MULTIPLE CHOICE	TRUE / FALSE	FILL IN THE BLANK
BENEFICIAL	SHORT ANSWER	ESSAY	PROJECTS
ENJOYABLE	PROJECTS	MULTIPLE CHOICE	MATCH

While it is interesting to note how the teachers feel about assessing and which are most frequently used, it is also alluring to note the students' perception. Table 4 and 5 show how 100 students feel about their assessment process.

Table 4

RESULTS OF 100 STUDENT SURVEYS ON FREQUENCY OF TEST TYPES

TYPE	ALWAYS	SOME	SELDOM	NEVER
MULTIPLE CHOICE	38	62	0	0
TRUE / FALSE	0	22	71	7
FILL IN THE BLANK	14	62	24	0

Table 5

100 STUDENT'S TEST TYPE PREFERENCES

FAVORITE	# OF STUDENTS	LEAST	# OF STUDENTS
PROJECTS	44	ESSAY	45
MULTIPLE CHOICE	36	FILL IN THE BLANK	30
FILL IN THE BLANK	20	MATCHING	25

Appendixes C and D are a compilation of the students' math scores in traditional and rubrics. Appendixes E and F list the reading scores.

Table 6 shows the math scores using the traditional and rubric scoring system. There was correlation of .62, a moderate correlation. The commonality of 38.44 indicated that only 38% of the time will these scores are assessing the students the same way.

Table 6

MEAN STANDARD DEVIATION AND CORRELATION OF MATH SCORES

SAMPLE	MEAN X	STANDARD	r	r
REGULAR	79.14	10.64	0.62	38.44
RUBRIC	1.45	0.42		

Table 7

MEAN, STANDARD DEVIATION, AND CORRELATION OF READING SCORES

SAMPLE	MEAN X	STANDARD	r	r
REGULAR	79.26	16.04	0.26	0.0676
RUBRIC	1.62	0.57		0.07%

As can be seen in Table 7 there was a low correlation of .26 between the regular and rubric procedures for assessing reading. The commonality of .07% indicates that the two assessment procedures have little in common.

Discussion

The hypothesis of this study was that teachers cannot use one type of assessment to fairly evaluate a student's abilities but must utilize alternate assessment strategies to meet student's individual needs. The results of this study indicate that this was true, and therefore, the hypothesis was correct.

Though slight differences appear when comparing scores, as in Table 6 and 7, the correlations were low to moderate. The results also showed that the two types of tests do not measure the same goal. In mathematics, only 38% of the time were rubrics and traditional scores comparable, in reading, only .07%

The surveys conducted also supported the hypothesis. Teachers utilize various types of techniques and feel that is necessary for success. Students enjoy and benefit from different types of assessing according to their surveys. Results indicate the need for variety.

Educators and administrators need to work together for the benefit of the students in order for complete success to be achieved. Although there is a growing trend towards alternative assessment, educators must not forget the previous methods used - which were also successful.

**ASSESSMENT PROCEDURES :
RELATED RESEARCH**

In order to understand why alternative assessment is in debate, we must understand why teachers need to assess their students.

According to researchers Lester and Kroll (1991), we must evaluate our students for a number of reasons. How can a teacher teach effectively unless she evaluates prior knowledge? Evaluation shows the teacher how much her students know so she can alter her lessons appropriately. By evaluating, a teacher can learn her students' ideas, beliefs, and attitudes so she can prepare them to be better humanitarians. A student needs to know that the work he does is of value. And, finally, a reason to collect data is due to the constricting guidelines of the State - grades are necessary for accountability.

How does a parent feel about assessment in their child's classroom? Cynthia Garnett (1992) has written an article explaining why she feels, that as a parent, standardized testing is not always appropriate in a classroom. Throughout the school year, teachers teach towards the constructivist method, which develop higher order skills. But, in the spring, teachers and schools are forced to become positivists and teach towards the standardized tests and must change gears in order for their students to achieve high scores. Ms. Garnett feels that "assessment needs to put value on students' work, be more than a test, and a continuous, dynamic, and often informal

process which best appreciates that children's achievement is tied to their sense of themselves and their sense of their place in this world.”

Standardized testing can also be a positive tool for evaluators argues Walsh and Betz (1985). We can assume the following:

- * All standardized test words have similar meanings to all students.
- *Individual's test behavior is consistent over time
- *The test measures what it is supposed to measure
- *An individual's test scores will be equal to his true score, plus the error

Fred Newmann (1996) also agrees with Ms. Garnett's ideas about standardized testing. He feels that “authentic assessment often emphasizes certain procedures of assessment considered more likely to elicit a more complex intellectual performance than the standardized norm referenced test.” And if this is true, students will benefit from authentic assessment ideals and gain higher order skills in all areas of study.

In 1993, the National Education Goals Panel constructed the Malcolm Report or “Promises to Keep: Creating High Standards for American Students”. This report tried to define content standards and performance

standards in the United States. According to the Panel, content standards specify “what students should know and be able to do” while performance standards specify “how good is good enough.”

The Panel agreed that assessments are imperative in giving the schools accurate information for planning curriculum, grading, and allow us to be fair in categorizing the students. But they felt that today’s assessments provide misinformation about the individual student, classrooms, schools, and districts. So although the Panel agrees that some sort of assessment is needed, today’s assessments are imperfect and needs to be fine tuned in order to be fair to all students and schools in the United States.

According to the Goals Panel, performance standards include performance tasks and traditional tests when assessing students. The report stated that “ A performance standard indicates both the *nature of the evidence* (such as an essay, mathematical proof, scientific experiment, project, exam, or combination of these) required to demonstrate that the content standard has been met and the *quality of student performance* that will be deemed acceptable (that merits a passing of “A” grade)”. The performance standard can also be defined as a distinct score on a rubric for a specific task.

Marc Tucker, co-director of the Standards Project, (1992) stated that unless a task is given to the students and then completed, a performance cannot be assessed. Content standards will then become the target for creating assessment.

Critics of the Goals Panel argued that performance tasks are too specific. In order to fairly assess a child, 10 to 36 performance tasks are necessary for each subject area. This is usually not done in our schools. Richard Shavelson (1998) agrees with this notion that if a child can do one question correctly, he might not be able to do another correct. Therefore, many items are needed to appropriately test the child.

Traditional testing is the second most used assessment tool in the United States. The number of items assessing a particular topic are between 3 and 11. Although this is a fair number, it is still not a complete picture of the students' abilities claims the critics.

Utilizing traditional or authentic test data appropriately is a concern many have. According to Mike Schmoker, author of the book, *Results*, analyzing test data should be a primary method to plan the school year's curriculum. School districts need to look at the test data and change their curriculum to fit the needs of the students. But most schools find that if they follow this idea of changing their curriculum with each individual

grade, the students do not have a continuity from year to year. The teachers also have many problems trying to figure out what students were previously taught. According to Robert Marzano, (1998) “Keeping no secrets” and “Teaching to the tests” for performance tasks and traditional tests are absolutely necessary to adequately prepare students to meet performance standards.

There has been a lot of research on alternative assessment in the past ten years. Explanations and examples of assessment, along with actual classroom research has been in the forefront of educational research. Although most articles seems to agree that some type of alternative assessment needs to be put in place in the schools, many don't agree on specifically which one. And assuming we eliminate the current traditional methods in the classroom, educators need to agree on which direction the American classroom should go.

One of the reasons the United States is having difficulty deciding between authentic and traditional is the lack of training by the educators and administrators. For example, only 3 out of 5 states require principals to become competent in assessment in order to receive their license. (Stiggins, 1998) So how can a teacher feel comfortable if her evaluator does not?

Richard J. Stiggins believes that there are 3 obstacles to cross before authentic assessment can be a success. The first barrier is the belief that standardized testing scores are necessary to see students' academic growth. And if we are to utilize these types of tests, a more effective one needs to be created. The second barrier is the fear teachers have for being held accountable for student achievement. Teachers need to teach to their students, not to the tests. A third barrier is the way the community and parents feel and react to the standardized tests. Due to their own personal experiences with the standardized tests, they feel that these scores are essential to learning. And until they feel otherwise, scores will be extremely important to them. And the fourth barrier, according to Stiggins, is the lack of clarity about what high school graduates are expected have learned from their school experience. If these barriers are not overcome, effective assessment will never be reached.

In a study conducted by Kamii and Lewis in 1991, data was collected on 87 pupils in 4 Alabama second grade classes in late spring. Two classes had students which were taught in the traditional mathematics style, while the other two were taught in the constructivist primary program. The traditional classes utilized textbooks, workbooks, and manipulatives. The constructivist program asked students to "invent their own procedures for

solving computational and story problems.” They also played games where the students had to use numerical thinking and strategies. The researchers compared the four classrooms using standardized test results and found that the students in the constructivist classroom scored in the 79th percentile in the mathematics section, while the traditional students scored 85th or above. In conclusion, traditional students did better on the State test. But, the researchers took the data a step further. They individually interviewed all of the students and asked them to explain why the answers they had marked were the correct answer. The researchers found that the achievement tests evaluated only if they can solve the problems, not if they understood why it was the correct answer. The final conclusion was that although the traditional students knew more of the answers, they did not understand the idea behind the answer. Whereas the constructivist group seemed to have acquired more background knowledge which might be useful in future learning.

There was a research study done by Farenga in 1998 at Dowling College, New York which was designed to measure students’ out of school science experiences. 349 elementary students, 173 males and 176 females, from grades 3 to 6 were chosen. The survey used was the Elementary

Science Experience Scale or ESES. This survey asked the students 35 questions and looked at three types of experiences.

- (1) Science Related Inquiry Skills - using a globe, map, telescope, ruler
- (2) Life Science - visiting a museum or park
- (3) Physical Science - working on a car, fix an electrical item

The results provided evidence that supported the idea that out of school experiences were relevant to a classroom. These experiences gives students advantages over others when they are studying science. The survey allowed educators to become aware of students' prior knowledge to maximize and enhance their learning.

Although there are many advocates of alternative assessment, there are those who have their doubts about this new idea. J. Terwilliger (1997) believes that alternative assessment “denigrates knowledge and basic skills and that there is little data to support the validity of such assessments.” He also feels that “high quality intellectual performance in the real world requires little knowledge.”

Advocates of alternative assessment can argue with Terwilliger using a national study of 24 elementary, middle, and high school classrooms in mathematics and social studies. Researcher Newmann (1996) have found

that teachers who assigned their students higher level performance tasks regularly were at much higher levels of achievement than those who were assigned less authentic tasks.

While it is important to assign authentic tasks to improve students' skills, it is also imperative to ask the students which task they felt was beneficial to them and how changes can be made for improvement. A research study was done at a university in southeast Georgia from the Fall 1992 through Spring 1996. This study consisted of end of term surveys from 19 graduate level classes. Throughout the course, students were given various and frequent types of evaluations. After the tests and throughout the semester, individual instructor consultations were held. The instructor then utilized the suggestions from the students to improve his teaching. Results found that for the students who had taken advantage of the consultations and due to the various types of testing, grades were higher and felt that the classroom was a better team environment which was beneficial to all. (Griffin, 1998)

With all of these new ideas on how to teach authentically, teachers should know their assessment choices according to Dr. Marlow Ediger (1998)

Types of Tests: Norm Referenced and Criterion Referenced

skills

*multiple choice = guessing and lack of higher order

*short answer = highly factual and not real higher order

*matching tests = measures facts and not thinking

*essay = best higher order skills, creative and critical thinking, BUT lack reliability when evaluator grades

There needs to be a compromise when it comes to assessing students. Dr.

Ediger seems to feel that the compromise is a portfolio in which all types of assessment examples are used.

Portfolios can include any of the following:

*criterion referenced tests

*standardized tests

*personality tests

*teacher observations

*teacher written test items, such as true/false,

completion, multiple choice, matching, and essay

*anecdotal records

*sociograms

*student work

A portfolio can be defined as a tool which will possibly replace the standardized tests one day. It is a showcase for students' work. It can include many types of assignments, projects, reports, and writings. The National Council of Teachers of Mathematics, 1991, feels that a mathematical portfolio should focus on these items:

- *Student Thinking
- *Growth over time
- *Mathematical Connections
- *Students views of themselves as mathematicians
- *The problem solving process

A study was completed in Delaware over a three year period using the portfolio practices of four elementary school classrooms (Roe,1997). The portfolios to be put in place were at a summer literacy program for at risk literacy students and were mandated by the school district. While the researchers looked at the portfolio practices, they also studied the attitude change in the teachers. The researchers felt that the personalities of the teachers will alter the way they keep their portfolios. And, the results found that the portfolios changed as well as the teachers feelings towards their

work throughout the three years. The changes the researchers found were as follows:

1. Attributes of the Portfolio - The folders became more “streamlined” and “fine tuned”
2. Student Involvement - Teachers wanted the children to place the work they felt good about in their folders. Previously, the work was chosen by the teacher and used as proof of the student’s abilities.
3. The Contribution of a Portfolio to Instructional Decision Making - Create curriculum that the children enjoy while learning, and then place the work in the folders.

The researchers found that over time, the teachers moved from “having” portfolios to “doing” portfolios. The teachers utilized the information positively in their working with the students. They felt that they understood their students literacy problems better when they could refer to their portfolios that they felt were beneficial.

Although there are those who feel that portfolios are a great compromise in the assessment battles, there are those who disagree. One concern from the Delaware research (Roe, 1997) was how to incorporate portfolios with instruction in the everyday classroom. Does the portfolio

contribute to more individual consideration of a child's literacy needs ? And do child's literacy achievements exceed expectations when a portfolio is used ? The study realized the importance of a portfolio in the classroom but understood that there are still many quirks to be altered in the program before portfolios will be completely beneficial to both student and teacher.

Vermont was the first state to use portfolios in a statewide assessment program in 1991. Fourth through eight grade students were required to keep portfolios in mathematics and composition. The state would monitor the progress of the students and see how it changed instructional practice.

Requirements were as follows:

1. Teachers would establish a rubric
2. Students would chose their best piece of work to present to the evaluators. That piece, along with the rest of the portfolio would be evaluated on categories of purpose, organization, detail, tone, and grammar.
3. Mathematics work: problem solving and calculation work

Along with the portfolio, students would be required to complete a "uniform test" which was a timed writing sample and included a variety of mathematical problems. The portfolio and uniform test would make up Vermont's assessment policy.

The 1991-1992 results of Vermont's new standards reported a low reliability rating and limited amount of useful data.(Nidds, 1997) Due to these problems, Vermont officials could only use this assessment data at a statewide level. And because of the large student numbers, the information for individuals was less important making the data useful only for entire groups of students. By 1992-1993, improvements were made. The state tried to improve the quality of the portfolios by training their teachers more intensely, making more reliability checks, and controlling and supporting the program more. Still with all of the improvements, Vermont found that the portfolio system is not reliable enough for local and state level decisions. For example, how can educational budgets be decided without specific data ? Why would a town council vote to spend money on an idea they are not comfortable with ? Comfort equals data. Portfolios seem to work well when their are no statewide standards or accountability.

Great Britain has been a leader in the program of authentic assessment involving portfolios.(Nidds, 1997) And even their results have not been as positive as advocates would like. The teachers are reported to have a negative feeling towards this type of assessment. Most teachers "reported that major disruptions had occurred to normal classroom practice, and half of those surveyed felt that the (authentic assessments) were totally

unmanageable.” An example given was that it took 82 to 92 hours “to plan for the assessments, collect needed material for administration, do the assessments, mark them, and record the marks.” Time equals money and such a huge investment is not always possible. Great Britain also felt that the portfolios were unreliable when comparing students throughout the country. Again, researchers feel that portfolio information is best when it is used for individual purposes, not entire towns.

Another types of alternative assessment is Value - Added Assessment (Sanders, 1998) This type of assessment measures the school system, teachers, and administrators by analyzing how well the students academics are progressing throughout the school year. And although the standardized test scores are used, Tennessee is analyzing them quite differently than before. The Tennessee Value - Added Assessment System , or TVAAS, measures each student every year with scales that match the school’s curriculum, follow the child’s yearly progress, and decide whether the objective was met. There are many problems with this type of overall assessing such as students moving, missing tests, and teachers have no control over the types of students they have in their classes.

In 1991, the TVAAS required every student to complete a statewide test in several academic subjects. Grades 2- 8 were tested in math, reading,

language arts, science, and social studies using the CTBS/4 Test. In addition to these tests, the state required an end of the year test for the high school students. By the beginning of 1993, reports have been given to the schools and public on the effectiveness of the school system. The teachers and administrators use this information to see where their school is lacking or improving. The research has found that since 1991, the statewide average in 8th grade math, language arts, and science have risen slowly. Social Studies has remained the same, while the reading comprehension have slowly declined. Between 1992 through 1996, Tennessee was one of the few states where the math scores improved immensely.

The TVAAS research has found many interesting findings since this program was put into place. Some findings are:

- *Gains in the scores are unrelated to race and socioeconomic factors.

- *When students move to other schools, an academic loss is expected in the receiving school. But research has found that this makes little difference in the overall scores.

- *Gifted and Talented students gain less academic growth than the lower average student. Possibly due to teachers paying more attention to the failing student.

*Teachers make the most difference. A good teacher can enhance a child while a poor teacher can cripple their educational growth for the year.

The Tennessee Assessment team has found that there are huge variations between the schools in their state. Changes must occur to improve the educational standards in all schools. Examples to improve variations may include testing students from 2nd grade through 11th, report findings publicly in order to involve the town, and provide assistance to the teachers and administrators so they can do a better job to help their students.

There are many critics to the Value - Added Assessment program. An advocate of the Value- Added Assessment program, Samuel Bratton (1998), who is an administrator in the Knox County School system feels this type of assessing can only help the students. But, critics have stated that standardized, multiple tests are out dated. Mr. Bratton feels that a school needs to do what is right for them, not just jump on the bandwagon. Is it fair to evaluate teachers on the test scores of their students ? Tennessee feels that this information will only help to improve their teachers. Many opposers also feel that this type of testing is expensive. But Mr. Bratton concedes that this is less expensive than the new types of alternative assessing.

In determining whether alternative assessment is beneficial to the students in the classroom, it would be wise to look at how teachers are assessed by their evaluators. In the past and even today, observations done by administrators were standard paper and pencil evaluations. Many are still given graded “report cards” on their performance. If teachers are turning towards authentic assessment for their students, shouldn’t evaluators change their methods for teachers?

Teach for America is a national teacher’s group comprised of recent education graduates. T.F.A. developed a Performance Assessment Institute (PAI) in 1994 with 200 members. These teachers commit for 2 years to teaching in rural or urban public schools. The PAI uses an assessment tool called a portfolio to evaluate whether or not these teachers are meeting the requirements to obtain their permanent teacher certification. The PAI addresses two questions: (1) To what extent has the teacher attained the goals he/she established in student learning and growth. (2) By what means has the teacher impacted student learning and growth. These evaluations allow the teachers more freedom and as long as the students themselves have grown, no standardized tests scores will matter to the children, teachers, or parents. (Jonson, 1998)

G.V. Glass (1974) proposed an “observational judgmental” teacher evaluation that contained multiple resources.

- *a systematic observation from trained evaluators

- *pupil evaluations of teachers

- *data that included professional training and examinations of teacher’s knowledge

T.L. McGreal (1983) also agreed that multiple sources were needed to evaluate a teacher. He stated that “teaching and learning are complex acts that occur in many forms and contexts.” He suggested 6 sources to evaluate.

- *parent evaluation

- *peer evaluation

- *student performance

- *self-evaluation

- *student evaluation

- *artifact collection

Kenneth Peterson, Dannelle Stevens, and Richard C. Ponzio (1998) wrote an article to discuss the need for various teacher type evaluations in order to be fair to the individual. They researched and concluded the following:

*More time and effort needs to be in place for the evaluation process.

*A total representation of the teacher is not possible unless evaluations begin to include more areas.

*Teachers are dissatisfied with evaluations and think they are not representative of the teaching they do.

Teachers today are used to receiving their report card evaluations. If the teachers' evaluations are to change to include a more objective attitude, teachers will be upset with the new suggestions and not the perfect grade. In order to alleviate this disappointment, new rewards for teachers need to be put in place, such as promotions based on merit reviews.

Based on the above research, it is clear that although alternative assessment has its critics, it can be a crucial tool in the modern classroom.

REFERENCES

Bol, Linda. "Influence of Experience, Grade Level, and Subject Area on Teachers' Assessment Practices" *The Journal of Educational Research* (July/August, 1998) : Vol. 91, 323-330.

Bratton, Samuel. "How We're Using Value - Added Assessment" *The School Administrator* (December 1998) : 30- 32.

Dougherty, Eleanor. "What Principals Need to Know about Standards" *Principal* (November 1998) : Vol 78, #2, 12-16.

Ediger, Marlow. "Measuring How Much Students Have Learned." *Education* (Spring, 1998) : 406-415.

Farenga, Stephen & Joyce, Beverly. "Development and Analysis of a Scale to Assess Students' Out-of-School Science Related Experiences" *Education* (Spring, 1998) : 381-383.

Fleming, Kandace. "Teachers' Choice of Test-Item Formats for Classes with Diverse Achievement Levels" *The Journal of Educational Research* (March 1998) : Vol 91 #4, 222-228.

Garnett, Cynthia. "Testing - Do Not Disturb? A Concerned Parent's View of Testing." *Mathematic Teacher* (February, 1992): 35-37.

Glass, G.V. "A Review of Three Methods of Determining Teacher Effectiveness" *Evaluating Educational Performance* (1994) : 11-32.

Gordon, Rick. "A Curriculum for Authentic Learning" *Education Digest* (March 1998) 4-8.

Griffin, Bryan W. "Monitoring and Improving Instructional Practices (And Are Student Evaluations Valid?) *Journal of Research and Development in Education* (Fall, 1998) : Vol. 32, #1, 2-6.

Herman, Joan. "The State of Performance Assessments" *The School Administrator* (December 1998) 17-20.

Jonson, Kathleen. "Promoting Teacher Excellence: A Comparison of Two Performance Based Teacher Assessment Frameworks" *Education* (Summer, 1998) : 499-508.

Lester, Frank K., & Kroll, Diana. "Evaluation : A New Vision." *Mathematics Teacher* (April, 1991): 3-7.

Marzano, Robert. "Rethinking Tests and Performance Tasks" *The School Administrator* (December 1989) 10-12.

McGreal, T.L. "Successful Teacher Evaluation" Alexandria, Va. Association for Supervision & Curriculum Development. (1983).

Newmann, Fred. "An Exchange of Views on "Semantics, Pyschometrics, and Assessment Reform: A Close Look at Authentic Assessments" *Educational Researcher* (Aug.- Sept. 1998) : 19-21.

Nidds, John & McGerald, James. "How Functional is Porfolio Assessment Anyway ?" *Education Digest* (January, 1997) : 47-50.

Peterson, Kenneth, Stevens, Dannelle, & Ponzio, Richard, "Variable Data Sources in Teacher Evaluation" *Journal of Research and Development in Education* (Spring 1998) : Vol 31 #3, 126-131.

Popham, W. James. "Farewell Curriculum: Confessions of an Assessment Convert." *Phi Delta Kappan* (Jan. 1998) : 380 - 384.

Roe, Mary. "That was then and This is now: A Longitudinal Study of Teachers' Portfolio Practices." *Journal of Research in Childhood Education* (1997) : Vol. 12 No. 1, 16 - 25.

Sanders, William. "Value - Added Assessment" *The School Administrator* (December 1998) : 24-27.

Steele, Diana & Arth, Alfred. "Lowering Anxiety in the Math Curriculum" *Education Digest* (March, 1998) : 18-23.

Stenmark, Jean. *Mathematics Assessment*, National Council of Teachers of Mathematics (1996)

Stiggins, Richard J. "Confronting the Barriers to Effective Assessment."
The School Administrator (December, 1998) : 6-8.

Terwilliger, J. "Semantics, Psychometrics, and Authentic Reform: A Close
Look at Authentic Assessments." *Educational Researcher* (1997) 26(8), 24-
27.

APPENDICES

APPENDIX A

ALTERNATIVE ASSESSMENT QUESTIONNAIRE

Please answer the following questions:

1. Years of teaching experience _____
2. Years of teaching at George Washington _____
3. Grade teaching _____
4. Student class size _____
5. When evaluating your students, which assessing methods do you utilize ?

Always Sometimes Seldom Never

Multiple Choice

True/False

Fill in the Blank

Short Answer

Essay

Matching

Project

Text Book Made Tests

Your Own Made Tests

Rubric

Traditional grades (100 points)

From the above selections, which assessment method do you feel is the easiest to correct ? _____

most beneficial to learning ? _____

most enjoyable for the student ? _____

APPENDIX B

STUDENT QUESTIONNAIRE ON ALTERNATIVE ASSESSMENT

1. Age _____
2. Grade you came to Hillside _____
3. From the following list, which type of test do you see the most of ?

Always Sometimes Seldom Never

- Multiple Choice
- True/False
- Fill in the Blank
- Short Answer
- Essay
- Matching
- Projects
- Text Book Made Tests
- Teacher Made Tests
- Rubrics
- Traditional Grades (100 points)

Rank in order your 3 favorite types of tests

- 1.
- 2.
- 3.

Why?

Rank in order your least favorite types of test

- 1.
- 2.
- 3.

Why?

Why do you think that teachers give you test ?

If there were no tests, how could a teacher find out if you learned the information ?

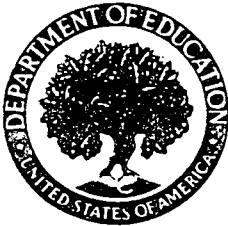
Thank you for your input !!

READING RUBRIC

CLASS NAME OR PERIOD NUMBER		READING RUBRIC															
CLASS SUMMARY		2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1
Average	Highest Score	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
Lowest Score		0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	

Type the names of assignments and tests in the spaces below.

STUDENTS	INC.	GRADE	AVG.													
G		B	2.50%	3.00	1.00	3.00	1.00	2.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
P		B	2.40%	1.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	1.00	1.00
X		B	2.40%	2.00	3.00	2.00	2.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
D		B	2.20%	3.00	3.00	3.00	3.00	2.00	2.00	1.00	1.00	1.00	1.00	2.00	2.00	2.00
T		B	2.20%	3.00	3.00	3.00	3.00	0.00	0.00	0.00	1.00	1.00	1.00	3.00	3.00	3.00
J		B	2.00%	3.00	3.00	3.00	1.00	3.00	3.00	2.00	2.00	1.00	1.00	2.00	1.00	1.00
C		B	1.90%	2.00	1.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	1.00	1.00	1.00	1.00
Z		B	1.90%	2.00	1.00	3.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00	2.00	2.00	3.00
F		B	1.80%	1.00	1.00	3.00	1.00	2.00	2.00	3.00	3.00	3.00	3.00	1.00	1.00	1.00
S		B	1.80%	3.00	1.00	1.00	1.00	3.00	3.00	3.00	3.00	3.00	1.00	1.00	3.00	1.00
AA		B	1.80%	2.00	2.00	2.00	1.00	1.00	1.00	2.00	2.00	3.00	3.00	0.00	0.00	2.00
H		B	1.70%	1.00	1.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	2.00	2.00
U		B	1.60%	2.00	1.00	1.00	3.00	2.00	2.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
J		B	1.60%	3.00	3.00	0.00	1.00	2.00	2.00	1.00	2.00	2.00	1.00	1.00	1.00	1.00
Q		B	1.60%	3.00	1.00	2.00	1.00	2.00	2.00	1.00	2.00	1.00	1.00	1.00	2.00	1.00
L		C	1.50%	3.00	1.00	2.00	1.00	2.00	2.00	1.00	1.00	1.00	1.00	3.00	1.00	1.00
M		C	1.50%	3.00	1.00	1.00	1.00	1.00	1.00	2.00	2.00	1.00	1.00	1.00	1.00	1.00
Y		C	1.50%	2.00	3.00	3.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00	3.00	1.00	1.00
K		C	1.40%	3.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	2.00
R		C	1.40%	1.00	1.00	1.00	2.00	2.00	2.00	1.00	1.00	1.00	1.00	2.00	2.00	2.00
E		C	1.30%	3.00	1.00	1.00	3.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00
N		C	1.30%	1.00	1.00	2.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	2.00	2.00	2.00
U		C	1.20%	2.00	3.00	2.00	2.00	2.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
B		C	1.10%	1.00	3.00	2.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00
O		C	1.00%	0.00	1.00	1.00	0.00	2.00	2.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
W		C	1.00%	2.00	2.00	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
A		C	0.90%	2.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00
V		C	0.90%	2.00	2.00	0.00	2.00	0.00	0.00	2.00	2.00	0.00	0.00	1.00	1.00	1.00



U.S. DEPARTMENT OF EDUCATION
OFFICE OF EDUCATIONAL RESEARCH AND IMPROVEMENT (OERI)
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

ERIC/RCS
2805 East Tenth St.
Bloomington, IN 47405

TM029844

REPRODUCTION RELEASE (Specific Document)

I. DOCUMENT IDENTIFICATION

Title: TRADITIONAL ASSESSMENT VS ALTERNATIVE ASSESSMENT
Author(s): DANA BELLE
Corporate Source (if appropriate): KEAN UNIVERSITY
Publication Date: MAY 1999

II. REPRODUCTION RELEASE

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche and paper copy (or microfiche only) and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the options and sign the release below.

CHECK
HERE



Microfiche
(4" x 6" film)
and paper copy
(8 1/2" x 11")
reproduction

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY
DANA BELLE
(PERSONAL NAME OR ORGANIZATION
KEAN UNIVERSITY
AS APPROPRIATE)

OR



Microfiche
(4" x 6" film)
reproduction
only

"PERMISSION TO REPRODUCE THIS
MATERIAL IN MICROFICHE ONLY
HAS BEEN GRANTED BY

(PERSONAL NAME OR ORGANIZATION
AS APPROPRIATE)

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed in both microfiche and paper copy.

SIGN
HERE

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction of microfiche by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: Dana Belle Printed Name: DANA BELLE
Organization: KEAN UNIVERSITY
Address: 153 ILLER DRIVE Position: STUDENT
MIDDLETOWN NJ 07748 Tel. No.: 732 706 7307
Zip Code: 07748 Date: MAY 1999

III. DOCUMENT AVAILABILITY INFORMATION (Non-ERIC Source)

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents which cannot be made available through EDRS.)

Publisher/Distributor: _____
Address: _____
Price Per Copy: _____ Quantity Price: _____

IV. REFERRAL TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER

If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:



DOCUMENT TRANSFER FORM

Date Received: 05-12-99

Date Transferred: 05-12-99

TO:

ERIC/ *TM*

FROM:

ERIC/ CS- ERIC Clearinghouse for Reading, English, and Communication

The enclosed document, *Traditional Assessment Versus Alternative Assessment*, appears to be more within your scope than ours.

The document was obtained by:

- solicited contribution
- unsolicited contribution
- staff contribution
- ERIC Facility

Correspondence information is as follows:

- all relevant material is attached
- no relevant correspondence; however the name and address of the contributor or contact person is (if known):

author notified of transfer

Roger Sensenbaugh
Clearinghouse Contact Person