

**AN INVESTIGATION INTO THE RELATIONSHIP  
BETWEEN FOURTH GRADE DEGREES OF READING  
POWER ACHIEVEMENT AND THE FOURTH GRADE NEW  
YORK STATE ENGLISH LANGUAGE ARTS  
ASSESSMENT ACHIEVEMENT**

**THESIS**

Submitted to the Graduate Committee of the Department  
of Education and Human Development  
State University of New York  
College at Brockport

In Partial Fulfillment of the Requirements for the  
Degree of Master of Science in Education

By

Julie L. Wells

State University of New York  
College at Brockport  
Brockport, New York

August 1999

SUBMITTED BY:

Julie L. Wells 7/14/99

APPROVED BY:

Pat E. Smith 7/14/99  
Thesis Advisor Date

Geneva L. Bezy 7/13/99  
Second Faculty Member Date

Scott O. Robinson 7/14/99  
Director of Graduate Studies Date

## Abstract

The purpose of this study was to determine if there is a statistically significant relationship in achievement between the traditional standardized fourth grade DRP reading test and the NYS English Language Arts Fourth Grade assessment.

Academic standardized testing is becoming more performance based and is letting go of the traditional multiple choice assessment. New York State has adopted 28 learning standards for seven of the content areas. The language arts content area has four learning standards that have a great emphasis on higher level thinking and processing skills.

To be able to assess the students properly, the state has developed the English Language Arts Assessment that requires students to create a more authentic written response by making more insightful connections between the reading passages and personal experience.

The subjects were one hundred and thirty-one rural, heterogeneously-grouped fourth grade students. The ages of the students range from 8-10 years. The students' reading abilities range from advanced readers to students who receive Title One support.

A correlation was used to determine if there was a statistically significant difference between the NYS ELA and DRP test scores.

All the findings indicate a strong positive correlation between the DRP reading test and the NYS ELA assessment. With 95% confidence,

this study can expect that students who scored the highest on the ELA assessment should also score high on the DRP reading test, but may not have the same exact score. The same should also be true for the low and middle range scores. This expectation can be permitted due to the finding that essentially the skills measured on both tests are related.

## Table of Contents

<b><u>Chapter 1</u></b> Statement of the Problem	
Purpose of the Study	1
Need for the Study	1-3
Definition of Terms	3
<b><u>Chapter 2</u></b> Review of the Literature	
Purpose of the Study	4
Standardized And Performance Assessments	4-6
Standardized Reading Tests	6-9
Standard-based Curriculum	9-10
Performance Testing	10-12
Summary	12
<b><u>Chapter 3</u></b> Design of the Study	
Purpose of the Study	13
Null Hypothesis	13
Methodology	13-14
Analysis of Data	14
<b><u>Chapter 4</u></b> Analysis of Data	
Purpose of the Study	15
Null Hypothesis	15
Analysis of Data	15-16
Summary	17
<b><u>Chapter 5</u></b> Conclusions and Implications	
Purpose of the Study	18
Conclusions and Implications	18-19
Implications for the Classroom	19-21
Implications for Further Research	21-22
Summary	22
<b><u>References</u></b>	23-25

# **CHAPTER I**

## **Statement of the Problem**

### **Purpose of the Study**

The purpose of this study was to determine if there is a statistically significant relationship in achievement between the traditional standardized fourth grade DRP reading test and the NYS English Language Arts Fourth Grade assessment.

### **Need for the Study**

As our country and state adopt a standard based educational system there is need to find more accurate ways to assess the students' learning and program development. Academic standardized testing is becoming more performance based and is letting go of the traditional multiple choice assessment. New York State has adopted 28 learning standards for seven of the content areas. The language arts content area has four learning standards that have a great emphasis on higher level thinking and processing skills.

To be able to assess the students properly, the state has developed the English Language Arts Assessment that requires

students to create a more authentic written response by making more insightful connections between the reading passages and personal experience. The test will have a section in which the students are asked to listen to a short reading, take notes and use graphic organizers to answer short essay questions. The students will also be asked to answer multiple choice comprehension questions looking for skills in: vocabulary development, ability to use inferential thinking skills, and finding factual information. The New York State English Language Arts (ELA) assessment looks and feels quite different from the typical standardized reading test given in the past. Those tests are often made up of multiple choice questions looking for one correct answer to fit in a cloze passage. The differences in structure between the two assessments are quite extensive.

For educators, administrators, and parents, the goal is finding the relationship between these two types of standardized reading tests. States are requiring students to become highly literate thinkers. In order to provide the best instructional atmosphere for the students, it is necessary to learn which standardized test allows the student to demonstrate their abilities in the most productive way. The intent of this study is to find out if each student's results show a significant relationship between the two assessments. Once the relationship, if

any, has been discovered, direction can be placed on improving reading achievement.

### **Definitions of Terms**

**Standardized test-** An instrument composed of empirically selected items that has definite instructions for use, adequately determined norms, and data on reliability and validity.

**Standard-Based Education-** All educational content areas are based on a set of standards which are evaluated by a set of performance tasks or benchmarks.

**Performance Assessment-** An assessment where the learner is required to show their competency by performing a series of related tasks.

## **CHAPTER II**

### **Review of the Literature**

#### **Purpose of the Study**

The purpose of this study was to determine if there is a statistically significant relationship in achievement between the traditional standardized fourth grade DRP reading test and the NYS English Language Arts Fourth Grade assessment.

#### **Standardized and Performance Assessments**

"What should *all* our students know? What should *all* our students be able to do with this knowledge?" (Wheelock, 1997, p.29) Teams made up of communities, school districts and state education departments are researching these important questions regarding students' learning. "Therefore, many have reasserted the position that assessment and curriculum should be untied in the serving of learning." (Bergan & Feld, 1993, p.41)

" The process of an assessment, however, is not just focused on evaluating student accomplishment. Rather, the heart of assessment is a continuing flow in which the teacher (in collaboration with the student) uses information to guide the next steps in learning. The

educator must ask, "What should I do to help the student progress?" (Neill, 1997, p.35) Many states, including New York, are developing standard based educational programs and assessments in an attempt to allow teachers to consistently revisit their students needs to produce highly literate thinkers and learners. Standard based curriculum is a set of goals and benchmarks that the students will achieve. In helping the students develop higher level thinking skills, teachers now have new avenues to exploring and facilitating learning.

Teachers are asking essential questions (ex. When is fair, fair, and when does it become unfair?), and exploring big ideas to motivate and challenge learning. Standard based education allows for the unknown concept and theme to be questioned, broken down and revisited again and again. As stated by Wheelock (1997), "emerging standards envision learning that shifts toward a greater focus on thinking skills- reasoning, problem solving, making connections, communicating-as the context for learning basic skills and facts" (p.29).

Today's language arts instruction comes from an eclectic pool of resources. Many teachers use trade books, cross curriculum integration, group thinking, cooperative learning, projects, performances, community and parental involvement, technology, and texts to make meaning and insightful connections to learning. According to Villaume, Worden, Williams, Hopkins, & Rosenblatt in 1994, they felt that students were not readers who could identify,

interpret and develop answers from their reading. With many of the changes in programs and state requirements, "Teachers can now see students as builders and explorers who use resources including texts, peers and experience to discover meaning and solve self-defined problems" (Guthrie, Mitchell & Van Meter, 1994, p.266).

### **Standardized Reading Tests**

Unfortunately, most standardized reading tests do not test in the nature of current teaching. In April of 1994, Kapinus wrote, "Concerns focus on problems such as assessment tasks that bear (sic) little or no resemblance to those encountered in good instruction or the world beyond the classroom"(p.578). Another article written by the National Association of Education of Young Children (NAEYC) in 1988, suggests that "standardized testing has led to the adoption of inappropriate teaching practices" (p.43).

In 1987, Bussis & Chittenden indicated that "such testing procedures may direct teachers' attention away from what matters to developing readers" (p.302). Standardized testing limits children's education possibilities and discourages social and intellectual development in the areas of cooperation and creativity (Allison, Haas, & Haladyna, 1998). A study conducted by Dally, Martens, Kilmer & Massie in 1996, found that the more overlap there is between teaching

conditions and testing conditions, the more likely students will generalize when answering questions.

Well-intended practices for testing strategies are not always successful when preparing for standardized reading assessments. "Recent research suggests that repeated experience with standardized testing, particularly for low achievers, have (sic) a cumulative, negative impact on students that can be summarized in three general trends: growing disillusionment about test, decreasing motivation to give genuine effort and increasing use of inappropriate strategies." (Taylor & Walton, 1997, p.67) In addition, Allison, Haas, & Haladyna in 1998 suggest that the level of motivation will begin to decrease and test anxiety will increase, as the student becomes older. Teaching to the standardized test, neglecting curriculum and higher level thinking and processing skills is not working with the ever-changing educational system.

Standardized testing of reading achievement still measures reading exclusively for phonics and word recognition, comprehension, and isolated skill acquisition. One disadvantage of this type of testing is the use of multiple choice questions. "Children are taught to provide the one 'right' answer on the answer sheet, but are not being challenged to think" (NAEYC, 1988, p.43).

On the other hand, performance assessment measure reading in the same areas as standardized testing but takes the test many steps

above and beyond literal thinking. Performance assessment examines the students' levels of comprehension, background knowledge, problem solving strategies, vocabulary, transfer of learning and real life written responses to the readings.

One necessary component to reading is having a wealth of background knowledge or schema. Teachers are consistently building their students' knowledge and providing activities to expand their cognitive ability with regard to the world around them. Unfortunately, as Bussis & Chittenden reported in 1987, tests seek to minimize the influence of background knowledge by posing questions that stick strictly to the information presented in reading comprehension passages.

What are the reasons for such a discrepancy between a standard-based education and assessment, and standardized testing in today's society? In the 19<sup>th</sup> century, standardized tests were created to ensure that all students were receiving the same standard of education (Allison, Haas, & Haladyna, 1998), meant that all students were learning the same way, in the same fashion, in every school.

Standardized tests are intended to measure a school's program, and compare achievement made from one school to another. In addition, at risk schools can be identified, given extra support to improve instruction, and increase future test scores. Standardized testing does very little to go beyond literal thinking, learning and assessing

students. Little has changed from the 19<sup>th</sup> century view on standardized testing but a new way of thinking is beginning to emerge due to students' unsettling test scores and overall academic performance.

Since the age of standardized testing has evolved, people have been critical of this form of testing. Even H.G. Wells, author of War of the Worlds, objected over the influence of external examiners deciding which curricula should be taught. Pearson and Valencia (1987) quote him as saying: The examiner pipes and the teacher must dance- and the examiner sticks to the old tune. If the educational reformers really wish the dance altered they must turn their attention from the dancers to the musicians. (p.731)

### **Standard-based Curriculum**

Today, the term's "standards" or "standard based learning" has little comparison to the past. The new standard-based approach is intended to generate life long learners who have an appreciation to continue to excel in all aspects of intellectual growth. A standard based curriculum expects the learner to demonstrate higher level thinking and processing skills, use problem solving strategies, look for and take a challenge or risk, and seize ownership in their educational career. Educators are now looking at the whole child and including reading, writing, speaking, listening, and social interaction into the realm of

literacy development. As noted by Afflerbach, Kapinus & Delain (1995), "When standards are accompanied by performance indicators or learning goals at specific grade levels, understanding of the discipline can be even richer" (p. 292). Therefore utilizing the students' rich understanding of the academic disciplines and creating a test around the learning that is taking place will evoke a true assessment to the expected learning. Teaching to a test is no longer necessary or a concern since a higher standard of learning is always being valued.

### **Performance Testing**

When students are enabled to receive reading instruction as active learners and are measured with a performance assessment, students will become sophisticated readers who have a sense of ownership, awareness of their individual reading habits and strategies (Pearson & Valencia, 1987). "Performance-based assessment has brought a new slant on these issues, particularly as they relate to evaluating student achievement of the 'outcomes of excellence' previously mentioned" (Afflerbach, Kapinus & Delain, 1995, p.440). Performance assessments are more likely to represent activities that reflect the actual goals and instructional activities of the classroom and the real world outside the classroom. (Strickland, 1995)

The tasks on performance assessment utilize open-ended items that will allow the students to answer by drawing pictures, completing

graphic organizers, and long and short written responses. Kapinus (1994) noted a marked improvement in performance assessment scores over traditional multiple choice and fill in the blank items. Students were able to answer items by drawing on background knowledge, personal experience, creativity and various reading strategies to design their response instead of selecting it from a cluster of choices (Afflerbach, Kapinus & Delain, 1995).

Teachers, administrators, districts and states need to remember to look at the whole child and not let one assessment determine everything. Kapinus cited a word of encouragement in 1994, a principal wrote,

“We finally have a ‘test’ that we should be ‘teaching to.’ We must focus more on the learning process to gather, organize, and interpret information rather than memorization of content facts. And this assessment program... is a step in the right direction. Continue to emphasize futuristic needs; don’t be pushed into backing up...”. (p. 580)

With the ever changing world and the role the educational system plays in today’s society, school districts, states and teachers are increasingly becoming more accountable for their students’ learning and progress. “ Pressure to demonstrate accountability through children’s performance on standardized tests not only changes how teachers teach and what children study, but seems to be changing our very understanding as of the nature of learning and achievement.” (Meisels, 1993, p.35) Eachgroup states a need for a more objective, authentic assessment but standardized testing and scores still out

weigh authentic performance measures. Despite trends toward performance-based assessments, multiple choice tests are commonplace and will be in demand into the next century states Taylor and Walton in 1997. Many school reform advocates criticize the use of state wide multiple-choice tests, but 41 states still continue to use them.

### **Summary**

After a review of the research of both traditional and performance assessments with the underlying thought of a standard-based education, there appears to be strong indications of a need for more authentic forms of assessment. As in all debates, there are always strong opinions for both sides. The development of assessment instrument is only one part of the picture to consider in initiating a new era uniting assessment to curriculum and education. ( Bergan & Feld, 1993)

## **CHAPTER III**

### **Design of the Study**

#### **Purpose of the Study**

The purpose of this study was to determine if there is a statistically significant relationship in achievement between the traditional standardized fourth grade DRP reading test and the NYS English Language Arts Fourth Grade assessment.

#### **Null Hypothesis**

There is no statistically significant correlation between the Degrees Of Reading Power reading scores and the NYS ELA assessment scores.

#### **Methodology**

##### **Subjects**

The subjects were one hundred and thirty-one rural, heterogeneously-grouped fourth-grade students. The ages of the students range from 8-10 years. The students' reading abilities range from advanced readers to students who receive Title One support.

### **Instruments**

The two instruments that will be used for this study are the Degrees Of Reading Power reading achievement test and the NYS Fourth Grade English Language Arts Assessment.

### **Procedures**

As fourth grade students they took the NYS ELA assessment in January of 1999. Following that assessment, the students were assessed in May, 1999 using the Degrees Of Reading Power reading achievement test. The researcher examined the correlation between both tests to see if there was a statistically significant correlation between the two tests.

### **Analysis of Data**

A Pearson Product-Moment Correlation was used to determine if there was a statistically significant correlation between the DRP reading achievement test and the NYS ELA Assessment scores.

## **CHAPTER IV**

### **Analysis of Data**

#### **Purpose of the Study**

The purpose of this study was to determine if there is a statistically significant relationship in achievement between the traditional standardized Degrees of Reading Power reading test and the NYS English Language Arts Fourth Grade assessment.

#### **Null Hypothesis**

There is no statistically significant correlation between the Degrees of Reading Power (DRP) reading test scores and the NYS ELA assessment scores.

#### **Analysis of Data**

The Pearson Product-Moment Correlation Coefficient ( $r$ ) formula was used to find the relationship between the DRP reading test and the NYS ELA assessment scores. Once the calculated ( $r$ ) variable was found it was then compared to the critical value of the Pearson Correlation Coefficient of a two-tailed test at the Alpha .05 level of significance or 95% confidence interval.

## Findings and Interpretations

The Pearson Product-Moment Correlation Coefficient calculated to the value  $(r) = .746$ . The critical value of the Pearson Correlation Coefficient at the alpha .05 or 95% confidence interval with 100 degrees of freedom (df) was equal to a critical value  $(r) = .195$ . Since the calculated  $(r)$  was larger than the critical value, this study will have to reject the null hypothesis and conclude that there is statistically significant positive relationship between both reading tests. As one set of reading scores increased the other test also increased, which indicates a strong positive correlation between both tests. Table 1 below demonstrates this strong positive correlation, as the NYS ELA standard score matched with the DRP converted standard score. Essentially both tests are measuring the same kind of skills.

Table 1

Correlation of NYS ELA and the DRP Assessment Test scores

	<u>NYS ELA Standard Score</u>	<u>DRP Raw Score</u>
<b>Mean</b>	646.33	45.09
<b>Standard Deviation</b>	30.11	10.22
<b>Number of Students</b>	131	131

Obtained Pearson Value  $(r) = .746$

---

Critical  $(r)$  Value = .195

## **Summary**

All the findings indicate a strong positive correlation between the DRP reading test and the NYS ELA assessment. With 95% confidence, this study can expect that students who scored the highest on the ELA assessment should also score high on the DRP reading test, but may not have the same exact score. The same should also be true for the low and middle range scores. This expectation can be permitted due to the finding that essentially the skills measured on both tests are related.

## **CHAPTER V**

### **Conclusions and Implications**

#### **Purpose of the Study**

The purpose of this study was to determine if there is a statistically significant relationship in achievement between the traditional standardized DRP reading test and the NYS English Language Arts Fourth Grade assessment.

#### **Conclusions and Implications**

The results of this study conclude that there is a statistically significant positive relationship between the DRP and ELA assessment test scores. With 95% confidence, this study rejected the null hypothesis. This study can conclude that both tests are essentially measuring the same related skills. Therefore, students who perform high on one test will likely score as high on the second test, but may not have the same exact score.

There are many factors that may have affected the results of this study. Even though both tests were essentially testing the same reading skills, the test format was quite different. In the past students have taken many reading tests that did not require them to answer

questions in written form. The ELA assessment made students more accountable to respond to questions by utilizing their writing ability. Their lack of exposure to the standards may have influenced the performance level to become lower than expected with little variation between the scores.

Another factor that might explain why the NYS ELA assessment had such a strong positive correlation to the DRP is the lack of opportunity to write real life responses within strict time limits. Many students are taught the writing process which requires the students to go step by step taking many days to complete a solid written piece. The students did not have the luxury to go through each writing step slowly and carefully. This hindered many students in completing the short essay questions. Children need to learn the skill of answering these types of questions in a quick, precise way.

### **Implications for the Classroom**

Although there was a significant positive correlation in both tests, many outcomes for the classroom could come from this research. To begin, educators recognize that reading and writing achievement are needed for a student to be successful in both language arts areas. In general, in order to be successful in writing the student would also need to be a good reader. The same assumption is also true for the weak reader. If a student has weak reading skills, it could also reflect

weak writing skills. This study indicated that students who did well on one test would also do well on the other and implied that both tests were essentially measuring the same skills. In view of this finding, it would make sense that each student would score comparable on each test no matter the format.

If further research is conducted along with the findings of this study, it may show that teaching style is not a factor when looking at the students' test scores. Children who are in a classroom with a teacher who uses a performance based teaching approach and assessment versus the traditional teaching approach and assessment will not be at an advantage or disadvantage when taking either of these tests. Depending on their achievement, they should score in the same range on either test.

The most commonly asked question throughout the school year is "How are my students doing and how will they do on the standardized reading test?" In essence, the teachers could give practice DRP tests and find out in general how each student is doing according to the skills tested on each test. The teacher could then see what reading skills the student might need work on in order to be successful on either test. The teacher must keep in mind that the writing component of the ELA is not found on the DRP and would need to assess the student's work by making observations of the student's writing. In the

areas the student is lacking writing skills, the teacher will need to provide further writing instruction.

### **Implications for Further Research**

This study may have had limitations that further research and time could reveal. To begin, this was the first time the NYS ELA assessment was given. A longitudinal study looking at these students at the 8<sup>th</sup> grade ELA assessment and 12<sup>th</sup> grade Regents test may yield different results.

Another study on this topic might look at the students who are presently in kindergarten when they reach fourth grade. By then they will have had more instruction in a standard based curriculum. The result could be higher test scores on the ELA and a greater variation within the student population.

The last reason to research this study further is the fact that the scores were not distributed equally which made it difficult to show a statistical difference between the two tests. A more heterogeneous range of students' abilities may have different results. A selection from various districts may distribute the scores in a wider range. As the scores disperse and become more varied it could cause a change in the correctional value. This may begin to show a statistical difference between the ELA and DRP assessment.

## **Summary**

The intention of this study was to see if there was a relationship between two types of reading achievement tests. The DRP is a standardized reading test that uses multiple choice questions in a cloze format. The ELA is a new standardized test, which uses a variety of performance assessment tasks that require the student to use higher level thinking skills. The Pearson Product-Moment Correlation Coefficient ( $r$ ) Formula was used to measure the relationship between each test. At the 95% confidence interval, this study can conclude that there is a strong positive correlation or relationship between both reading tests. Therefore, it can be implied that both tests measure the same kind of skills. As one test will increase or decrease the same will be true for the other test.

## REFERENCES

- Afflerbach, P., Kapinus, B. DeLain, M.T. (1995) . Equity and performance-based assessment: An insider's view. The Reading Teacher, 48 (5) , 440-442.
- Afflerbach, P., Kapinus, B., Guthrie, J.T., Van Meter, P. Mitchell, A. (1994) . Performance assessments in reading and language arts. The Reading Teacher, 48 (3) , 266-271.
- Bergan, J.R. & Feld, J.K. (1993). Developmental assessment :New directions. Young Children, 48 (5), 41-46.
- Berliner, D. C. & Biddle, B.J. (1996) . Reality-based education standards for all. Education Digest, 62 (3) , 25-28.
- Bormuth, J.R. (1968) . The cloze readability procedure. Elementary English, 45 236-249.
- Bussis, A.M. & Chittenden, E.A. (1987) . Research currents: What the reading test neglect. Language Arts, 64 (3) , 303-308.
- Cordon, L.A. & Day, J.D. (1996) . Strategy use on standardized reading comprehension tests. The Journal of Educational Psychology, 88(2) , 288-295.
- Crowell, D.C., Au, K.H. (1981) . A scale of questions to guide Comprehension instruction. The Reading Teacher, 34 389-393.
- Crowell, D.C., Au, K.H., & Blake, K.M. (1983). Comprehension Questions: Differences among standardized test. Journal of Reading, 26 (4) , 314-319.
- Daly, E.J., Martens, B.K., Kilmer, A., Massie, D.R. (1996). The effects of instructional match and content overlap on generalized reading performance. Journal of Applied Behavioral Analysis, 29 (4), 507-518.

- Flood, J. & Lapp, D. (1993) . Clearing the confusion: A closer look at national goals and standards. The Reading Teacher, 47 (1) , 53-61.
- Haladyna, T., Haas, N., Allison, J. (1998) . Continuing tensions in standardized testing. Childhood Education, 74 (5) , 262-273.
- Kapinus, B. (1994) . Looking at the ideal and the real in large-scale reading assessment: The view from two sides of the river. The Reading Teacher, 47 (7) , 578-580.
- Keck, M.A. (1996). Correlations between teacher assessment of second graders reading abilities and their performance on subtest of cognitive skills. Unpublished master's thesis, University of Brockport, New York.
- Meisels, S. (1993). Remaking classrooms assessments with the work sampling system. Young Children, 48, 34-40.
- Neill, D.M. (1997). Transforming student assessment. Phi Delta Kappan, 79 (1), 34-58.
- NAEYC position statement on standardized testing of young Children 3 through 8 years of age. (1988, March) . Young Children, 40, 42-47.
- Salinger, T. (1995/1996) . IRA, standards, and educational reform. The Reading Teacher, 49 (4) , 290-298.
- Strickland, D.S. (1995). Reinventing our literacy programs: Books, basics, balance. The Reading Teacher, 48 (4), 294-302.
- Taylor, K. & Walton, S. (1997). Co-opting standardized tests in the service of learning. Phi Delta Kappan, 79 (1), 66-67.

Teale, W.H., Hiebert, E.H., Chittenden, E.A. (1997) . Assessing young children's literacy development. The Reading Teacher, 40 (8) , 772-777.

Valencia, S. & Pearson, P. (1987) . Reading assessment: Time for a change. The Reading Teacher, 40 (8) , 726-732.

Villaume, S.K., Worden, T., Williams, S., Hopkins, L., Rosenblatt, C. (1994). Five teachers in search of a discussion. The Reading Teacher, 47 (6), 480-487.

Wheelock, A. (1997) . Standards-based reform and the middle grades. Educational Digest, 62 (5) , 29-33.