Spring 5-15-2009

Authentic Assessment versus Traditional Assessment: A Comparative Study

Nikki Brawley
Coastal Carolina University

Authentic Assessment vs. Traditional Assessment: A Comparative Study

By

Nikki Brawley

Early Childhood Education

May 2009

Submitted in Partial Fulfillment of the
Requirements for the Degree of Bachelor of Science
In the Honors Program at
Coastal Carolina University
May 2009

Assessment is defined as "almost any form of measurement and appraisal of what students know and can do." These forms can include, but are not limited to: tests, reports, observations, and questioning. There are two main subgroups used when describing assessment: traditional and alternative. The term traditional assessment refers to a paper and pencil based test used to determine what a student knows and can recall, i.e. the Palmetto Achievement Challenge Tests (PACT). The term alternative assessment refers to "almost any type of assessment other than standardized tests." (McAfee & Leong, 2-3) This subgroup can be divided into various types of alternative assessments. Alternative assessment includes, but is not limited to, observations, checklists, student portfolios, and self-evaluations. The type focused on throughout this paper is authentic assessment, with the purpose of determining whether authentic assessment requires higher-order thinking skills than traditional assessment. The main focus of the paper is to determine which test shows more about what the student knows about the intended topic.

Authentic assessments are those tests which are used to see if students can apply the knowledge they have learned in a real-world setting. (Frey & Schmitt, 406) Another definition for authentic assessment is any assessment that is "part of children's ongoing life and learning in the classroom…and other typical school and center settings." One way to understand the distinction between traditional and authentic assessments is by comparing the test given to receive a driver's license. The first part of the test is a multiple choice test; this is the traditional assessment. The second part of the test is the actual act of driving a vehicle; this is the authentic assessment. (McAfee & Leong, 2-3) Authentic assessments "examine student performance on worthy intellectual tasks," whereas traditional assessments simply require the ability to recall knowledge. (Wiggins, 1-6)

Theories

Many educators draw on theorists to validate the actions they take in the classroom. They also support their use of authentic assessment by basing their own assessments on educational theorists. These theorists include Burrhus Frederic Skinner, Albert Bandura, Jean Piaget, and Lev Semenovich Vygotsky.

Skinner was a behaviorist. Behaviorists define learning as a change in the behavior of a student. He believed that learning could only occur if the student could demonstrate the behavior. Bandura believed that students could learn through observation and did not have to be demonstrated, unless conditions were set up to elicit the behavior. Piaget believed that during a certain period of knowledge construction, student behavior could change but only after students had developed schema. Schema is defined as a cognitive way to organize knowledge. It can be pictured as an elaborate network of abstract mental structures which represent one's understanding of the world. (Barry, Naus, & Rehm, 722) Vygotsky believed in a "Zone of Proximal Development." He said that in this Zone, a behavior may be achieved with help from an adult before the child could demonstrate the behavior independently. (Bergen, 99)

Well-designed, authentic assessments draw upon the theory of B.F. Skinner by "measuring behavior change immediately after a learning event." They support the theory of Bandura by eliciting a learned behavior. (Bergen, 99) By using "the developmental period in which mistakes occur to understand how knowledge is developing," assessments address the learning theory of Piaget. Vygotsky's theory of learning is concerned with how a student performs "both with and without adult facilitation." (Bergen, 99) Using standards and having different ways to demonstrate reaching those standards, learning from mistakes, and

acknowledging adult facilitation are all steps to help authentic assessments satisfy the various theories about learning.

Rubric Usage

Research has shown that using rubrics as a way to assess students is an acceptable practice. Project LINK is a program being tested in the Head Start program to improve child outcomes. It involves four different components. The first of these components is the use of authentic assessment. Another component is that each child has an individualized plan for learning. The third component is the use of specific curriculum and instructional plans. The final component is the development of an individual student portfolio. Rubrics are used to assess the students. Observers rated the students on a scale from zero to two. A score of zero means that a child did not demonstrate the desired skill. A score of one means that the child could partially demonstrate the skill or demonstrate it with assistance. A score of two means that the child could demonstrate the skill according to the criteria set by the teacher. (Brown, Hallam, & Frontczak, 275-276)

Why Assess?

There are two general purposes for assessments. Formative assessments are used to gather information. This information is "then used to shape and improve" the programs used in the classroom. Summative assessments are done at the end of a predetermined length of time. They are used to "determine the effectiveness of a program." (McAfee, 3)

Authentic assessment can be used to find out what children know and can do. Also, they can provide insights on students' "attitudes, interests, and approaches to learning. The information provided is used to guide and assist children's growth, development, and learning." (McAfee, 9)

To summarize, teachers assess students to monitor development and learning. They also use assessments to guide classroom planning and decision making. Assessments can also be

used to identify children who might benefit from special help, in addition to reporting and communicating with others. (McAfee, 34).

Goals of Assessment

The main purpose of assessments should be to provide teachers with the necessary information to better their instruction in an attempt to address the needs of all of the students in their classroom. Assessments should reflect the abilities of the child and not just the demographics of the school. Assessments are intended to improve instruction. Authentic assessment is based on the idea that assessments should be performed on what the students are learning and not just what they are expected to learn. The assessments should be customized to the specific classroom and group of children. (Christie, 565-567)

It is believed that assessment results should be used not to compare individuals or specific populations (McAfee, 181). Instead, they should serve three specific functions. Assessments should be used to make "sound decisions about teaching and learning," to identify "significant concerns for individual children," and to "help programs improve their educational and developmental interventions." (McAfee, 28)

Developmental Appropriateness

Developmentally appropriate practices result from the process of professionals making decisions about the well-being and education of children based on at least three important kinds of information or knowledge. One kind of information is "what is known about child development and learning—knowledge of age-related human characteristics that permits general predictions within an age range about what activities, materials, interactions, or experiences will be safe, healthy, interesting, achievable, and also challenging to children." (NAEYC position statement)

Another kind of information is "what is known about the strengths, interests, and needs of each individual child in the group to be able to adapt for and be responsive to inevitable individual variation." The third kind of information used to make decisions is "knowledge of the social and cultural contexts in which children live to ensure that learning experiences are meaningful, relevant, and respectful." (NAEYC position statement)

Developmental appropriateness applies to everyday classroom assessment and the curriculum. This is indicative of the fact that when designing assessment there are three key factors which should be taken into account. One of these factors is the age of the students; the age of the student determines how they learn and assimilate knowledge. Another factor is the developmental level of the student. There are developmental milestones which must be attained before students are mentally, physically, and emotionally prepared for learning certain concepts. The final factor is the student's life experiences. Some students have not had the same experiences as their classmates. Their understanding of concepts is based on their past experiences. (Cress, 96)

Bloom's Taxonomy

Bloom's Taxonomy was first proposed by Benjamin Bloom in 1956. His taxonomy was divided based on three domains: affective, psychomotor, and cognitive. The cognitive domain was hierarchically divided. The levels were knowledge, comprehension, application, analysis, synthesis, and evaluation. (McAfee, 59) Bloom's Taxonomy was revised in 2001. The levels in the cognitive domain were changed to remembering, understanding, applying, analyzing, evaluating, and creating. The difficulty or complexity of thinking required progresses from remembering to creating. (Krathwohl, 215) Since learning occurs in the cognitive domain, that will be the focus of the research.

Bloom's Taxonomy is used to provide teachers with a way to define higher order thinking. By using the Taxonomy's levels, the teacher is able to determine the level of thinking required to complete the assessment. This provides the teacher with a clearer picture of the child's cognitive development in the classroom. Applying the Revised Taxonomy to assessment causes the view of assessment to "extend beyond discrete bits of knowledge and individual cognitive processes to focus on more complex aspects of learning and thinking." (Airasian & Miranda, 249)

Bloom's Taxonomy is used to identify learning objectives and the various levels of learning. A study conducted at Duke University used Bloom's Taxonomy to design discipline-specific assessments of critical thinking. The questions used were designed to elicit higher-order thinking skills. (Bissell & Lemons, 67)

Advocates

Authentic assessment, specifically performance assessment provides students with the opportunity to demonstrate what the knowledge they have acquired about a subject. This may be a "more valid indicator of students' knowledge and abilities" than traditional tests. (Sweet, 1-5) Students apply the knowledge they have, instead of just recalling it. When authentic assessments are based on what students are learning at that present time, the curriculum is allowed to drive the test. (Sweet & Zimmerman, 1-5)

Advocates say that with authentic assessments, there is actual "observable products and understandable evidence" of student performance. The teacher can discuss the strengths and weaknesses of each student using understandable terms and real-life situations. Authentic assessments can also provide understandable evidence for the parents of students. (Wiggins, 1-6)

Supporters of more authentic tests have come to believe that if "teachers are going to teach to the test, the tests…should be worth teaching to." Why not make tests that focus not only on the basic skills students learn but also on the development of higher-order thinking skills? They believe assessment should be part of instruction, not something that happens once all of the learning is finished. (Guskey, 51-54)

Critics

Critics believe that state standardized tests do not improve instruction because the results of state tests are not available until after the student or group of students have risen to the next grade. Standardized tests do not provide for the differences inherent in students from various areas, i.e. socioeconomic levels. (Christie, 564-567) They also think that standardized tests are given power by "political and educational policy makers and by societies."

Many critics feel that standardized tests are misused and overused. It is also felt that the tests have "undue influence on what is taught." (McAfee, 8) In fact, they believe that pressures facing teachers lead to more concentration on the short term and focus on lower-level thinking skills as identified by Bloom's Taxonomy. (Holmes, 87)

Critics also believe that standardized tests force the teacher to teach to the tests. Many teachers feel that standardized tests encourage the skewing of instruction to focus on the more basic skills which the tests assess. Due to this skewing, the curriculum suffers and is thus narrowed to fit what the test is assessing.

Tests

Some traditional assessments, like standardized tests, do not provide scores to the classroom teacher in time to adjust teaching methods. Authentic assessments test the knowledge of the student. More importantly, they test the ability of students to apply that knowledge.

The use of authentic assessments in the classroom has support from the major theorists in the education field. Bloom's Taxonomy has also been used in research conducted on assessment. The complexity of thinking required for the completion of an authentic assessment is slowly becoming a major issue in the education field. Thus the question seeking an answer through the following testing is which test is designed for the use of higher-order thinking skills.

Testing

The research, or data collection, portion of the project occurred at a Horry County, SC elementary school in a third grade classroom. The use of a third grade classroom as an early childhood classroom is supported by the Spadoni College of Education at Coastal Carolina University, which defines Early Childhood as grades pre-K to third (2008). The class was composed of twenty-four students. The students ranged from eight to nine years old. The class was composed of the high performing math students from two of the third grade classes.

The standard being assessed was the South Carolina Mathematics standard that says students "will demonstrate through the mathematical processes an understanding of methods of multiplying whole numbers." (SC Standards) One of the indicators of that standard is that students will "recall basic multiplication facts through 12 x 12." (SC Standards) To assess the standard, two tests were given.

The classroom teacher gave students a multiplication test that consisted of thirty problems. The nines multiplication table was the focus of the assessment. The teacher typically used this traditional assessment to judge the students' knowledge of multiplication facts.

The students were then individually given an authentic assessment designed to check their understanding of multiplication facts. The authentic assessment was based on a rubric created from the questions on the traditional test and the performance required to demonstrate students' knowledge. The rubric had three sections. The student could get up to 2 points per section. The first section was to identify the teacher created multiplication problem. The second section was to use the blocks to create a visual representation of the given multiplication problem. The third section was the explanation of the use of the blocks. The total possible score a student could receive was six points.

Rubric for Authentic Assessment

0	1	2
Could not identify the problem represented by the blocks, created by the teacher.	Identified what problem was created with some guidance.	Identified the problem the teacher created.
Could not use the blocks to create a visual representation of the given multiplication problem.	With some guidance from the teacher, the student was able to create a visual representation of the multiplication problem given.	Created a visual representation without any assistance from the teacher.
Could not explain why they used the blocks in that way.	With some guidance, student could explain their use of the blocks.	Easily explained the use of the blocks with no guidance from the teacher.

To ensure the anonymity of the students, both tests and the results were coded by the classroom teacher. Each student was given a letter to identify them by the teacher. She labeled the two tests with each student's letter to preserve their anonymity. The sheets used to code the assessments were then shredded so no child's name would be associated with the results. The test results were compared to see how well the students understood the mathematics concept and to see what levels of Bloom's were used to complete the test.

On the traditional test, students were asked to complete thirty multiplication problems. To complete this test, students were required to use the lowest level of Bloom's Taxonomy. The assessment was designed in a manner that required students to recall or remember information.

On the authentic test, students were asked to identify the multiplication problem represented by the blocks made by the teacher. They were then asked to create a visual representation of a given multiplication problem using the blocks. Finally, they were asked to explain why they used the blocks in that way and what they represented.
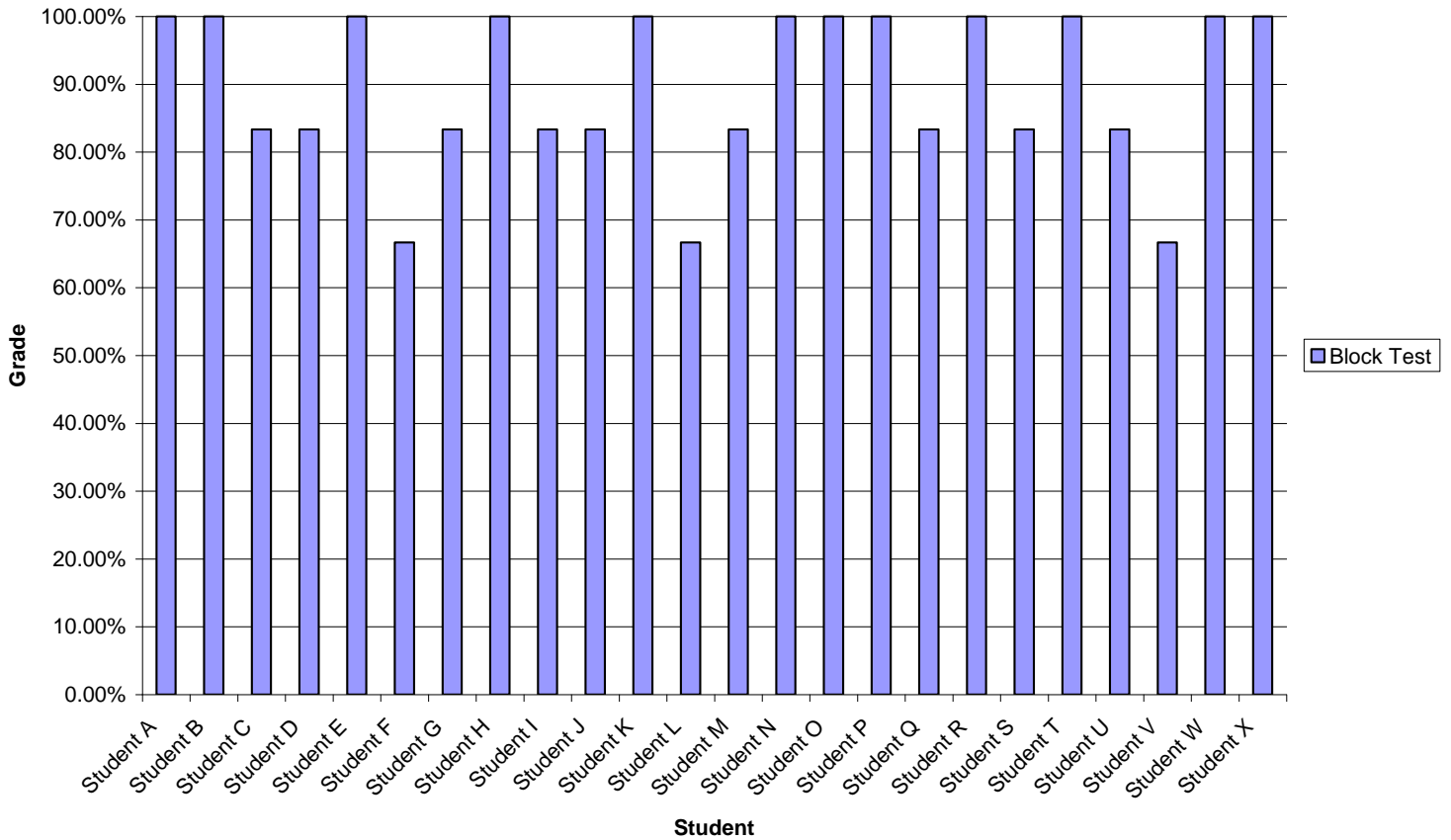
Results

Eight students made 100 on the traditional assessment. Two students scored in the 80s. Seven students scored in the 70s. Two students scored in the 50s and two in the 30s. One student scored in the 60s, one in the 40s, and one in the 20s. The passing score was 70. Of the twenty-four students tested, seven students fell below the passing line. Twelve students made 100 on the authentic assessment. Nine students scored in the 80s. Three students made a score in the 60s. Again, the passing score was 70. Of the twenty-four students who took the authentic assessment, three students fell below the passing line set by the teacher.

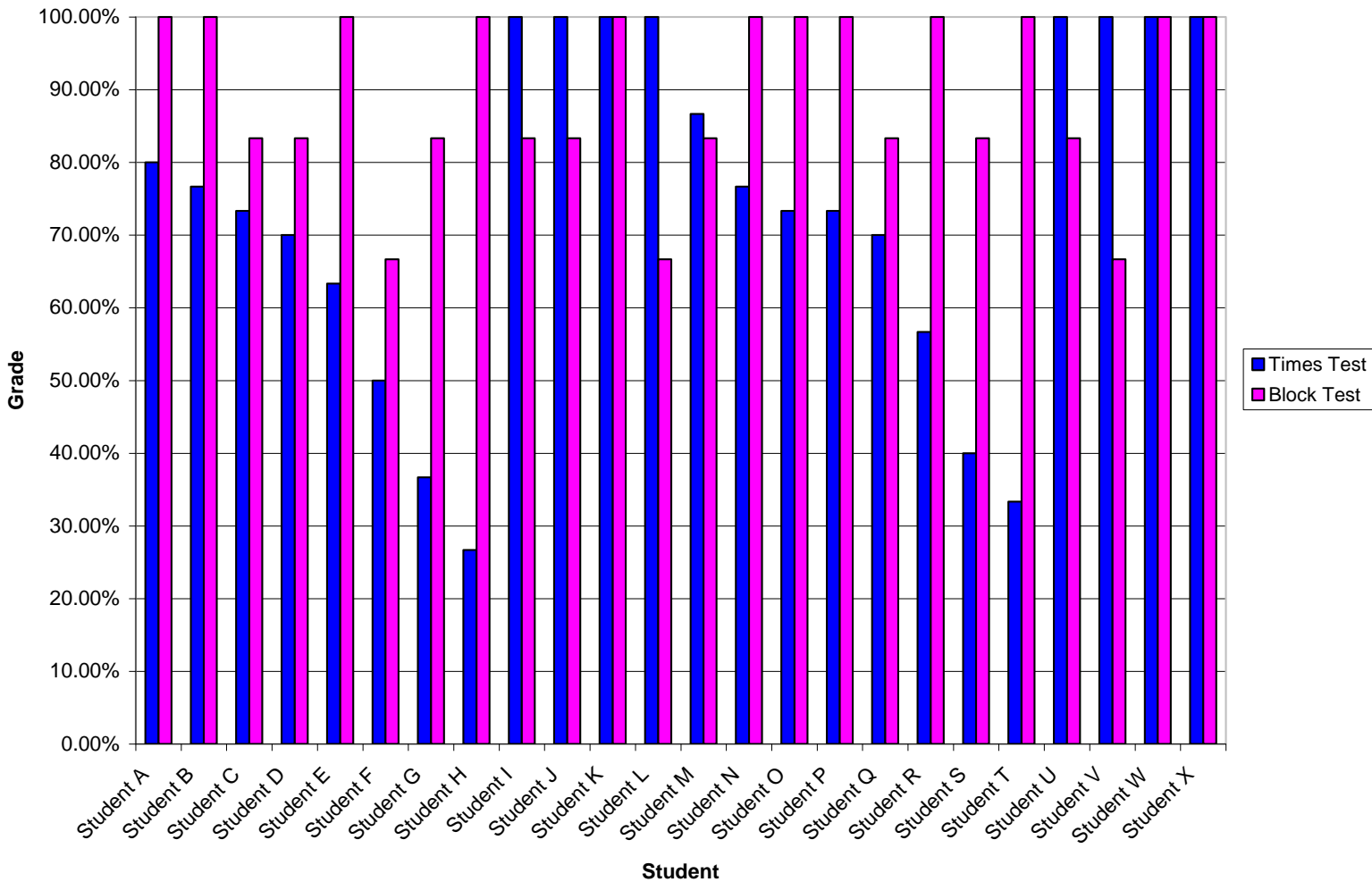**Traditional Assesment**
Multiplication Test
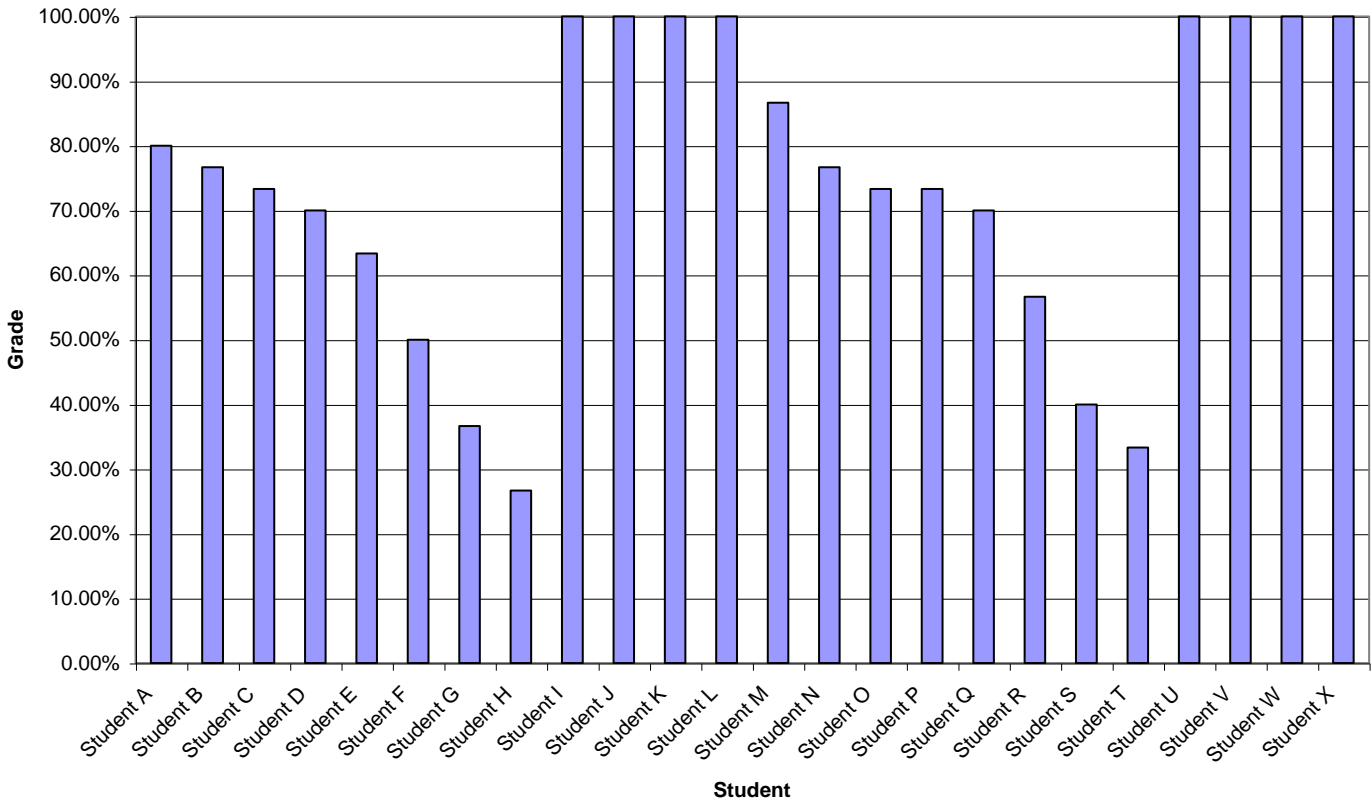
Authentic Assessment
Block Test

Students K, W, and X made 100 on both assessments. These results show that the students not only recall the information but they also understood the concept. Fifteen students improved their scores from the multiplication test to the authentic assessment or block test.

Comparison of Tests

One specific student, student H, scored 26.6 on the traditional assessment. When student

H participated in the authentic assessment, they scored 100. This shows that while the student

may take time to recall information about multiplication facts, he/she demonstrated five of the

six levels of Bloom's Taxonomy. In other words, student H demonstrated higher order thinking

on the authentic assessment.

Six students had scores drop from the traditional assessment to the authentic assessment.

Student L scored 33.3 points lower on the authentic assessment than on the traditional

assessment. He/she was able to recognize or identify the problem created by the teacher. The student required some guidance with creating his own visual representation of a multiplication problem. He/she also needed some extra guiding questions to help with his explanation of how he used the blocks. The student's performance on the traditional assessment would show that they know their multiplication facts; however, the performance on the authentic assessment appears to show that the student has limited understanding of the concept.

Statistical Significance

When comparing the average scores from the two testing procedures, there was a statistically significant difference of 15.14 points (p=0.015), wherein the authentic assessment result was, on average, higher than the traditional assessment result within this data set. This suggests that a systematic function of the particular assessment instrument was responsible for this difference, rather than a variance found by chance.

Conclusion

For the most part, traditional assessments are designed to provide quick statistical data. The assessments do not typically require higher-order or higher-level thinking as defined by Bloom. Authentic assessments can be and typically are designed to encourage higher-order thinking skills.

Assessing a student's knowledge is a necessary part of the education profession. The question to be addressed is if teachers are taking time to assess, why not design an assessment that will further student thinking and better prepare them for using their knowledge during their lifetime?

The results from this research support the idea that authentic assessments, when designed properly, are a better way to determine the higher-order thinking skills required to complete a

task. Standardized testing has become the focus in many classrooms in recent years. The problem with this overemphasis is that most standardized tests are not designed to elicit higher-order thinking skills.

While both traditional and authentic assessments provide information about student knowledge, the skills required for completion are very distinct. Traditional assessments rely on the lower levels of Bloom's, which include remembering, understanding and applying. Authentic assessments, when designed to do so, rely on the higher levels of Bloom's, including analyzing, evaluating, and creating. The research done comparing authentic and traditional assessments in this third grade classroom adds validity to the idea that, when properly designed, authentic assessments demonstrate not only the student's knowledge of the topic but also their ability to utilize it.

Works Cited

National Association for the Education of Young Children.(2003) Retrieved April 17, 2009,

    from NAEYC Position Statement Web site:

    http://www.naeyc.org/about/positions/pdf/CAPEexpand.pdf

Airasian, P. W. & Miranda, H. (2002). The Role of Assessment in the Revised Taxonomy.

    Theory in Practice, 41 (4), 249-254. Retrieved November 26, 2008, from JSTOR.

Barry, E., Naus, M., & Rehm, L. (2006). Depression, implicit memory, and self: A revised

    memory model of emotion. Clinical Psychology Review, 26(6), 719-745. Retrieved

    January 12, 2009, from Academic Search Premier database.

Bergen, D. (1993). Authentic performance assessments. Childhood Education, 70 (2), 99-

    100. Retrieved August 30, 2008, from Educator's Reference Complete via Gale.

Bissell, A., & Lemons, P. (2006). A New Method for Assessing Critical Thinking in the

    Classroom. Bioscience, 56(1), 66-72. Retrieved January 12, 2009, from Academic Search

    Premier database.

Christie, K. (2007). Making Assessments More Meaningful. Phi Delta Kappan, 88 (8), 565-567.

    Retrieved September 9, 2008, from Academic Search Premier.

Cress, S. (2004). Assessing Standards in the "Real Kindergarten Classroom. Early Childhood

    Education Journal, 32(2), 95-99. Retrieved January 12, 2009, from Academic Search

    Premier database.

Frey, B., & Schmitt, V. (2007). Coming to Terms With Classroom Assessment. Journal of

    Advanced Academics, 18(3), 402-423. Retrieved January 12, 2009, from Academic

    Search Premier database.

Grisham-Brown, J., Hallam, R., & Pretti-Frontczak, K. (2008). Preparing Head Start Personnel

     to Use a Curriculum-Based Assessment: An Innovative Practice in the Age of

     Accountability. Journal of Early Intervention, 30(4), 271-281. Retrieved January 12,

     2009, from Education Research Complete database.

Guskey, T.R. (1994). What you assess may not be what you get. Educational Leadership, 51 (6),

     51-54. Retrieved September 9, 2008, from Academic OneFile via Gale.

Holmes, P. (2002). Assessment: New ways of pupil evaluation using real data. Teaching

     Statistics, 24 (3), 87-89. Retrieved November 4, 2008, from Academic Search Premier.

Krathwohl, D. (2002). A Revision of Bloom's Taxonomy: An Overview. Theory Into Practice,

     41(4), 212. Retrieved January 12, 2009, from Academic Search Premier database.

McAfee, O, & Leong, D (2007). Assessing and guiding young children's development and

     learning. Boston: Pearson Education, Inc.

Sweet, D., & Zimmerman, J. (1992). Performance Assessment (ERIC Document Reproduction

     Service No. ED/OERI-92-38). Education Research Consumer Guide, 2, 1-5.

Wiggins, G. (1990). The Case for Authentic Assessment (ERIC Document Reproduction Service

     No. EDO-TM-90-10). ERIC Digest, 1-6.