2007

Using standardized test reading comprehension software to improve student academic achievement in reading comprehension

Andy James Kubitza

USING STANDARDIZED TEST READING COMPREHENSION SOFTWARE TO IMPROVE STUDENT ACADEMIC ACHIEVEMENT IN READING COMPREHENSION

A Project

Presented to the

Faculty of

California State University,

San Bernardino

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

in

Education:

Instructional Technology

---

by

Andy James Kubitza

December 2007

USING STANDARDIZED TEST READING COMPREHENSION SOFTWARE TO IMPROVE STUDENT ACADEMIC ACHIEVEMENT IN READING COMPREHENSION

A Project

Presented to the

Faculty of

California State University,

San Bernardino

---

by

Andy James Kubitza

December 2007

Approved by:

[redacted]

Dr. Brian Newberry, First Reader — 13.NOV.07 Date

[redacted]

Dr. Randall Wright, Second Reader

ABSTRACT

By 2014, the federally mandated No Child Left Behind Act of 2001 expects all schools and school districts to have 100% of their students perform at or above the Proficient Level on the state's Language Arts and Mathematics standardized tests. The mandate and these tests have thus transformed the teaching of reading and mathematics into high stakes subject areas that have become the focus of much debate. This experimental quantitative design research study examined whether a web-based Standardized Test Preparation Intervention for Reading Comprehension was more effective and efficient in improving student academic achievement in reading comprehension than a paper-based Standardized Test Preparation Intervention. Fourth grade elementary students in an urban elementary school district were measured by the Woodcock-McGrew-Werder Mini-Battery of Achievement Reading Test prior to and after the use of the Standardized Test Preparation Interventions for Reading Comprehension. Results showed there was not a significant difference in Reading Comprehension improvement among users of the web-based Standardized Test Preparation Intervention versus users of paper based Standardized Test Preparation Intervention. The results of this study

suggest that while a web-based Standardized Test Preparation Intervention for Reading Comprehension may help to improve student's academic achievement in reading comprehension, and thus perform better on a standardized test, a paper-based version of Standardized Test Preparation Intervention for Reading Comprehension may be just as effective.

ACKNOWLEDGMENTS

I would first like to acknowledge and thank my wife, Sandee, who has put up with me and endured the many trials and tribulations throughout this journey of discovery. Thank you for your patience, and your ability to ignore my messes, and for the extra time and work that you have had to put in to keep our home running efficiently. A special thank you to my mom, DeAnn, for instilling in me the desire to "shoot for the stars". I would like to especially thank Gordon and Carol Ray for their support throughout my life and academic endeavors. I appreciate your tolerance of all the complaining and whining I am sure that I have put you through during this experience. I want to say Hi to Alex, my six-year-old son. I am done. I do not always have to be sitting at the computer anymore. And *yes*, I *can* play. I would like to thank Dr. Brian Newberry for everything, but most notably for the introduction to HTML, JavaScript, web design, PHP and MySQL, Constructivism and Behaviorialism. Thank you for opening the door to iMovie, an incredible experience. I would like to thank Dr. Randall Wright for his guidance and leadership while serving as professor and reader. Thank you for your advice and knowledge of the subject areas and your ability to communicate often-abstract ideas

in a clear and concise manner. A big thank you goes to the San Bernardino City Unified School District for helping me execute my study. Your commitment to the well being of our teachers and students inspires us all greatly. Another big thank you goes to the parents and students of the San Bernardino City Unified School District's Highland Pacific Elementary School who participated in our study: Thank you for your genuine interest, input, effort and constructive criticism during this journey. I would like to thank Dr. Paul Shirk for all of his help at the district level. I also would like to thank Principal Brad McDuffee for allowing me the freedom to explore more efficient and effective methods of educating students. Thank you again to Alex who managed to pull me away from this thesis project to enjoy a sweet distraction such as dinner or a weekend at the park. You helped me refocus my priorities whenever I lost sight of what was important in life.

DEDICATION

To my wife Sandee, who gave me her heart when I was young. You are the best thing that has happened to me. Thank you for seeing in me all the wondrous abilities that took me so long to see.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

BACKGROUND

Introduction

In 1983, a report entitled *A Nation at Risk* (National Commission, 1983) determined the public school system was not producing students with the skills needed to meet the challenges of the day. By 2002, the federally mandated *No Child Left Behind Act* (NCLB) was enacted with the expectation that all schools and school districts must have 100% of their students perform at or above the Proficient Level on a state's Language Arts and Mathematics standardized tests by 2014. School districts, such as the San Bernardino City Unified School District, have repeatedly reported low growth or failed to demonstrate any progress toward achieving state academic standards as measured by Adequate Yearly Progress (AYP) scores. Thus, these schools and school districts have been identified as "in need of improvement", and the state has placed them under a "Corrective Action Plan." The school improvement status or "Corrective Action Plan" has placed even more pressure on teachers, students and staff to perform better and raise standardized test scores. Therefore, the accountability demands of the NCLB Act have

put great pressure upon teachers to teach to the standards as set forth by the state of California. NCLB expects all students to meet those standards, and to demonstrate their proficiency of those standards by performing at or above the Proficient Level on a state's Language Arts and Mathematics standardized tests.

Existing research shows that low reading levels in students greatly affects their academic success (Armbruster, Lehr, & Osborn, 2003). Low reading and comprehension skills hinder the student's ability to perform well in school, as well as on multiple-choice standardized tests. This is especially true in low socio-economic, inner-city schools.

In most public schools, reading comprehension is assessed in three typical formats. The first and most difficult format for students is the essay question. Usually this question requires the reader to write a short essay on a given topic. This is exemplified by the numerous book reports, research papers and other writing projects that require students to read and respond to narrative or expository text.

The second format used to assess reading comprehension is the open-ended short answer question that so predominates the textbooks of Science and Social

Studies from Third grade through Graduate School. The textbook is generally read first and then at the end there are a number of comprehension questions to answer. These short answer questions are answered in one sentence to one paragraph in length. Students are afforded daily opportunities to work with these two methods in reading and reading comprehension.

The third format used to assess reading comprehension is not so readily apparent in the daily school regimen. The multiple-choice question format for assessing reading comprehension may be the easiest of the testing formats for students to do, yet it is often the most controversial of the three methods. However, this is the format that has been adopted by the state of California, in the form of the California Achievement Test (CAT6), to assess students in Language Arts and mathematics. Unfortunately, the students have had very few opportunities to practice with this format.

In an attempt to provide a resource to practice the multiple-choice test format, to help these students improve their reading comprehension skills, and to help students perform better and raise their standardized test scores, a web-based Standardized Test Preparation Intervention for Reading Comprehension was designed and

developed. This work evaluated this intervention in terms of how effective the intervention helped the students to develop reading comprehension skills and perform better on standardized tests. This study measured the result of this web-based Standardized Test Preparation Intervention for Reading Comprehension program and determined whether such a program improved the reading comprehension skills of students.

At the inception of the study, an initial assessment of the students' reading levels was conducted. After completion of this first assessment, the correct responses were tallied and this raw data was used to set a baseline score from which to evaluate the results.

Two groups, a control group and an experimental group, were formed. The experimental group practiced with the web-based Standardized Test Preparation Intervention for Reading Comprehension program while the control group practiced with a paper-based version of the same intervention.

After the three week implementation period, the students were assessed a second time. After completion of this second assessment, the correct responses were tallied and this data was used to determine any growth in the students' reading levels.

The two groups were then reversed. For the next three weeks, the control group became the experimental group and practiced with the web-based intervention while the experimental group became the control group and practiced with the paper-based intervention.

At the end of the second three-week time of implementation, a third and final assessment was then given to measure the students' reading level and determine whether the students had improved their reading comprehension as a result of practicing with the intervention.

The purpose of this study was to investigate whether a web-based Standardized Test Preparation Intervention for Reading Comprehension was more effective than a paper-based version of the intervention to remedy, to some degree, the low reading comprehension levels of students within a target classroom of elementary school students in the San Bernardino City Unified School District.

Statement of the Problem

Reading and reading comprehension have been an important issue in American education for a long time. In fact, it has been said that reading and reading comprehension are the heart of education. When students

can read, students are able to succeed across the curriculum. The institution of the federal program No Child Left Behind, has made student academic achievement in reading a top priority. The standardized testing that has been done in context with the law has indicated that students across the nation have poor reading and reading comprehension skills. The students in the San Bernardino City Unified School District were much the same.

Within the San Bernardino City Unified School District, many students tended to lag or under-perform academically. They experienced a great deal of difficulty when attempting to access grade level texts across all core subjects. As a result, grades and performance declined, student interest waned, and ultimately, students entered a cycle of poor performance and a general lack of academic success. The central problem stemmed from the fact that reading and reading comprehension levels were low.

This problem has developed over many years, and many factors have contributed to the students' poor academic achievement in reading comprehension. Many of the students came from low-income families that received aid and free or reduced lunches. These low-income students comprised 88% of the student population. The attendance rate was

very poor. Student mobility was very high, and many of the students' parents had poor educations or did not value education. While class size was a reasonable 20 students per teacher in first through third grades, that number soared to 34 students per teacher in grades four through six. The students themselves were unmotivated to learn. They did not read outside of the "forced" situations of specific classroom activities. They experienced quick frustration when classroom texts seemed too difficult, and they quickly assumed that there was something fundamentally wrong within them. They imagined that they were "stupid," or they labeled themselves as "dumb." All of these perceptions discourage students from reading and ultimately trap students in a paradigm of recurring academic failure, for if a student fails to read, a student simply fails—in all aspects of education and across the curricular spectrum.

Students in the author's fourth grade class also had low, inadequate reading comprehension skills, which interfered with their academic achievement. Based on standardized test scores and teacher observation, there was a lack of comprehension in all types of reading. The students were able to read the words, but were having difficulty in understanding what they read and answering

questions about what they read. Students were having problems with fiction and nonfiction, as well as expository and non-expository reading formats. Non-expository reading has been defined as reading in formats such as stories, fairytales and myths. Expository text, on the other hand, is often first encountered in the fourth grade when students began studying topics such as Science and Social Studies. Expository text has generally been deemed as more difficult for students to comprehend because of the vocabulary and the student's lack of background knowledge.

Students were also having difficulty in comprehending other types of text. Directions or steps in a series or process, poetry, information in a letter, and information in an advertising format are all examples of what the students need to be able to read and comprehend. In addition, many students have shown a need for better understanding of information presented in a Table of Contents, Indexes and Timelines. These literary forms are encountered throughout elementary school from Kindergarten through the sixth grade, and are tested by the state's California Achievement Test.

There was also a need for the students to practice reading comprehension with long reading passages. This was

especially apparent in the early primary grades (first, second, and third). Students at second grade, for example, had to be able to read and comprehend a reading passage as long as 300 words. Based on the California Achievement Test Released Questions, students at the fifth grade level were expected to read and comprehend reading passages of 600 words.

In addition, students needed to be able to read and comprehend both single and multi-paragraph reading passages. Single paragraph reading passages were often more prevalent at the primary level, yet they made up less than 30% of the reading passages students were expected to read. Students at the primary level, essentially second and third grade, needed to be able to read multiple paragraph reading passages that average 6 paragraphs in length. Reading passages, at the primary grade level, ranged as high as ten and eleven paragraphs in length. At the upper grade level, grades 4 - 6, students needed to be able to read long multi-paragraph reading passages that ranged from four to fifteen paragraphs. Based on the California Achievement Test Released Questions, the average reading passage that an intermediate grade level student needed to be able to read and comprehend was seven to eight paragraphs in length.

Students must also be able to answer a wide variety of types of comprehension questions. The student needed to be able to answer questions from all six of the main categories compiled in Bloom's Taxonomy. These questions ranged from simple Knowledge questions such as recalling information from the text to more complex and difficult questions such as from the Application category in which students were asked to solve problems and apply information to produce a result. Students also needed to be able to answer Compare and Contrast questions that asked the student to read two or more reading passages and then compare them.

Students across the nation have exhibited poor reading and reading comprehension skills. Many factors have contributed to why students have exhibited poor academic achievement in reading comprehension. The California Achievement Test (CAT6) has shown that there were a variety of different reading comprehension skills and reading formats that needed to be learned and mastered by fourth grade elementary students in order for them to be successful in education. The central problem was that students were unable to succeed in education because they have not learned and mastered the different reading comprehension skills or reading formats.

Purpose of the Project

Teachers in California have been assigned the responsibility of insuring that every student in their class score at or above the Proficient level on the state standardized test or suffer severe penalties. Teachers in the elementary school setting were already challenged daily with a variety of factors, such as classroom management, poor attendance and students with different ability levels, that prohibited their ability to teach academic skills and content. Faced with these prospects, the elementary teacher, particularly those working in inner city, low socio-economic elementary schools, have had to come up with their own reading and reading comprehension interventions. Invariably, teachers have had to make adjustments "on the run" and "reinvent the wheel" for every new challenge presented by each new class.

In response to these new challenges, many different approaches have been used to improve reading comprehension skills and help raise academic achievement in reading comprehension. However, the results of the California Achievement Test (CAT6) have shown that in many areas little progress has been achieved (California Department of Education, 2007a). It was imperative that an intervention be developed to help improve the reading

comprehension skills of students. Successful intervention was needed to place requisite academic tools such as reading skills within the reach of students and ultimately increase student performance. More importantly, it was needed to improve the future of our students by opening opportunities that would otherwise be closed to them.

This work sought to address the need for such interventions, and to determine whether a web-based intervention or a paper based intervention was the more effective and efficient intervention to improve students' reading and reading comprehension skills. Computers have an untapped potential in the development of reading and reading comprehension skills. It was initially hoped and expected that the web-based intervention would be the more effective and efficient intervention to improve the students' reading and reading comprehension skills.

The purpose of this experimental quantitative design research study was to investigate whether a web-based Standardized Test Preparation Intervention for Reading Comprehension was an effective and efficient intervention to remedy, to some degree, the low reading comprehension levels of students within a Fourth grade classroom of elementary school students in a large urban elementary school district in Southern California.

Research Question

This work asks the question: "How will the use of a web-based Standardized Test Preparation intervention by 4th grade elementary students affect their academic achievement in Reading Comprehension as measured by the Woodcock-McGrew-Werder Mini-Battery of Achievement Reading Test vs. students who work with a paper-based version of the Standardized Test Preparation intervention?"

Significance of the Project

Students must display the ability to answer comprehension questions correctly at a higher percentage rate. Every student needs to be able to answer correctly a variety of comprehension questions based on Bloom's taxonomy and comprehend what they read at or above the student's grade level. In other words, according to the California Department of Education, (2007b, para. 6), all students in grades one through six [must be able to] "...read and understand grade-level-appropriate material. They [must be able to] draw upon a variety of comprehension strategies as needed, including generating and responding to essential questions, making predictions, and comparing information from several sources." In grades five and six they must also be able to "...describe and

connect the essential ideas, arguments and perspectives of text by using their knowledge of text structure, organization, and purpose" (California Department of Education, 2007b, para. 6). These elementary reading and reading comprehension skills are essential not only to the student's academics, but also and more importantly to the student's quality of life when they reach adulthood.

Reading is the heart of education. When students can read, students are able to succeed across the curriculum. It is imperative that intervention be made in the case of students in the San Bernardino City Unified School District to ensure that all students can read and read with comprehension at a much higher level or rate than has ever been expected before. Successful intervention will develop reading skills within all students and ultimately increase student performance.

Successful intervention will have developed reading skills to such an extent that students can stop learning how to read, and begin to read to learn. Students, who learn to read well, will benefit by having new information, knowledge and opportunities that would otherwise have been closed to them. Students will grow up and be able to give back to the community rather than be a burden upon it. This is of the greatest importance and

significance to our students, our schools and our society. Successful intervention, most importantly, will improve the future of all people.

Hypothesis

This study examines the hypothesis that fourth grade elementary students will perform better in Reading Comprehension as measured by the Woodcock-McGrew-Werder Mini-Battery of Achievement Reading test using a web-based Standardized Test Preparation intervention than students who work with a paper-based version of the Standardized Test Preparation intervention.

Students will perform better with the web-based version of the Standardized Test Preparation intervention because it offers immediate grading and feedback on the reading passage. In addition, with the web-based version, the students will receive positive reinforcement through the presentation of Rock and Roll songs when they successfully complete a reading passage. However, we must consider the rival hypothesis. It is possible they will perform worse, or students may perform better with the paper-based version of the Standardized Test Preparation intervention because it is in a familiar form.

Null Hypothesis

We must consider the null hypothesis. Which is: There will be no difference in student reading comprehension performance as measured by the Woodcock-McGrew-Werder Mini-Battery of Achievement Reading Test between the students using the web-based version or the paper based version of the Standardized Test Preparation intervention.

Limitations and Delimitations

During the development of the project, a number of limitations and delimitations were noted. These are presented in this section.

The study was delimited by examination of reading comprehension of students within the Inland Empire. The study was further delimited by its examination of students within the elementary school setting of San Bernardino City Unified School District, and then further delimited to fourth grade elementary school students. The study was also delimited by an examination of reading and reading comprehension as related to the mandatory standardized testing facet of the *No Child Left Behind Act* (NCLB).

The study was limited by the small number of participants who took part in the study. The study was also limited by the amount of time that was allotted to

practice with the intervention. The study was further limited in that the study was a pilot study by a first time researcher.

Definition of Terms

The following terms are defined as they apply to the project.

Standardized Testing: a test administered and scored in a standard manner. The tests are designed in such a way that the questions, conditions for administering, scoring procedures, and interpretations are consistent and are administered and scored in a predetermined, standard manner.

Adequate Yearly Progress (AYP): A nationwide accountability system mandated by the No Child Left Behind Act of 2001 that requires each state to ensure that all schools and districts make Adequate Yearly Progress.

No Child Left Behind Act of 2001 (Public Law 107-110): Commonly known as NCLB, is a United States federal law that reauthorizes a number of federal programs that aim to improve the performance of U.S.'s primary and secondary schools by increasing the standards of accountability for states, school districts and

schools. It also sets the accountability system called Adequate Yearly Progress (AYP).

Academic Performance Index (API): The cornerstone of California's Public Schools Accountability Act of 1999. It measures the academic performance and growth of schools on a variety of academic measures.

CHAPTER TWO

REVIEW OF THE LITERATURE

Introduction

In order to raise the academic achievement of the student both on standardized tests and in reading comprehension, the instructor needs to gain a full understanding of the problem by determining what has been done and what has been discovered to be effective in making gains. In other words, the instructor needs to make use of research, and be grounded in sound educational theory. Teachers need to study diverse learning theories to optimize the learning of their students. With a thorough knowledge of the psychology of learning, teachers may do a better job of teaching to pupils of all levels (Ediger, 1999). If an instructional program ignores what is known about educational theory, then learning is left to chance (Smith, 1999). Chance is just what has been happening for the past 25 years. Teachers, administrators and school districts have swung on the proverbial educational pendulum back and forth, and the students have suffered for it. Educational philosophies such as Cognitivism and Constructivism have been distributed in whole or in part to elementary teachers and have generally

failed to significantly raise standardized test scores. Now, with the legislation of the No Child Left Behind education law, the state of California has begun an accountability system in which students are evaluated with a test steeped in Behaviorism. While there may be many different educational philosophies, Behaviorism just naturally seems best suited to take on the job of raising the academic achievement of our students.

Behaviorism

Behaviorism is defined in a variety of different ways. Behaviorism is an approach to psychology based on the proposition that behavior is interesting and worthy of scientific research. In simple terms, behaviorism is the study of the observable behavior of man. It examines the observable actions and reactions of an individual. Behaviorism was the first psychology that looked at human behavior and how humans learn (Smith, 1999). Behaviorism is a philosophy predicated on change and lifelong learning. Change in observable behavior is considered learning. According to the behaviorist philosophy, there is not a point at which learning stops for any living human being.

This basic philosophy has been around for a long time. Its origins have been traced back to the work of Aristotle. "One of the earliest explanations of learning came from Aristotle (384-322 B.C.). He said that we remember things together (1) when they are similar, (2) when they contrast, and (3) when they are contiguous" (Woolfolk, 1995, p. 199). However, it is John Broadus Watson, an American psychologist, who is recognized as establishing the psychological school of behaviorism. His article, "Psychology as the Behaviorist Views It," published in the Psychological Review in 1913, is considered a landmark in the founding of behaviorism.

John B. Watson argued for the value of a psychology that concerned itself with behavior in and of itself, and not as a method of studying consciousness. "Psychology, as the behaviorist views it, is a purely objective, experimental branch of natural science which needs introspection as little as do the sciences of chemistry and physics" (Watson, 1913, Summary). This was a significant shift away from Structuralism, the psychology of the day, which used the method of introspection (the thinking about one's own internal state) and regarded the study of behavior of no value. With behaviorism, Watson put the emphasis on the external behavior of people and

their reactions to given situations, rather than on the internal, mental processes of those people. In other words, Watson called for the study of the observable behavior of men and animals, not of their experiences, thoughts or feelings. In his opinion, the analysis of behaviors and reactions was the only objective method to get insight into human actions.

John B. Watson proposed his Behaviorist theory based on the works of Russian physiologist Ivan Pavlov. Ivan Pavlov accidentally discovered the phenomenon of classical conditioning (learned reflexes) in his study of the digestive system of the dog, and subsequently investigated the phenomena in detail. This animal training model is known as stimulus-response or Classical Conditioning.

Classical Conditioning

In classical conditioning, subjects are taught to react automatically and involuntarily to a stimulus that previously had a different or no effect on them (Woolfolk, 1995).

Classical conditioning was discovered by Ivan Petrovich Pavlov, a Russian physiologist, in 1904 when he won the Nobel Prize for his work in animal digestion. It focused "... on the learning of involuntary emotional or

physiological responses such as fear, increased heart beat, salivation or sweating ..." (Woolfolk, 1995, p. 199). These involuntary actions are often called respondents. Pavlov's discovery repeatedly paired two stimuli together. A neutral stimulus that first had no effect on a subject would be paired with a second stimulus that caused a response. Through repetition, the neutral stimulus would become learned or associated with the other unrelated stimulus and cause a response of some kind (Shaffer, 1994). For example, an animal can be taught to salivate at the sound of a bell (Diaz-Rico & Sandlin, 1995).

Research has confirmed that classical conditioning has a strong role in shaping a person's attitudes and prejudices. Classical conditioning may also be involved in the shaping of our fears, phobias and other emotionally related responses (Shaffer, 1994).

Stimulus and Response

Stimulus and response is the first of three main ideas associated with behaviorism. In the idea of stimulus and response, all response can be traced back to a stimulus. The repeated pairing of the stimulus and response can then cause them to become associated. The

learned associations between stimuli and response are the building blocks of human development (Shaffer, 1994).

Watson understood the concept of stimulus-response and applied it to children and students. He believed that children had no inborn tendencies, but rather were shaped by their environments. Watson believed that only the simplest of human reflexes (for example, sucking and grasping) are inborn and that all significant aspects of one's personality are learned (Shaffer, 1994). Therefore, children were largely influenced by their parents and other significant people in their lives. For this reason, Watson believed that parents must train their children and instill good habits.

Though other well-known twentieth-century behaviorist researchers such as Edward L. Thorndike and Clark L. Hull acknowledged that behavior was either the only method or the easiest method of observation in psychology, it was B.F. Skinner who brought Behaviorism to the forefront of American educational philosophy.

Burrhus Frederic Skinner carried out experimental work in comparative psychology from the 1930s to the 1950s, and remained behaviorism's best known theorist and proponent until his death in 1990. He developed a distinct kind of behaviorist philosophy, which came to be called

radical behaviorism. He also started a new version of psychological science, which he called behavior analysis or the experimental analysis of behavior. This branch of psychology aimed to develop a theory of behavior based on principles of learning. Skinner conducted research on shaping behavior through positive and negative reinforcement. He also developed a behavior modification technique called operant conditioning.

Operant Conditioning

The second major idea of behaviorism is that of Operant Conditioning. In classical conditioning, learned involuntary or unintentional responses are elicited by a conditioned stimulus. Operant (or instrumental) conditioning is quite different. Operant conditioning requires the learner to first give a response of some kind, and then the learner can begin to relate this response with the resultant outcomes, or consequences (Shaffer, 1994).

Operant conditioning involves the use of positive and negative consequences (Slavin, 1991). When a response or act is followed by a reinforcing consequence, then the future probability of the response increases. When a response or act is followed by a punishing consequence,

then the future probability of the response decreases (Skinner, 1963). Thus, operant conditioning is a learning situation in which voluntary behavior is strengthened or weakened by consequences or antecedents (Diaz-Rico & Sandlin, 1995). Consequences are defined as events that follow an action. Consequences reveal not only the first attainments of learning, but also the fluency of that learning which is an important consideration for future learning that builds upon previous learning. There are two main types of consequences in operant conditioning: reinforcement and punishment (Shaffer, 1994).

Reinforcement

A crucial contribution in the area of operant conditioning by Skinner was his clarification of the concept of reinforcement: "... Skinner (1953) proposed that the vast majority of behavior ... is motivated by external stimuli-reinforcers and punitive events - rather than internal forces, or drives" (Shaffer, 1994, p. 80).

Reinforcers are consequences that promote operant learning by increasing the likelihood that the response will occur in the future (Shaffer, 1994). In other words, a reinforcer is a consequence or event that follows an action and promotes or strengthens a behavior or causes

the behavior to increase in frequency or duration. Reinforcement can be further defined as an event that was experimentally observed to increase the rate of response for that subject at that particular time. This event is referred to as a reinforcer. Food, water, brain stimulation, sex, social contact, and drugs are all reinforcers that have been used in operant research with animals.

There are two main categories of reinforcers. Primary reinforcers satisfy biological desires. Food, water, security, warmth, and sex are examples of primary reinforcers because they all satisfy basic human needs (Slavin, 1991). Primary reinforcers are very effective because they do not have to be associated with other reinforcers.

Secondary reinforcers are only effective when they are associated with primary reinforcers. For example, "Grades have little value to students unless their parents notice and value them, and parents praise is of value because it is associated with love, warmth, security, and other reinforcers" (Slavin, 1991, p. 104).

There are three basic types of secondary reinforcers. Social reinforcers, the first type, come in the form of praise, smiles, hugs or attention. A second type, activity

reinforcers, are given when access to games, toys, or other fun activities are used to reward a child or student. Reinforcers that are more sophisticated are referred to as token or symbolic reinforcers, the third type. Items such as points, grades, and most importantly money, can be earned by individuals and then exchanged for other reinforcers (Slavin, 1991).

Reinforcement can be either positive or negative. A positive reinforcer strengthens a behavior when it is presented to a subject. A negative reinforcer increases the likelihood of a behavior happening again when it is withdrawn (Skinner, 1963). It is important to remember that both positive and negative reinforcements are used to increase a desired behavior.

Positive reinforcement is a pleasant or positive incentive. They are often thought of as rewards for a particular behavior. Positive reinforcement strengthens a behavior by awarding a desired incentive after the occurrence of a preferred behavior.

There are many different types of positive reinforcement. Informational consequences, such as identifying improvements or accomplishments with special marks or certificates are often used. There are both material reinforcers such as prizes or small gifts and

non-material reinforcers such as that of verbal praise or a smile on a teacher's face. In the educational system, teachers may use what is called "artificial reinforcers" such as stickers or free time. Skinner preferred natural reinforcers to artificial reinforcers such as "tokens" or verbal expression. Yet contemporary applications of behavior analysis commonly make use of artificial consequences.

The use of positive reinforcement in the classroom requires three things. Prior to instruction, the teacher needs to announce what the students are to learn, the amount of learning that must be acquired before the pupil secures the reward as well as the reward that will be given to the student if the goal is achieved. Working for this reward is the motivator for the learner. The reward must motivate pupils to achieve the goal at a high rate. If the reward does not motivate the student to achieve the goal, then the reward is of no use and should be replaced by one that does motivate the student.

Positive reinforcement has been determined to be the most desired procedure in education because of its positive long-term effects. It is appropriate for all learners. The use of positive reinforcement is also highly flexible with little in the way of problematic side

effects. The main concern, however, is to insure that the short-term consequences are aligned with long-term consequences. In other words, it is important to be reasonably sure that the immediate behavior that is increased does not lead to delayed aversive consequences. Unfortunately, positive reinforcement is often more difficult to deliver, especially on a consistent basis, than negative reinforcement or punishment.

Negative reinforcement is a disagreeable stimulus that is removed from the situation once a specific response has taken place. It is very important to understand that negative reinforcement is not punishment. Negative reinforcement strengthens a behavior by removing an unpleasant stimulus. "To illustrate: we have all been in cars in which an obnoxious buzzer sounds until we buckle our seatbelts. "Buckling up" becomes a stronger habit through negative reinforcement - that is, we learn to fasten the seatbelt because this act ends the irritating noise" (Shaffer, 1994, p. 81).

Behavior can be shaped with the repetition of positive or negative reinforcement (Smith, 1999). The effect appears to depend upon the rate at which reinforcement is given. In general, the more often reinforcement is given, the faster the conditioning takes

place (Skinner, 1948). In addition, consequences that immediately follow the desired behavior are more effective than consequences that are delayed for any amount of time (Slavin, 1991). This is sometimes referred to as Immediacy of Consequences. "A smaller reinforcer given immediately generally has a much larger effect than a large reinforcer given later" (Kulik & Kulik as cited in Slavin, 1991 p. 107-108). It must be remembered, however, that repeated attempts and extra time are also needed before the effects of positive reinforcement can be seen.

Punishment

Punishment is an unpleasant or unwanted consequence. Punishment involves the presentation of an unpleasant stimulus when a child, student or subject does something that is deemed a wrong or bad behavior (Shaffer, 1994). Punishment, therefore, is a method for coping with undesirable behavior.

The purpose of punishment as a reinforcer is not to strengthen a behavior, but to reduce the occurrence of a student's negative behavior (Shaffer, 1994). Its aim is to reduce or discourage undesirable behavior, and decrease the likelihood that it will happen again. Punishment does not eliminate a particular behavior. Punishment only

suppresses a behavior, and does not direct a student or child toward positive behavior.

There are two types of punishments. The first type of punishment is called presentation punishment. Presentation punishment is referred to as Type I punishment. Presentational punishment is used to decrease a behavior by presenting a negative stimulus or consequence after the behavior has occurred.

There are a variety of different types of presentation punishment. Punishments can be anything that weakens or suppresses an undesirable behavior. Punishment can range from a simple caution or warning to social isolation (timeout), in which the student is removed from the classroom environment for 5 to 10 minutes. They can be demerits, extra work, running laps or the demand to write the rules 100 times. They can range from a scolding to a reprimand. A reprimand is a criticism for negative or undesirable behavior. Spanking may even be recommended and used as a presentation punishment for dangerous behavior.

The second type of punishment is called removal punishment. Removal punishment is referred to as Type II punishment. Removal punishment is used to decrease the behavior by removing something pleasant. Punishment through the loss of reinforcers is referred to as response

cost. When privileges are taken away after a student or child has behaved in an undesirable manner, the teacher or parent is using removal punishment (Woolfolk, 1995).

The topic of punishment is very controversial due to its many negative ramifications. Although some contemporary behaviorists have argued for extenuating circumstances that justify its use, this is not the view shared by most researchers and educators. Contrary to popular belief, Skinner himself did not advocate the use of punishment. He repeatedly condemned the use of punishment. His research suggested that punishment was an ineffective way of controlling behavior, leading generally to a short-term behavior change. The effectiveness of the punisher tends to "wear off" as those who receive them get used to them. Punishers may then find they need to increase the intensity of the punishment for it to continue to be effective. Finally, when punishment is stopped, the student or child's behavior gets worse again (Slavin, 1991). In addition, "... punishment may have some undesirable side effects, such as making the child angry or resentful toward the punitive agent. There is even some evidence that punishment can backfire and produce effects opposite to those intended ..." (Shaffer, 1994, p. 82).

Punishment is meant to suppress negative or undesirable behavior. It does not teach new behavior or produce any long-term changes in behavior (Shaffer, 1994). Therefore, positive, rather than negative, reinforcement has proven to be more effective in bringing about lasting changes in behavior.

Reinforcement Schedule

A reinforcement schedule or schedule of reinforcement is a term that refers to how often reinforcement is given (Slavin, 1991). A reinforcement schedule proposes that reinforcement should occur in a consistent, structured format. There are many different types of reinforcement schedules.

With a continuous reinforcement schedule, every time a desired behavior occurs, a reinforcer is given to the student. When people are learning a new behavior, they will learn it faster if they are reinforced for every correct response (Woolfolk, 1995).

On the other hand, an intermittent reinforcement schedule only presents a reinforcer after some, but not all of the instances when a desired behavior occurs. In order to maintain a behavior that has been mastered, students should be reinforced only occasionally for the

behavior rather than every time. This helps maintain the behavior without the student growing to expect constant reinforcement (Woolfolk, 1995). There are four types of intermittent reinforcement schedules: fixed-interval, variable-interval, fixed-ratio, and variable-ratio.

A fixed interval reinforcement schedule awards reinforcement to the student after a set period of time (Woolfolk, 1995). This type of schedule offers predictability to the student. A weekly spelling test is a good example of the use of a fixed interval reinforcement schedule.

The variable interval reinforcement schedule awards reinforcement to the student after varying lengths of time (Woolfolk, 1995). Pop quizzes are a good example of variable interval reinforcement. This type of reinforcement has an element of unpredictability so that the student has to exhibit greater persistence with the desired behavior in order to achieve the desired reward.

In a fixed ratio reinforcement schedule, reinforcement is given after a predetermined amount of work has been completed (Woolfolk, 1995). A computer program that shows students a reward after every seventh correct problem is a good example of the use of a fixed ratio reinforcement schedule.

In a variable ratio reinforcement schedule, reinforcement is given after a student has completed a varying amount of answers. The student may receive reinforcement after seven attempts the first time, after 15 attempts the next time and after only 2 attempts the next time. Pay offs from slot machines are based on the principle of variable ration reinforcement (Diaz-Rico & Sandlin, 1995).

People give more effort and work at a faster pace when they are paid based on ratio reinforcement schedule rather than on interval reinforcement schedule (Woolfolk, 1995). For example, when paid by the piece (piece work) rather than paid by the hour people work harder and faster in order to maximize their pay off. On the other hand, people will quickly give up when the reinforcement does not come when expected or does not meet their expectations. To encourage effort and work, variable reinforcement schedules are most suitable (Woolfolk, 1995). However, it is important for educators to make sure that they try to reduce the reinforcement schedule, so that the students exhibit the appropriate behaviors and learn on their own, without relying on reinforcement (Smith, 1999).

Programmed Instruction

The third of the three major ideas of behaviorism is that of programmed instruction. Programmed instruction was first introduced by B. F. Skinner at Harvard in 1954. It was based on the principles of operant conditioning and the theories of learning. It was intended to free teachers from repetitive drills found in basic academic subjects such as spelling and arithmetic. Skinner believed programmed instruction was superior to traditional instruction because the students were rewarded immediately for correct answers rather than waiting for a teacher to correct written answers.

Programmed instruction became very popular in the 1960's. In fact, by 1962, even the United States President's Science Advisory Committee (PSAC) supported the use of programmed instruction and other behavioral science methods to improve the quality of American education (Casas, 2003). The field of education embraced this new teaching method. However, most programmed instruction was put into book form. This was called programmed textbooks. Programmed instruction in book form had one major disadvantage. It could not prevent students from looking at the answer before writing their own answers. "The research indicates that no university or

school district adopted the technology as a standard medium for instructing all students" (Casas, 2002, p. 18). Furthermore, it was felt by teachers and administrators at the time that the new technology was not helping students to learn the material. So, by 1968, the printing of programmed instructional materials for the classroom was stopped by educational publishers (Vargas, 2005).

However, programmed instruction continued to be used as an instructional method. "... Individually Prescribed Instruction (IPI) [was] probably the most widely used programmed instruction method when this approach was at its peak of popularity in the mid-1970s" (Slavin, 1991, p. 299).

Programmed instruction has been shown to facilitate the learning of content because it incorporates many important principles of learning (Fernald & Jordan, 1991). These principles include clear behavioral objectives, small steps, logical sequencing, active responding, immediate feedback, and drill and practice.

The construction of programmed instruction begins with determining what exactly is to be learned, and then stating this as a behavioral objective. These behavioral objectives are the content that is to be learned. They need to be written as plainly and specifically as

possible. These objectives also need to be written so that they can be specifically measured. If an objective can be measured, it can then be determined if the objective has been achieved.

The performance criteria then need to be determined and clearly spelled out (Fernald & Jordan, 1991). In other words, the performance criteria need to be unmistakably stated so that all stakeholders understand what is expected, and what is considered a passing grade.

All lessons and learning opportunities are designed and aligned with the stated objectives so that the students can achieve mastery of the stated objectives.

Finally, the assessment procedures are clearly written or orally stated so that they are meaningful to the test taker. Testing is also designed to correspond with the behaviorally stated, specific objectives of instruction so that what is tested is what the students have actually practiced and learned (Ediger, 2000).

Programmed instruction is a method of presenting and teaching information. Material is presented in a logical sequence of steps in an individualized instructional manner. The students work on the self-instructional materials at their own speed and at their appropriate level (Slavin, 1991). The overall organization is designed

so that students can achieve mastery of the content. The theory is that students will learn faster and retain information longer when the material is presented in a series of small, logically related steps.

Essentially, programmed instruction consists of small lessons that must be mastered in order to progress on to the next level. The programmed instructional materials are broken down into small subskills, so that students may work on the material by themselves in a step-by-step manner. This system allows the student to build upon previous learning and minimizes the chance of making errors or confusion at each step (Slavin, 1991). These steps are called frames.

The content is then organized into a sequence from easy or simple problems to more difficult problems that deal with complex information or skills. Skills need to be taught sequentially, so that new learning by students can be based upon what has been previously taught (Ediger, 2000). Therefore, the subject matter is presented in a sequence of distinct, controlled steps at which the student is tested and expected to pass before moving on to the next level. In order to do that, the learner must master the basic skill or information of a step before moving on to steps that are more difficult.

The setting of performance criteria to determine whether the student has mastered a particular step is often quite controversial. Many test makers have allowed students to progress on to the next step by exhibiting a success rate of 70 percent or less. In programmed instruction, the student must be correct approximately 90 percent of the time when responding to a test items, the student can then make continuous progress with increasingly complex items. Skinner, on the other hand, felt that all the material had to be learned with a 100 percent degree of accuracy before the student could move on to the next level or task.

It is generally agreed that the more time a learner spends interacting with well-organized instructional content the more the student will learn (Sulaiman & Dwyer, 2002). In programmed instruction, a student is given many opportunities to practice a variety of different skills, and to demonstrate mastery of these skills. In fact, the learner must repeat the assignment until the set performance criterion is reached. Thus, the student is given as much time that is needed to master a particular skill. Progress in programmed instruction therefore, is achieved at the students own pace. Students work through the programmed material by themselves at their own speed.

After a student has become proficient with a skill step, the student is given questions to test for understanding and mastery of that skill step. The tests are used at the end of each learning step and are directly related to the specified behavioral objectives of instruction (Kim, 1992). The test items should be based on a list of objectives and then created using a model in which items are selected from template and then varied in difficulty to shape the general character of the test (Kim, 1992).

In order to make progress the student must answer the test questions correctly, either one at a time or in a group where a defined percentage must be attained. Progress to the next level or assignment is only made when the learner has met the set performance criteria. The learner is required to pass each section before continuing on to the next section. This technique encourages practice and mastery of the information or skill.

There are two main types of programmed instruction. The original type was designed by Skinner and was referred to as Linear Programmed Instruction. This type of programmed instruction allows for advancement through the steps only in a particular order when a correct answer is given.

The second type of programmed instruction was called "intrinsic" or "branching" programmed instruction. It was developed in 1958 by Norman Crowder. Programmed instruction is sometimes considered boring for both the student as well as the teacher. Crowder attempted to alleviate this problem by introducing branching to programmed instruction. If a student's answers indicate that he already had some knowledge of the subject, then he may be directed to a branch of more advanced material. With this, Crowder tried to relieve the problem of boring monotony that is encountered with the repetition of all the small steps that are necessary to attain mastery of a certain subject or skill.

In branched programmed instruction, there were many possible answers, and the units or steps of instruction were larger. The learner's possible responses were based on a multiple-choice format. Based on the chosen response the program would branch to the appropriate unit or step (Clark, 1999). The student first had to study a small unit of material. When ready, the student would be asked a question. If the student answered correctly, the student was given new material to study and a new question. If the student answered incorrectly however, the student was directed to a branch of the program with review material

that explains his error. However, branched programmed instruction allowed the students to skip over what they already knew or to go back to review a lesson when they felt it was needed.

In programmed instruction, the role of the learner is an active one. The learner must put effort into the work and respond frequently in order for the reinforcement to be obtained. During the actual learning process, the students should be on task and working hard so that the behaviors, such as writing or calculations in math, can be observed. Educators must understand that learning only takes place unless there is a change of behavior as displayed through higher test scores, or improvement in performance of an activity (Smith, 1999).

The role of the teacher in programmed instruction is to set clear objectives for both the short term and the long term. The teacher selects both immediate and long-term consequences, and the teacher arranges positive reinforcements to occur when the learner improves or progress to the next step. The teacher also has the responsibility to arrange the environment, and to ensure that punishment is not delivered to the learner. If the student is having difficulty with the material, it is the responsibility of the teacher to make each step smaller or

put the information into smaller pieces so that the student can make progress. This is sometimes referred to as feedback.

Feedback is a very important concept in programmed instruction. After answering the question or group of questions, the student should then receive instant feedback. Feedback is usually in the form of telling the learner whether the answered responses were correct or incorrect. The teacher can also determine where the student has made mistakes and then teach what the student needs to meet the desired objective. The first responses of each sequence are prompted with large amounts of feedback, but as performance improves, less and less help should be given.

During the 1960s and 1970s, there were many different programmed instruction techniques developed. However, these techniques in general failed to help students achieve at a higher rate (Slavin, 1991). Other research has shown that programmed instruction was often more successful than traditional instruction because it recognized the different abilities and needs of individual children. Nevertheless, programmed instruction has proven effective in achieving certain learning outcomes. The military and private industry has used it successfully to

train its personnel for a wide variety of activities, skills and information. Programmed instruction materials have also been frequently used in special education to great effect (Slavin, 1991). Finally, programmed instruction that helped change the focus of education from teachers standing in front of a class and presenting information to the learner practicing with the information until mastery is achieved (Smith, 1999).

Programmed instruction has made a big impact on instructional technology and education. More and more, instructional designers have realized that tutorials must do more than present blocks of content with quizzes at the end. "Effective instruction requires learners to respond to what each screen of information presents and to get feedback on their performance before advancing to the next" (Vargas, 2005, para. 11). With the coming of the computer and the internet, the perfect machine that Skinner lacked is now available.

Teaching Machines

Since this work involved the use of a web-based intervention to raise the academic achievement of students, it was important to have a full understanding of the history of teaching machines and the field of

instructional technology, as well as the work that had been done and discovered to be effective in helping students to learn.

There were many definitions used to describe a teaching machine. A teaching machine was a mechanical aid to teaching that presented visual material such as a problem, and then reported immediately to the pupil whether he was right or wrong without requiring any labor on the part of the teacher (Skinner, 1955). A teaching machine was any mechanical, automatic device used for presenting a program of instructional material. A teaching machine was a tool that mechanically, electrically or electronically presented instructional curriculum at a rate controlled by the learners' responses. A teaching machine was an automatic device for implementing the teaching method known as programmed instruction. However, according to Benjamin (1988), the "... consensus definition might read as follows: A teaching machine is an automatic or [self-controlled] device that (a) presents a unit of information (b) provides some means for the learner to respond to the information, and (c) provides feedback about the correctness of the learner's responses" (para. 7).

Teaching machines were devices used to systematically present a programmed sequence of instruction to a student. They presented information and then posed questions to a student. The educational information presented by the teaching machine was called a program. In order to work, a teaching machine required the student to interact with it by selecting an answer to the posed question. After the student had put an answer into the machine, the machine then determined if the answer was correct or incorrect. The student then had to press a button or pull a lever to move each new step or unit into view.

There have been many educational devices patented as educational teaching machines. A simple cardboard device may be called a machine. More complicated machines used film, tape or recorded material to present a program. Many of these teaching machines have been designed to incorporate the ideas of B. F. Skinner. In fact, Skinner has often been credited for originating the idea of the teaching machine. However, as much as Skinner has been associated with teaching machines, he was not the first to attempt to build a machine to teach.

The origin of the teaching machine can be difficult to trace. There are many examples of devices used by the early Greeks that were used to teach. In the first century

A.D., teaching aids made of ivory were used by the Romans to teach the alphabet. By 989 A.D., however, it has been documented that a Celestial Teaching Machine was developed by Gerbert D'Aurillac, who was elected Pope Sylvester II in 999. It was an astronomical teaching machine. It was different in that it did not need a teacher to be at hand to operate (Buck, 2000). In 1809, the first patent of an educational teaching machine was recorded. It was a device used for teaching reading (Benjamin, 1988).

In 1866, Halcyon Skinner was granted a patent by the United States patent office for a machine that helped students practice spelling. Although billed as a teaching machine, it did not teach spelling, but rather only provided practice. A hand crank exposed a series of pictures that appeared at the top of the machine. The pictures such as a dog or a horse were what the student was expected to spell. The front of the machine had eight keys that were each attached to a wheel inside the machine. Each wheel contained the 26 letters of the alphabet and a blank space. However, there was not a system in place to determine whether the word was spelled correctly, and thus the student could misspell a word and never know (Benjamin, 1988).

In 1897, George Altman patented a machine that taught arithmetic. It did not present the learner with information. Therefore, it required a teacher to provide the needed information. However, it was a self-controlled machine that provided a way for the student to respond to questions. It also provided the student with feedback about the correctness of the given answer (Benjamin, 1988).

In 1911, Herbert Austin Aikins patented a device that could help teach a variety of subjects such as mathematics, spelling, reading, history and more. It was made up of wooden blocks that were placed into a wooden case. It was not automatic, self-controlled or a machine. However, the most interesting aspect of Aikins' device was that it was the first educational aid designed with and based on the psychological research of Edward L. Thorndike (Benjamin, 1988).

Further developments can be traced back to Edward Thorndike, an educational psychologist at Columbia University Teachers College. In 1912, Thorndike described the principles of computer-based instruction more than fifty years before the birth of the computer (McNeil, 2004).

In 1928, Michael S. Gleason patented an Educational Game Apparatus. It taught arithmetical, geographical and other facts to children. It also incorporated the idea of game playing, thus making it entertaining and fun so that the student would be engaged for longer periods of time (Gleason, 1928).

In the early 1920s, Sidney L. Pressey, an educational psychology professor of Ohio State University, developed a machine to provide drill and practice items to students in his introductory courses. The device was originally presented at the American Psychological Association (APA) in 1924. The patent application was first submitted in January 1926 under the heading "Machines for Intelligent Tests". The patent was granted in 1928. The teaching machine was initially developed to test students with automatic self-scoring features, but it soon became evident that it had the ability to teach new concepts as well as to test for the understanding of these concepts (Lumsdaine, 1959).

The machine that Pressey made looked like a typewriter with a window that displayed a question and four answers. Ludy Benjamin (1988) describes the machine as:

> A large drum with paper attached rotates and exposes typed or written material in a narrow window. The typed material is essentially a multiple-choice question with four alternatives labeled 1 through 4. On the side of the machine were four corresponding keys that the student pressed to input their answer. The four keys ... correspond[ed] to the four answers, and one of those is depressed by the subject [to input an answer]. (para. 14)

True-false questions could also be answered by using only the first two keys, (Pressey, 1926). In order to progress on to the next question, the user had to press the correct key (Glaser, 1960).

Pressey's teaching machine could function in two main ways. It could operate in either a "test" or "teach" mode. In the test mode, the student simply pressed the key to the corresponding answer. The machine recorded the response on a counter that was in the back of the machine and then automatically advanced displayed the next question (Benjamin, 1988; McNeil, 2004).

When the machine was set to the teaching mode, the user raised a small lever on the back of the machine. In raising the lever, the machine was prevented from moving forward to the next question until the student had

correctly answered the current question. This allowed the student to make multiple attempts at the answer on each question, until the right answer was chosen. All of the key presses were recorded and counted on the mechanical counter in the back of the machine (Benjamin, 1988).

Pressey also implemented the behaviorist idea of reinforcement. An additional attachment could be fitted to the machine. It dropped a small piece of candy into a container, if the student made the right amount of responses that had been set on the "reward dial". With the use of this attachment, the student was automatically rewarded when he/she reached the preset goal (Pressey, 1926).

In the second generation of Pressey's teaching machines, the Drum Tutor would add a few new features. The new machine would add an error window that displayed a cumulative count of the errors (the key presses). When a wrong choice made, the error count increased by one and the question remained in the item window. This indicated to the student to try again with another different response to the question. If the choice was correct, the machine automatically displayed the next question in the item window and the error count remained unchanged. Therefore, the student immediately knew whether his answer

was right or wrong, as well as how many times it took to correctly answer the questions (Stephens, 1953).

This machine was intended for drill and practice and had the unique feature of dropping a question from the testing routine once it had been correctly answered twice in succession. When a student pressed the right key, the drum revolved and turned up a new question. According to Pressey (1927), the machine presented:

> ... the questions in order and [went] through the series the second, third or further number of times. After the series had been gone through twice, the machine revolved past those questions, which had been answered correctly without the pressing of a wrong key. In addition, as an item was learned to the point where two successive right answers are made, it is thus thrown out. Finally, after every item has been mastered, the apparatus automatically stops and releases a small coupon, indicative of the fact that the exercise had been mastered. (p. 43)

The teaching machine could be adjusted by the tester so that the user would have to correctly answer a question two, three, or four consecutive times (Pressey, 1927).

By 1932, Pressey had become discouraged because of the problems of successfully marketing his machine and

disappointment in the lack of interest by education and the public (Benjamin, 1988). Pressey's teaching machine failed to capture the attention of the public. The machines promised a faster educational pace and the need for fewer teachers at the time of the Great Depression when more jobs, such as teachers, were needed.

By 1936, there were almost 700 different patents issued for educational devices. However, with the stock market collapse, the rate of unemployment at an all time high and the eventual entry of America into World War II, teaching machines had all but disappeared from the American consciousness.

Skinner attributed the failure of Pressey's machines "... in part to cultural inertia: the world of education was not ready for them. But they also had limitations which probably contributed to their failure" (Skinner, 1958, p. 1). They were primarily testing devices. Another possible limitation was that it used multiple-choice questions rather than allowing the student to construct their own responses (Lumsdaine, 1959). Skinner also noted that Pressey's machines were designed to be used after some learning had taken place (Skinner, 1958). Pressey's machines did not present information. The student had to study a textbook, or watch a film or lecture prior to

using the device (Lumsdaine, 1959). Skinner applied Pressey's findings to teaching machines for the workplace, the classroom and a variety of other settings. However, he included the idea that the teaching machine should present information in a manner that would be appealing to the student so that it would stimulate the student's interest, and thus would facilitate the learning process to a greater extent (Skinner, 1955).

B. F. Skinner became interested in teaching machines by accident. Vargas (2005) writes:

> When [his youngest daughter] was in fourth grade, on November 11, 1953, Skinner attended her math class for Father's Day. The visit altered his life. As he sat at the back of that typical fourth grade math class, what he saw suddenly hit him with the force of an inspiration. As he put it, "through no fault of her own the teacher was violating almost everything we knew about the learning process." In shaping, you adapt what you ask of an animal to the animal's current performance level. But in the math class, clearly some of the students had no idea of how to solve the problems, while others whipped through the exercise sheet, learning nothing new. In shaping, each best response is immediately reinforced. Skinner

> had researched delay of reinforcement and knew how it hampered performance. But in the math class, the children did not find out if one problem was correct before doing the next. They had to answer a whole page before getting any feedback, and then probably not until the next day. But how could one teacher with 20 or 30 children possibly shape mathematical behavior in each one? Clearly, teachers needed help. That afternoon, Skinner constructed his first teaching machine. (para. 10)

After visiting his daughter's classroom, Skinner built a primitive machine to teach arithmetic. This first teaching device, the Slider Machine, provided students with mathematical drill and practice (Casas, 2002).

To use the Slider Machine, the student had to bring a stack of cards with preprinted problem on them to it. They then inserted a card into the machine, which made a problem appear in a window. The student worked out the problem and then answered the question by moving sliders to set the numbers for their answer. When the student was done and wanted to learn if their answers were correct, they pressed a button. This caused the sliders to lock into place and turn on a light inside the machine. If the student's answer was correct, then the light was revealed

through a hole in the card. This light allowed the student to read the answer. The student then removed the old card. The student could then progress on to the next problem by placing a new card from the stack into the machine. If the student's answer was incorrect, the light did not shine through the hole in the card. The student then had to pull a lever that rearranged the sliders, and the student had to try again (Casas, 2002).

With this first machine, students had to already know how to work out various types of problems (Casas, 2002). Like Pressey's machines, it did not teach anything new. All it did was give more practice on skills already learned. Skinner's first teaching machine simply presented problems in random order for students to do, with feedback after each one.

The Disk Machine was Skinner's second teaching machine. It was small enough to sit on top of a student's desk (Casas, 2002). The machine was a small rectangular box with a lever on the left front and two small windows on top. The center window displayed the question. The other window, located near the right edge, was where the students wrote their answer. A 12-inch paper disk contained 30 printed questions. A student would insert the disk and close the machine. The machine would not work

until the cover was closed and locked. Once a student had begun, the machine could not be unlocked until he/she was done (Skinner, 1958). The student only saw one question at a time. The student answered the question by writing an answer in an exposed frame of a paper tape at the right. The student then raised the lever of the machine (Lumsdaine, 1959). Moving the lever caused the answer to be moved under a glass plate where the student could still see it, but prevented him from cheating and changing it (Casas, 2002). It also turned the paper disk to show the correct answer. The student now had to compare his answer with the correct answer, and decide if it was right or wrong. If the answer was correct, the lever was moved to the right. This movement punched a hole in the paper next to the response, recording the fact that it was determined to be correct. This also set the machine so that the question did not appear again when the student worked around the disk a second time. Whether the response was correct or not, a second frame appeared when the lever was returned to its starting position. The student then worked through the disk's problems a second time, but only with the questions that were not answered correctly. When the disk rotated all the way around without stopping, the assignment was complete (Skinner, 1958).

Skinner built a learning device that taught students new information in small easy steps whereby errors could be kept to a minimum. The information therefore had to be carefully sequenced and organized so that new learning could be built upon what was learned in previous steps. Skinner was not satisfied and soon would develop programmed instruction (Vargas, 2005). The term programmed learning was coined to describe information constructed in a systematic and logical manner (Skinner, 1958).

Another important difference in Skinner's machines from those of Pressey and others was the type of questions used. Pressey's machine asked the student to answer multiple-choice questions. Skinner objected to the multiple-choice question format because it meant exposing the student to a variety of wrong answers. In addition, the student did not have to create an answer, but rather just guess and select an answer. Skinner felt that students needed to learn how to construct their responses. Therefore, Skinner's second machine required the student to create his or her own answer (Benjamin, 1988).

By the mid 1950's, there were many others developing various types of teaching machines. Some were sponsored by the military; others were developed by companies looking to make a profit; and some were just the ideas of teachers

trying to help their students learn. Douglas Porter of Harvard University created a device that featured only five questions within a cycle. The questions were on dittoed sheets that were fed into the bottom of the machine. The students wrote their answers directly onto this sheet (Lumsdaine, 1959).

E.Z. Rothkopf developed a teaching device he named the Polymath (Rothkopf, 1958). He built a machine that incorporated a plastic electric tracer that the student used to draw a picture, diagram an electrical circuit or draw a route on a map. The machine could then automatically determine the correctness of the response of the student (Lumsdaine, 1959).

Another teaching machine was called the Magazine-loaded automatic projector. Designed by the Air Force, it presented films that demonstrated skills in the laboratory or training for technicians. It was pre-programmable and had start and stop features. The magazine contained the pre-threaded film (much like a VCR tape) and could be load into the machine by the user instantly (Lumsdaine, 1959).

R.M. Gagne created a microfilm projection machine for the Air Force. It was designed to teach troubleshooting for complex electronic equipment. The teaching machine

could select a question from several hundreds that were available on the microfilm. The machine could also put the questions in any random sequence or order. The order would be determined by the responses of the student (Lumsdaine, 1959).

The Subject-Matter Trainer was developed by Leslie J. Briggs. It presented a series of questions in sequence in a small window and the student had to press a button next to 20 different answers. When the student chose the right answer, a green light turned on. A buzzer would sound if a wrong answer was chosen. It was used to train personnel to identify components, terms and other paired relations (Lumsdaine, 1959; Briggs, 1958).

Another teaching machine developed by Leslie J. Briggs was the Card-Sort Device. It provided questions on individual cards that when answered were sorted into two piles depending on whether the question was answered right or wrong. The student worked through the wrong piled questions until there were no cards left. A red light was lit for an incorrect answer, while a green light was lit for a correct answer (Briggs, 1958).

By the late 1950s, teaching machines were in vogue. The launch of the Russian spacecraft Sputnik had fueled new interest in the sciences and the need to educate

students in new and better ways. Interest in teaching machines peaked within the educational community, the armed services and private industry. Title VII of the National Defense Education Act (NDEA) of 1958 promoted the development of technology projects, including the design and implementation of teaching machines and programmed instruction (Casas, 2003).

In the early 1960s, teaching machines were being used throughout the United States at universities and public schools. Grants from the U.S. Office of Education further fueled the use of teaching machines in U.S. schools. By 1962, there were fifty-nine companies building various types of teaching machines. Major companies such as AT&T, General Dynamics and Kodak were beginning to use teaching machines and programmed instruction for training their employees (Casas, 2002). Teaching machines had become big business and most of the machines used the basic Behaviorist ideas from Skinner's work (Benjamin, 1988).

By the late 1960s and early 1970s, however, teaching machines had been discarded due to various concerns. Initially, cost was a major problem negating use of computers in the classroom. School districts did not find teaching machines cost effective. Teaching machines were not used as they were intended. They were often only used

as an intervention for low performing students, enrichment or supplemental purposes (Casas, 2002). Early machines were limited in the types of programming they had to offer, as well as the reinforcements they could provide. Schools preferred to design their own programs rather than purchase commercial programs. Programs were not "user friendly". Some teachers were afraid of or did not understand the new technology. They felt it isolated the students from teachers and peers. Many teachers were forced to use the teaching machines and developed resentment toward them. Finally, people who had supported teaching machines were gradually becoming aware of a new and rapidly developing technology called the computer and the concept of computer-assisted instruction (CAI) (Benjamin, 1988).

Computer-Assisted Instruction

Since this work involved the use of a web-based intervention to raise the academic achievement of students, it was important to have a full understanding of the history of Computer-Assisted Instruction and the field of instructional technology, as well as the work that had been done and discovered to be effective in helping students to learn.

Computer-assisted instruction (CAI) refers to any instruction presented by a computer. The computer uses pictures, diagrams, audio and video to present content and subject matter. Computer-assisted instruction can present drill and practice exercises, teach through tutorials, test a student for comprehension and engage the student in a dialogue about a subject. The term can also apply to systems that provide interactive on-line instruction and testing (Suppes, 1988).

Computer-assisted instruction goes by many other names. It is often referred to as computer-aided instruction and computer-assisted learning. Other terms that apply to computer-assisted instruction are computer-based instruction (CBI), computer-based education (CBE), computer-based learning (CBL), computer-based teaching (CBT), computer-enriched instruction and computer-managed instruction (CMI).

One of the primary goals of computer-assisted instruction is to make the learning process more effective and efficient (Atkinson, 1969). The idea of computer-assisted instruction was to assess the students for initial understanding of a concept. Then the computer could present information, give practice to the students, test for comprehension of the subject and provide extra

instruction to students who needed (Slavin, 1991). In addition, computer-assisted instruction has been used to help the student master the educational objectives and standards that have become so prevalent in education (Traynor, 2003).

Computer-assisted instruction began as an attempt to copy teaching machines. CAI was an outgrowth of the Teaching Machine Project at the IBM Research Center in the late 1950's. During that time, the researchers developed the IBM 650 Inquiry Station. It was comprised of a large, mainframe computer linked to a typewriter. It was designed to teach arithmetic. Around the same time, IBM also developed the first computer language devoted exclusively to computer-assisted instruction. It was named Coursewriter.

The first successful computer assisted instruction was called PLATO (Programmed Logic for Automatic Teaching Operations). It was built at the University of Illinois with the assistance of Control Data Corporation (Benjamin, 1988). Debuting in 1960, it is considered the beginning of computer-assisted instruction. The system was composed of a computer, a slide selector, a keyboard, a storage device and a television screen. It was designed to teach thirty different courses to elementary school, high school and

college students. In one of these courses, a film was shown and then a series of 32 questions were given on the films content (Troutner, 1991).

At Stanford University in 1963, researchers set out to design and develop the first integrated CAI system. In order to construct the system, they had to assemble various components from a variety of different companies. They used a PDP-1 computer from the Digital Equipment Corporation. Audio was provided by a Westinghouse device. For a monitor, a cathode-ray display from Philco-Ford was used, and a device for film imaging was supplied by IBM (Atkinson, 1969). Six student stations were eventually developed and taught mathematics and language arts.

The IBM 1500 Instructional System was developed by IBM and Patrick Suppes at Stanford University in 1966. It is considered the first computer designed with education as its primary goal. It used a very sophisticated branching logic for the time that made decisions about what was to be presented next by evaluating the student's responses (Atkinson, 1969). The IBM 1500 Instructional System had an audio system, a projector to display pictures, a cathode ray tube (CRT) graphic display with a light pen, and a keyboard. Computer practice came in the forms of drill-and-practice, tutorial presentations and

dialogue interactions (Troutner, 1991). The instructional software used to power the system was an early version of IBM's Coursewriter (Reisman & Carr, 1991). It was designed to teach mathematics and reading to first grade students. The reading program taught students to identify letters of the alphabet and vocabulary. To answer questions, the students touched the screen with the light pen. The computer first determined where the student was pointing and whether or not the student's responses were correct. It then determined what information should be presented next to the student (Slavin, 1991).

The program proved to be successful. Students who received only twelve minutes of computer instruction a day for one year improved their reading scores by 1.2 grade level equivalents more than the control group who did not use the computer supplemental program.

The Stanford Drill-and-Practice System was developed at the same time as the IBM 1500. However, it used a different design philosophy. It was comprised of a large mainframe computer and Model 33 teletype student terminals connected by telephone lines. Approximately 500 student terminals were spread around the United States and the University of Stanford. It also did not have the

sophisticated branching logic of the IBM 1500 (Atkinson, 1969).

The 1970s showed continually growth in the field of computer-assisted instruction. In an index of computer-assisted instruction there were 1264 specific programs listed. That number more than doubles to 2997 available programs by 1978 (Suppes, 1979).

With support from the National Science Foundation, PLATO continued to be used throughout the 1970s both in Mathematics and in Language Arts. From 1973-1976, under the direction of Robert B. Davis, the work developed more than 100 hours of instructional material for teaching mathematics in Grades 4-6. Three strands of instructional material were developed: whole number arithmetic; fractions, mixed numbers, and decimals; and graphs, variables, functions and equations. The program was designed to include a wide range of student abilities, and each half hour selection was divided into three sections: a review of the previous material, a lesson to cover new material and a fun game. The program relied heavily on the graphic abilities of the PLATO terminals (Suppes, 1979).

From 1971 to 1976, PLATO also developed an Elementary Reading Curriculum Project. A series of behavioral objectives involved in learning to read was first

developed. In support of these objectives, approximately 80 hours of instructional material was developed. A computer-based curriculum management system, sequences of audiovisuals, and teacher feedback systems were then developed. Though the audio was an unreliable feature, the program successfully covered all the standard material and lessons as related to teaching the initial steps in reading (Suppes, 1979).

The largest CAI model was offered by Computer Curriculum Corporation (CCC). It offered a variety of courses for elementary through college students. It consisted of a mainframe computer and up to 96 teletype terminals linked by telephone connections could be used simultaneously (Suppes, 1979; Hayes, 1999). CCC developed drill-and-practice lessons that were meant to supplement the regular classroom instruction in reading and mathematics.

An interesting component of this model was that the computer kept records of everything the student did. The computer would then compare the student's work with the preset performance criteria. The computer could then move the student back to an easier level, continue to have the student practice at the present level or to advance the student on to the next level. The computer program also

used a random selection method to determine what the student would be working on within a given strand. Thereby, allowing the student to receive a mixture of content and exercises. In order to provide progress reports and grades, this model had the ability to furnish teacher reports showing the grade placement for each of the strands or content areas that the student was currently working on (Suppes, 1979).

The CCC program for mathematics was quite sophisticated for the time. It was designed around 14 mathematical strands that ranged from first grade through grade level 7.9. If a student was working at grade level 3.5, then the computer would randomly chose problems from all the strands with grade level appropriate material. The most fascinating aspect of the program was that the material was not pre-stored on the computer, but rather used an algorithm that randomly generated problems. Therefore, even if a student had to repeat a particular lesson several times, the problems would be different every time (Suppes, 1979).

Another major computer-assisted instructional model was called Time-Shared, Interactive, Computer Controlled, Information Television (TICCIT). TICCIT was developed by the Mitre Corporation and C. Victor Bunderson at Brigham

Young University. It received major support from the National Science Foundation from 1971-1976. It combined mini-computers and color televisions to deliver computer-assisted courses in English and mathematics to college students (Hayes, 1999). The programs were developed by experts rather than teachers and focused on learner-controlled experiences. Though, TICCIT did not provide enough remediation to help all students, it did prove effective when used as an adjunct to the classroom (Suppes, 1979).

During the 1970's, computer-assisted instruction was used in many other forms and at several different learning institutions. One of the best known was a mastery-based physics class implemented at the University of California, Irvine by Alfred Bork, Stephen Franklin and Joseph Marasco. CAI was also used at Ohio State University. It provided a drill and practice program to supplement the regular course lectures, and conducted some testing. In addition, Stanford University used CAI to teach an entire logic course (Suppes, 1979).

By the late 1970's, rapid developments in microprocessors and other components led to the availability of fully assembled microcomputers for the general public. By 1977, there was a shift to much smaller

computers when the Apple computer was developed. The Apple IIe soon followed in the early 1980's. With these developments, the personal computer was born and the use of computers for instructional purposes could be done in the comfort of a student's home.

By the 1980s, computers had become commonplace in the classroom and in the home. Continued improvement in the capabilities of hardware and software, as well as reduced costs for both, have made the computer accessible to nearly everyone (Benjamin, 1988). The use of online learning and hyperlinked materials also became more common. By 1984, distance learning became a major development in CAI with the introduction of CYCLOPS at the Open University UK.

By the 1990's, most schools had at least one computer. The decreasing cost and increasing availability of computers had led schools and teachers to be more interested in CAI (Slavin, 1991). In addition, computer-assisted instruction had blended with the world of E-learning. Computer-assisted instructional activities were now conducted via the internet and the World Wide Web, thereby connecting remote communities to much larger schools and universities.

In 1992, Patrick Suppes developed the Education Program for Gifted Youth (EPGY) at Stanford University. Computer-assisted instruction now was delivered with multimedia technology. Suppes combined text, audio and graphics to present lectures and lessons that were not available for various reasons. Unlike most of the previous CAI programs, this program was not designed to work as a supplement to a traditional classroom instruction, but rather was meant to stand-alone. EPGY was designed so that gifted students as young as 15 could take the class at home. The software presented demanding problems and then quizzed the students. Mastery of the subject was required to move on to new material. If a student needed help however, he had to contact an instructor through traditional means such as a telephone or through email (Rosenthal & Suppes, 2002).

By the turn of the century, the field of computer-assisted instruction was characterized by a tremendous range of developments that included web-based distance learning, discussion boards, e-mail, blogs, text chat, the World Wide Web, web sites and hypermedia. Hypermedia programs allow the user to integrate sound, animation, graphics and text through a variety of paths into one document. Hypermedia was designed to allow the

student more control over one's own learning (Soe, Koki, & Chang, 2000). By 2002, computer-assisted instructional assessment techniques had also developed. The computer could now record the exact time it took for students to make their response to a question as well as the sequences of the responses. This has enabled researchers to see a new aspect of the student's performance, as well as better determine the effectiveness of the program (Cope & Suppes, 2002).

There has been a great deal of research into the effectiveness of computer-assisted instruction. Researchers generally agree that computer-assisted learning is highly effective, but not all research concurs. For example, regular education students were generally found to make greater gains in test scores than special education students (Traynor, 2003). However, research conducted with learning disabled students, mentally retarded students, hearing impaired students, emotionally disturbed students, and language disorder students indicate that achievement levels in these students are far greater with computer-assisted instruction (Cotton, 1991).

Research has repeatedly found that computer-assisted instruction is most effective when used in combination

with regular classroom instruction (Slavin, 1991). Research indicates that students who learned subject matter using a computer in conjunction with traditional instruction retain the information far longer than students using only traditional forms of instruction (Cotton, 1991).

Researchers have also found that computer-assisted instruction enhances learning rate. Students actually learn faster with computers than with traditional instruction. According to Cotton (1991), students using computers "... learn as much as 40 percent faster than those receiving traditional, teacher-directed instruction" (Learning Rate, para. 1). In addition, students learned more material in the same amount of time that was given to students using conventional instruction (Cotton, 1991).

Computer-assisted instruction also has a positive impact on reading achievement (Soe, Koki, & Chang, 2000). Computer-assisted instruction holds great promise to increase student engagement in reading and to teach reading and reading comprehension skills. Computer-assisted instruction students have demonstrated higher rates of time on task than students who receive traditional instruction (Cotton, 1991). "For most students, the computer seems to have a motivating quality

all its own, so that they work longer and harder when using it than they would on comparable pencil-paper tasks" (Slavin, 1991, p. 306). Computer-assisted instruction captures the student's attention because of the interactive quality of the programs and with the use of graphics, animation, video and sound. Students using computer-assisted instruction also tended to have better attendance in school (Cotton, 1991).

Computer-assisted instruction seems to be more effective at teaching lower cognitive objectives than with higher cognitive objectives. Therefore, computer-assisted instruction seems to benefit younger elementary students more than older students in high school or college. It also seems to be more effective with lower achieving students than with higher achieving students. Economically disadvantaged students also receive great benefit from computer-assisted instruction. This could be due to the extensive drill and practice, privacy, immediate feedback and reinforcement features that are common to most CAI programs (Cotton, 1991).

Second language learners appear to be the only group not to greatly benefit from the use and interaction with computer-assisted instruction activities (Cotton, 1991).

Though research seems to indicate the benefits of computer-assisted instruction, there are still several problems in implementing it into today's public K-12 educational system. The first problem is the financial cost of fully implementing CAI into every classroom. Yet researchers continually point out the fact that the use of CAI is much less expensive than the variety of tutors and instructional interventions that are currently being used in schools today.

A second major problem concerns change. Schools and school districts are resistant to change (Suppes & Fortune, 1985). Never is the term status quo more appropriate than when used in describing the United States educational systems. Just as in Skinner's day when he attributed the failure of Pressey's machines in part to "cultural inertia" (Skinner, 1958), today we are still affected by the school district's and teacher's lack of desire to change. Some in education are still unsure about computer usage, while others are computer illiterate. Finally, teachers teach based on how they were taught. Therefore, teachers cling to old, outdated modes of teaching. One outdated method of teaching is the idea of a teacher standing in front of a class giving a lecture and then answering questions. In using computer-assisted

instruction, teachers would need to move away from being a presenter of information and become more of a troubleshooter and facilitator for the students (Suppes & Fortune, 1985).

In CAI, the teaching machine has been replaced by the computer. Computers can deliver information with a variety of different media at an extremely rapid rate. Now, the challenge in CAI is to develop new, more effective and efficient programs that will not only deliver the instruction, but also will encourage children to interact with it for longer periods of time. If CAI is to continue to grow and become more successful in helping students to learn, researchers and educators will have to discover which reinforcements are the most effective in gaining a student's attention and creating a desire to learn.

Reading and Reading Comprehension

In order to raise the academic achievement of the student both on standardized tests and in reading comprehension, it was important to have a full understanding of what research has been done in the fields of Reading and Reading Comprehension. It was also important to know what has been discovered to be effective in helping students to learn in these fields.

Reading and reading comprehension skills are the most important academic objectives in America's public schools today. The National Reading Panel (2000b) states, "Comprehension is critically important to [the] development of children's reading skills and therefore to their ability to obtain an education" (p. 4-1).

The importance of reading and reading comprehension cannot be overstated. Reading exposes people to the accumulated knowledge and history of human civilization. In addition to its basic and fundamental value, the ability to read has financial consequences to the person as well as to society in general. "Furthermore, American schoolchildren without high levels of reading comprehension face a difficult and uncertain economic future" (Johnson & Howard, 2003, p. 87). On the other hand, adults who read well tend to earn more money and are more likely to have higher-paying jobs. As society continues to grow it has become more technological, scientific and information driven. Most of what adults read contains information that needs to be understood (Duke, 2004). We now live in the information age, in which information is power. In order to achieve this, people must be able to read and comprehend. In fact, in all areas of everyday life, a higher level of literacy is needed.

The reading ability needed to comprehend materials important to basic daily living, such as voting forms, income tax forms, and driver's license tests has greatly increased with every passing year. Even to read the daily newspaper, a person needs to read and comprehend at approximately an eighth grade level. Though simplified reading materials have begun to be developed, the lack of sufficient reading ability definitely impairs a person's capacity to function in modern Western society.

This has brought increasing demands for literacy, and society has come to realize the need for better reading from its students. It has begun to demand that public education produce better results. Large urban schools in general, however, have not met this need. They have produced and continue to produce students who do not possess basic reading skills. "More than eight million students in grades 4 to 12 are identified as struggling readers" (Sternberg, Kaplan, & Borck, 2007, para. 1). "Slavin, Karweit, Wasik, Madden, and Dolan (1994) note that students who complete the third grade and lack reading skills are not likely to graduate from high school" (Johnson & Howard, 2003, p. 87), let alone contribute in a positive way to our future society.

Reading and reading comprehension are interrelated skills. In order for students to be able to comprehend what they are reading, they naturally have to be able to read. Reading is defined as an activity characterized by the translation of symbols, or letters, into words and sentences that have meaning to the individual. Reading comprehension, however, is much more than simply reading the words. The concept of reading comprehension is vast in breadth and depth, and it requires many different reading skills to be in place. Effective reading comprehension includes phonemic awareness, phonics, fluency, vocabulary as well as reading comprehension skills.

Reading Stages

Literacy is a process in which a student constantly develops and grows. To reach maturity in reading, an individual goes through a series of stages. The reading and reading comprehension stages are characterized by a series of sequential steps that begin with Reading Readiness and progress on to Beginning Reading. The Development of Reading Skills comes next and culminates with adult reading ability.

The first stage in reading is called Reading Readiness. The readiness stage is where the development of

reading and reading comprehension starts. This pre-reading period extends from the time a child is born to when a child first enters school and begins to be taught to read, usually around age six. During this period, the child learns to speak, to follow directions, to follow along with a story, to look at and understand pictures, to gain the ability to perceive differences in the sounds of words, and to become interested in stories and books. It is at this stage that young children begin to acquire an understanding of the language. They learn that spoken words are made up of separate sounds and that letters can stand for these sounds. In addition, they learn letter names and the sounds of the letters. It is during this time when young children become phonemically aware.

The next major stage in reading is known as the Beginning Reading stage or Word Recognition stage. This stage usually occurs between the first and the third grade. By the first grade, children begin to learn to read letters and to associate them with spoken words they know. They are taught to recognize basic site words and phonics. Fluency and comprehension of the text also begin to become important at this stage. The goal is for students to be reading beginning materials independently by the middle of the first grade, and that every student should be reading

independently and comprehending fully no later than the end of the third grade (California Department of Education, 2006).

Development of Skills is considered the next major stage in reading. This stage usually occurs between the fourth and the eighth grade with students ranging in age from 9-14. By the fourth grade or mid-elementary school years, the emphasis shifts from learning to read to reading to learn. Children are no longer taught to decode, but rather are expected to reading independently, and reading comprehension becomes emphasized. It is at this stage where comprehension can be taught as a series of skills. These skills include understanding word meanings in context, finding the main idea, making inferences about information implied but not stated, distinguishing between fact and opinion, predicting what might happen next in a story, and answering questions about various parts of the story such as plot, characters or setting. There is also a shift from reading stories to reading more difficult expository, content area materials such as Science and Social Studies. It is at this stage that the idea of teaching children new knowledge begins to become important. Students begin to read about facts, ideas and concepts from one general point of view.

In high school and college, students must understand facts and concepts from multiple points of view. This stage occurs after the eighth grade with students ranging in age from 14 to adult. The reading materials become more abstract and contain a larger, more technical vocabulary. At this stage, the student not only must analyze new information, but must also differentiate between different points of view and eliminate previously learned concepts and ideas that with new learning have been proven to be incorrect or misunderstood.

Improving Reading Skills

For over a hundred years, research has been conducted to determine the important aspects and skills that are necessary for reading. Traditionally, educators could not agree on the exact nature of the skills that were involved in reading and reading comprehension. Some researchers felt that reading was one entire skill that could not be separated into various parts. Other researchers felt that reading could be divided into different parts and taught separately. According to Armbruster, Lehr, and Osborn (2003), "While there are no easy answers or quick solutions for optimizing reading achievement, an extensive knowledge base now exists to show us the skills children

must learn in order to read well" (p. ii). These skills include phonemic awareness, phonics, fluency, vocabulary and reading comprehension skills.

Phonemic Awareness

Phonemic awareness is a required facet of reading that has often been overlooked. For a student to be able to learn to read it is critical that children relate sounds to letters.

Phonemic awareness is the ability to identify and work with the individual sounds in spoken words. It is the ability to manipulate the sounds of a spoken language. It is the understanding that the sounds of spoken language work together to make words. These individual sounds are called phonemes. They are the smallest parts of sound in a spoken word. There are 40 or 41 phonemes in the English language.

Phonemic awareness is displayed by children in several ways.

- Recognize which words in a set of words begin with the same sound ("Bell, bike and boy all have /b/ at the beginning.")

- Isolating and saying the first or last sound in a word ("The beginning sound of dog is /d/." "The ending sound of sit is /t/.")
- Combining, or blending the separate sounds in a word to say the word ("/m/, /a/, /p/ - map.")
- Breaking, or segmenting a word into its separate sounds ("up - /u/, /p/.") (Armbruster, Lehr, & Osborn, 2003, p. 2)

Phonemic awareness helps children to learn to read in many ways. It improves their ability to read words. It helps with reading comprehension. It also helps children to learn how to spell better (Armbruster, Lehr, & Osborn, 2003). However, for disabled readers, phonemic awareness instruction did not significantly improve spelling (National Reading Panel, 2000b).

Wile phonemic awareness can be learned, it is more important to understand that it can be taught. According to the National Reading Panel (2000a), scientific evidence shows "... that teaching children to manipulate phonemes in words was highly effective under a variety of teaching conditions with a variety of learners across a range of grade and age levels ..." (p. 7). The teaching of phonemic awareness to children significantly improves their reading

more than instruction that lacks any attention to phonemic awareness.

Phonemic awareness is usually taught to preschoolers, kindergartners and first graders. Small group instruction is considered the best way to teach young students and give them a solid foundation in phonemic awareness. Effective instruction includes teaching children to recognize the individual sounds within a word, recognize the same sounds in different words, recognize a sound that does not belong in a set of words, sequence spoken sounds and combine them to make a word, break words into separate sounds and then say each sound, recognize a word when one of it's sounds is removed, make new words by adding new sounds to a word, and substitute one sound for another to create a new word (Armbruster, Lehr, & Osborn, 2003).

Phonemic awareness instruction can be more effective when taught correctly. The best way is to teach children to relate the phonemic sounds to letters of the alphabet. Phonemic awareness instruction is also more effective when a teacher focuses on only one or two strategies rather than all of them. This allows children to thoroughly gain an understanding of a particular strategy without getting confused (Armbruster, Lehr, & Osborn, 2003).

To improve phonemic awareness, parents and teachers should include many different activities that focus on blending and segmenting words. Rhyming words is one of the first and best activities for young students to do. Children should also have many opportunities to identify and categorize word sounds. This should start with the first sound of a word and then move on to more difficult ones. It is important to remember, that children should have a solid understanding of phonemic awareness before moving on to phonics instruction (Armbruster, Lehr, & Osborn, 2003).

Phonics

The National Reading Panel (2000a) defines phonics instruction as:

> ... a way of teaching reading that stresses the acquisition of letter-sound correspondences and their use in reading and spelling. The primary focus of phonics instruction is to help beginning readers understand how letters are linked to sounds (phonemes) to form letter-sound correspondences and spelling patterns and to help them learn how to apply this knowledge in their reading. (p. 8)

In phonics, there is a predictable relationship between the letters in a written language and the sounds of a spoken language (Armbruster, Lehr, & Osborn, 2003). Simply stated, phonics teaches students the sounds that are associated with letters. This allows students to decode the written words and read.

Phonics or phonics instructions goes by many different names. Graphophonemic relationship, letter-sound association, letter-sound correspondence, sound-symbol correspondence, and sound-spelling are all names used to describe phonics.

Regardless of the label, the objective of phonics instruction is to enable children to learn and remember a system for reading words. A good phonics system is explicit, sequential and systematic. Used in this way, phonics instruction provides an effective way to produce growth in a student's reading, and has been proven more effective than phonics instruction that is non-systematic or no phonics instruction at all (Armbruster, Lehr, & Osborn, 2003).

Phonic instruction should be done as a whole class or as a small group activity for at least two years in the primary grades of Kindergarten, first and second grade. As the National Reading Panel (2000a) states, "The effects of

systematic early phonics instruction were significant and substantial in kindergarten and the 1st grade, indicating that systematic phonics programs should be implemented at those age and grade levels" (p. 10). Therefore, as Armbruster, Lehr, and Osborn (2003) declare:

> Systematic phonics instruction produces the greatest impact on children's reading achievement when it begins in Kindergarten or first grade. Both Kindergarten and first-grade children who receive systematic phonics instruction are better at reading and spelling words than kindergarten and first grader children who do not receive systematic instruction. (p. 14)

Phonics instruction is taught through the direct teaching of a set of letter-sound relationships in a clearly defined sequence. It includes both consonants and vowels. In phonics instruction, children are given frequent and numerous opportunities to work with and practice these relationships. The books and stories that the students work with provide opportunities to focus on specific decoding lessons that they have learned. In addition, children are given opportunities to write their own stories and spell words that use the decoding strategies that they have been taught.

Phonics instruction helps children to learn to read in several ways. Research shows that systematic phonics instruction has proven effective for students from Kindergarten through the 6th grade. It has also shown to be helpful for children with reading problems (National Reading Panel, 2000b). Phonics instruction improves the student's ability to recognize words, and it helps children to learn to spell better (Armbruster, Lehr, & Osborn, 2003). Kindergartners learned to read and spell better after receiving beginning phonics instruction. First graders also learned to read and spell better, but they also showed considerable improvement in their ability to comprehend text. Older children, while showing growth in their ability to read, did not show significant growth in reading comprehension (National Reading Panel, 2000a).

Fluency

Fluency is the ability to be able to read quickly and accurately. It is also the ability to recognize words instantly and to read smoothly with expression.

Fluency in reading is determined by two main factors. The first factor is called automaticity. Armbruster, Lehr, and Osborn (2003) state, "Automaticity is the fast, effortless word recognition that comes with a great deal

of reading practice" (p. 24). It is the ability to read a text rapidly and correctly. Teachers, therefore, must practice reading words with their students until they can read them automatically.

Expression is the second key factor in reading fluency. Expression is the act of putting feeling into what is read. Expression in reading is called prosody. To read with expression, a child must use appropriate phrasing, be able to divide the text into meaningful chunks, be able to pause at the correct time, and use stress patterns. The reader must also know when to pause appropriately within and at the ends of sentences, and when to change emphasis and tone (Morra & Tracey, 2006).

Studies have shown a strong correlation between fluency and comprehension. In 1997, Tan and Nicholson's research discovered that students who read fluently did better with reading comprehension (Pressley, 2001). Readers, who spend a lot of time and energy on the pronunciation of unknown words, do not have the time and energy to focus on the meaning of the text they are reading. This results in poor reading comprehension (Rasinski, 2003). Therefore, fluency in reading is the link between word recognition and reading comprehension. Fluency in reading allows the student to concentrate on

the message of the text rather than on the decoding of the words (Armbruster, Lehr, & Osborn, 2003).

There is common agreement that practice is essential to developing reading fluency. In fact, for most students substantial practice is required to attain fluency in reading (Armbruster, Lehr, & Osborn, 2003). Snow, Burns, and Griffin (1998) concluded that daily reading of many different texts that are at an appropriate level could possibly prevent children from becoming poor readers. Therefore, for the development of fluency, students need to have many opportunities to practice reading and have a high degree of success while reading. In order to have a high degree of success while reading, a student needs to read material that is at their reading level. Reading that is too easy does not improve reading skills, while practice that is too hard discourages the reader and leads to negative attitudes and giving up (Allington, 2001). According to Vygotsky (1978), the area between too hard and too easy is called the zone of proximal development. It is at this difficulty level where the students learn to read with maximum efficiency.

Traditionally, there have been two major approaches to teaching fluency in reading. The first approach is called repeated reading. Teachers have their students read

the same text aloud several times until the student's speed and accuracy are acceptable and only give guidance when it is needed. For most students, it is necessary to read a passage at least four times. (Armbruster, Lehr, & Osborn, 2003; Morra & Tracey, 2006). Research has shown that this approach has had a significant and positive effect on word recognition, speed, accuracy, fluency, and comprehension for a variety of elementary grade levels (National Reading Panel, 2000a).

The second major approach to teaching fluency in reading is called independent silent reading. Traditionally, programs such as Sustained Silent Reading (SSR) and Drop Everything and Read (DEAR) have been recommended. At this time, research has not yet determined if independent silent reading improves fluency (Armbruster, Lehr, & Osborn, 2003). In fact, independent silent reading may have a negative effect on readers who have not achieved fluency by allowing the student either to incorrectly read words or to not practice reading during the silent reading time.

Vocabulary

There is no doubt that vocabulary is very important to reading comprehension. Since 1924, research has shown

the importance of vocabulary in developing reading skills and knowledge (National Reading Panel, 2000a). In addition, students who have large vocabularies do well in reading comprehension (Pressley, 2001). On the other hand, students who cannot understand what they are reading often are unfamiliar with the vocabulary used in the text.

Vocabulary refers to the words that readers need to know and understand to comprehend what they are reading. These words are learned or acquired in two ways. Students learn most new words either directly or indirectly.

The research on vocabulary instruction has determined that "... most vocabulary is learned indirectly ..." (Armbruster, Lehr, & Osborn, 2003, p. 35). When students learn vocabulary indirectly, they learn new words from their everyday experiences such as talking with adults and other children. Children also learn vocabulary by reading on their own. In fact, the more students read, the more words they are exposed to the more words and word meanings they learn.

Scientific research shows that some vocabulary can be taught directly (Armbruster, Lehr, & Osborn, 2003). This is called direct or explicit vocabulary instruction. In direct vocabulary instruction, students are taught definitions to new words as well as how to pronounce the

new words. Direct instruction of vocabulary seems to be most effective when students understand that there is a meaningful purpose involved such as when the words are relevant to what the student is already learning. Thus, students are usually taught words that they have not come into contact with, but will soon see in the text that they are about to read. Instruction is also very effective when it is done over an extended period of time and with repeated coverage of the vocabulary words. Students should have many opportunities to interact with the words and various activities in which they are required to work actively with the words.

Computers are beginning to be used more often for vocabulary instruction. According the National Reading Panel (2000a), "The use of computers in vocabulary instruction was found to be more effective than some traditional methods in a few studies" (p. 14). Computers seem to help all students in vocabulary instruction. However, students in Preschool did not seem to benefit from computers as much as older students (National Reading Panel, 2000b).

Vocabulary is extremely important to the overall success of a student. Through vocabulary instruction, a student's reading comprehension can improve. In 1982,

Isabel Beck worked with fourth graders and taught them over 100 new vocabulary words. Her students outperformed other students who did not receive direct instruction in vocabulary on reading comprehension tests (Pressley, 2001).

Comprehension

Comprehension is the reason and purpose for reading. "Comprehension of text is now regarded as essential to reading ..." (National Reading Panel, 2000b, p. 4-39).

Reading comprehension is defined in many ways. As far back as 1917, E.L. Thorndike claimed, "... reading is reasoning" (National Reading Panel, 2000b, p. 4-40). Reading comprehension is also described as the level of understanding when reading a passage or text. It can be defined as the act of or capacity for understanding with the mind. Reading comprehension is the measurement of the reader's ability to make sense of written texts. Reading comprehension is the making of cognitive connections between the symbols we see and the words we say. In reading comprehension, the reader interacts actively with the text as he/she tries to determine the meaning of various kinds of text (Alyousef, 2005). In reading comprehension, not only does the reader need to understand

what has been read, but he must also make connections to what he already knows. The ultimate goal of reading is to be able to understand written material, to evaluate it, and to use it for one's needs. The simplest, most basic definition for reading comprehension is "understanding what you read".

Whatever the definition may be, it requires a great deal of practice to have good reading comprehension skills (Johnson & Howard, 2003). Reading comprehension is a learned skilled, and it can be improved through teaching and practice of reading comprehension strategies. Reading comprehension instruction generally begins at the third grade with most explicit reading comprehension instruction ending by the sixth grade (National Reading Panel, 2000b).

In order to maximize reading comprehension, students must first have a strong command of all the basic reading skills. Students need to be able to decode and read words well. They must also be able to read fluently and have a strong vocabulary in place (Pressley, 2001). It is essential that word-recognition and fluency skills be mastered to develop extensive reading comprehension skills.

In the late 1970s, Dolores Durkin found that there was very little comprehension instruction in the

classroom. She determined that, out of over 4,400 minutes of observed instruction, there was only 11 to 20 minutes of comprehension instruction. What she did find was that students were reading and then answered comprehension questions. In fact, Michael Pressley discovered that this trend has continued through the 1990's (Pressley, 2001; National Reading Panel, 2000b). There is a real need to teach children reading comprehension strategies instead of just having students answer comprehension questions. However, teachers are reluctant and unwilling to teach reading comprehension strategies (Pressley, 2001).

After the Durkin study, there was a great effort to learn what was important about reading comprehension. This period, from the 1970s to the 1990s, is referred to as the Comprehension Revolution. During this time, researchers developed an increased understanding that reading was more than just decoding. By the late 1990s, reading comprehension had become thought of as an active experience, rather than a passive one. Therefore, the emphasis was placed on students to pay attention and be on task as they read. Reading had also become conceptualized more as a cognitive act. Thus, as the student read, they were expected to be able to talk about what they read (Pressley, 2001). It was also determined that readers read

for a purpose. A reader can read to be entertained, to learn something or to find information to use (National Reading Panel, 2000b).

Research has discovered a variety of comprehension strategies to improve reading comprehension (Duke & Pearson, 2001). "Strategies are defined as learning techniques, behaviors, problem-solving or study skills which make learning more effective and efficient" (Singhal, 2001, p. 1). Comprehension strategies are sets of steps that readers use to make sense of what they read.

Reading comprehension strategies can be taught by a classroom teacher in an explicit and formal manner. The teacher first demonstrates and models a specific strategy. Then they teacher needs to provide guided practice to allow the student to work with the strategy and become independent with its use. The research findings are very positive for the explicit instruction of comprehension strategies. They help students to achieve significant gains in reading comprehension. However, if students are not purposely taught to use the comprehension strategies, the student will not acquire them. "More information is needed on ways to teach teachers how to use such proven comprehension strategies" (National Reading Panel, 2000a, p. 15).

There are a variety of comprehension strategies that research has proven to be effective for improving comprehension in reading. The National Reading Panel (2000a) concluded that there were seven reading comprehension strategies that have been proven to improve reading comprehension. These seven strategies are comprehension monitoring, cooperative learning, graphic and semantic organizers including story maps, question answering, question generating and summarization.

Good readers are able to self-monitor so that when they misread a word, they can go back and check for what they read incorrectly (Pressley, 2001). This is referred to as metacognition or comprehension monitoring. In comprehension monitoring, students are taught to think about and evaluate the text as they are reading. Students are also taught to pay attention to what they are reading and to recognize when they do not understand what they read. When students realize that they do not understand what they read, they are directed to go back and check to provide clarity and understanding (National Reading Panel, 2000b).

Cooperative learning has shown to promote reading comprehension in students. In cooperative learning students work with and teach other students the reading

comprehension strategies. Cooperative learning has been shown to help students learn more due to increased motivation (National Reading Panel, 2000b).

The third strategy to improve reading comprehension is to use graphic organizers to illustrate concepts. Graphic organizers help students remember what they have read (National Reading Panel, 2000a). A graphic organizer is a diagram or picture such as a map, graph or chart. One of the most popular in elementary school is the "web". These graphic organizers look like a spider web in which lines connect a central concept to a variety of related ideas. Graphic organizers can be used with both narrative and expository text. They especially help students to learn in the areas of science and social studies (Armbruster, Lehr, & Osborn, 2003).

The fourth strategy to develop comprehension is answering questions. Research shows that teacher questioning strongly supports and advances students' development in reading comprehension (Armbruster, Lehr, & Osborn, 2003). There are three types of question-answer instruction. Text explicit question-answer instruction is the most basic. In this type of instruction, students look back in the text to find the answer to the question. Text implicit question-answer instruction is more difficult.

Here, the student has to recognize implied information based on the text. The third and most difficult type of question-answer instruction is called sciptal. In this type of instruction, the answer is not found in the text at all. A reader must use background knowledge to answer the question (Armbruster, Lehr, & Osborn, 2003).

Generating questions is also a strategy for developing comprehension. When generating questions, the student must first read and understand a sentence or paragraph and then compose a question based on the knowledge learned from what was read. Generating questions require the reader to integrate information, such as the main idea, from one paragraph and then relate it to other important information in the text (Armbruster, Lehr, & Osborn, 2003).

The fifth strategy to develop reading comprehension is referred to as recognizing story structure. It helps students comprehend stories and answer questions on the content of the story. It is used exclusively when students read narrative text. This strategy has the greatest effect on poor readers. On the other hand, good readers do not seem to benefit as much from it (National Reading Panel, 2000).

Summarizing is the sixth strategy to develop reading comprehension. In summarization, the reader must read and understand the text and then write the main ideas in their own words. In order to summarize a reading passage, a student has to have some writing skills. Summarization helps in the recall of the main idea and details of what was read (National Reading Panel, 2000a).

In general, the evidence suggests that teaching a combination of reading comprehension strategies is the most effective in improving reading comprehension. When students use these strategies appropriately, they will increase their comprehension in reading (Pressley, 2001). "When used in combination, these techniques can [also] improve results in standardized comprehension tests" (National Reading Panel, 2000a, p. 15).

One of the major research finding discovered was that students need more instructional time with reading and extensive reading opportunities. In order to promote reading comprehension, students need to read much more in class, and then be given opportunities to talk about their reading (Duke, 2004).

There is some evidence for teaching reading comprehension strategies to younger primary students. Students in Kindergarten through the third grade made

great improvement in reading comprehension when given instruction in reading comprehension strategies (Pressley, 2001).

In the elementary school setting, it has become increasingly evident that teachers rely too much on teaching reading with narrative stories. Students need to be reading with and practicing expository text such as that used in science and social studies books (Pressley, 2001). The literature also suggests that teaching comprehension in the context of specific academic areas such as social studies can be very effective. "If this is true of other subject areas, then it might be [more] efficient to teach comprehension as a skill in content areas" (National Reading Panel, 2000a, p. 15).

No Child Left Behind and Educational Testing

In today's educational setting, in order to raise the academic achievement of the student, the instructor needs to have a full understanding of the issues and problems that face education with the enactment of the No Child Left Behind Act of 2001. It is also important to understand the role of educational testing as used under the law and how it has impacted education.

America's educational system was designed for another time, a time when workers only needed a basic education to obtain a good standard of living. However, the world has changed to a global economy. American workers are now in direct competition with workers from other countries. Work that required low-skilled labor has quickly been exported to countries where workers can be paid much less than workers in the United States. Fewer and fewer jobs can be found in the United States that require only low-skilled workers. In addition, computers and machines can do the low-skilled, repetitive work faster and cheaper. These new technologies have taken the low-skilled jobs that once used to employ thousands of workers. Jobs that used to be performed by people, such as those found in the American auto and textile industries, have been automated and converted to computer-operated machines (National Center on Education and the Economy, 2007).

In the future, jobs will demand more highly skilled workers. According to the National Center on Education and the Economy (2007), "This is a world in which a very high level of preparation in reading, writing, speaking, mathematics, science, literature, history and the arts will be an indispensable foundation for everything that comes after for most members of the workforce" (p. 6).

Beyond this [workers] will have to be good at both analysis and synthesis, self-disciplined and well organized. The National Center on Education and the Economy (2007) state:

> It is a world in which comfort with ideas and abstractions is the passport to a good job, in which creativity and innovation are the key to a good life, in which high levels of education - a very different kind of education than most of us have had - are going to be the only security there is. (p. 6-7)

Unfortunately, for many, America's public schools have not delivered a sense of security. For decades, America's public school students have been "socially" promoted to the next grade without having to prove that they had acquired the skills and knowledge associated with the grade that they were in, let alone the more advanced skills of analysis, synthesis and self-discipline. This philosophy, in conjunction with the rise of global economics and the lack of well paying jobs, has helped create a staggering crisis in America that has manifested itself in crime, unemployment and hopeless poverty (Paige, 2003; Hanson, 2006). These problems are especially bad for America's minorities where:

> 60 percent of black fourth-graders and 56 percent of Hispanic fourth-graders failed to demonstrate even 'basic' reading skills, while 46 percent of black fourth-graders and 38 percent of Hispanic fourth-graders failed to show mastery of even 'basic' math skills. (Hess, 2004, para. 5)

Since America's public schools are driven by American tax dollars, taxpayers and lawmakers have increasingly demanded that schools prove that they deserve to continue to receive public funds. This demand for proof has been termed the Accountability movement. Public schools are now expected to be accountable and to show that their efforts are producing results. As Petterway and Kritsonis (2006) say, "The time when public schools are allowed to operate without proven success is over" (p. 3). In the past, schools and teachers could claim that their students were learning. Now, with the authorization of the No Child Left Behind Act, America's public schools can no longer just say that their students are learning, they must prove that their students are learning.

The No Child Left Behind Act of 2001 (Public Law 107-110), commonly known as NCLB, is a United States federal law. It was signed on January 8, 2002 by President Bush three days into his first term. It authorizes a

number of federal programs with the goal to improve the performance of U.S. elementary and secondary schools by increasing the standards of accountability for states, school districts and schools.

NCLB is an attempt to reform and improve America's public educational system. With the understanding that the current public school system is failing to adequately educate America's youth or to prepare them to meet the demands of the 21st century, No Child Left Behind was designed with several goals in mind (Petterway & Kritsonis, 2006). First, NCLB was passed to raise current academic standards. Many schools and teachers have developed over time, what President Bush describes as the "... soft bigotry of low expectations ..." (U.S. Department of Education, 2007, p. 1). Therefore, it was also designed to encourage the development of even higher academic standards in American education. Expectations will continue to rise, and work for students will get even more demanding in order to prepare American students for the emerging global economy. Another major goal of NCLB is to make America's schools do a better job of educating its students. NCLB also seeks to emphasize the importance of being able to read well by the end of the third grade. In fact, NCLB expects every student to test at or above the

Proficient level in the tested areas of Reading/Language Arts and Mathematics by 2014. Therefore, the purpose of NCLB is to raise the academic achievement of all students in the United States.

In order to meet these goals, NCLB requires states to implement many different features. NCLB requires all states to implement a statewide accountability system for all schools and their students. This system must first establish a challenging set of academic content standards for the core academic subjects such as Reading/Language Arts, Mathematics, and Science. The second feature required under NCLB is a definition of student proficiency. These performance standards define the minimum expected levels of attainment necessary for a student to be deemed proficient within a specific academic subject. NCLB also requires that schools annually measure the students' progress in Reading/Language Arts, Mathematics and Science. Therefore, a state testing system must be established to determine if their students are meeting the states standards (Barton, 2004; Petterway & Kritsonis, 2006). Perhaps the most controversial of the features is that NCLB sets a timeline for the progress that school districts, schools and students must meet (Wenning, Herdman, Smith, McMahon, & Washington, 2003).

The current timeline expects all school districts and schools to have all of their students (100%) test at the proficient level by 2014 (Hanson, 2006).

NCLB holds students accountable for results. Schools are held responsible for ensuring that their students improve and make progress toward reading and math proficiency. Schools must now show that their students are improving and that the school is moving toward the goal of having 100% of their students reaching the proficient level or face sanctions. Districts and schools must make adequate yearly progress (AYP) toward proficiency goals each year. After two years of not making the AYP goal, the school will be labeled as an Improvement School. After four years of not making the AYP goal, the school will be subject to corrective action such as replacing staff. After five consecutive years of not making Adequate Yearly Progress, the school will be subject to major restructuring measures such as replacing the entire staff. The idea is to get the teachers to do a better job of educating their students and having the students perform better on standardized tests (Petterway & Kritsonis, 2006; Hanson, 2006).

There are many opposing viewpoints and complaints concerning the implementation and use of NCLB.

1) The law makes schools test too much.
2) The law encourages teaching to the test. However, former Secretary of Education countered, "... there is nothing wrong with 'teaching to the test', if you are teaching something that students need to learn" (Paige, 2003, para. 85).
3) The law narrows the curriculum and takes the fun and creativity out of education.
4) The law is poorly funded or not fully funded.
5) The law punishes the educational system.
6) The law unfairly labels schools as failing.
7) The law has unreal expectations of having their entire student population test at or above grade level by 2014.
8) The law is set up to make the educational system fail so it can be restructured, reorganized or privatized.
9) One of the main concerns is the rigid use of only testing to determine if students are making progress (Hanson, 2006). Many groups are trying to incorporate other measures in addition to standardized testing to measure what a student has accomplished.

10) Another complaint is that standardized tests only test basic skills rather than higher-level thinking skills (Hanson, 2006).

11) The law has created value conflicts within teachers. Teachers must struggle with the ideas of NCLB and the new requirements for teaching and their own personal principles.

According to Margaret Spellings, head of the U.S. Department of Education, these are just "... myths and misconceptions ..." (Ashby, 2007, p. 2). The government knows of these complaints and refuses to accept them as excuses (Ashby, 2007). NCLB requirements are not going to be eased or forgotten. In fact, NCLB demands will continue to become more and more demanding, and it will affect all aspects of teaching and education. Therefore, it is extremely important to understand the complexities of NCLB and to understand how these accountability measures will affect every teacher and student in their daily educational activities.

The NCLB accountability system is a structure. The foundation of that structure consists of four distinct aspects:

1. The content standards.
2. The performance standards.

3. The curriculum.

4. The test (Barton, 2004).

Content Standards

The first aspect in the NCLB accountability system is referred to as the content standards. Standards are important. They define what is necessary and expected. All throughout society standards have been established to improve the American people's quality of life. Standards have been set for medical procedures, financial dealings, food production, water quality and the construction of buildings. These standards have been put in place to protect our children and to make our daily lives better. It is hoped that content standards will also make American's lives better and the American educational system more effective (Kendall & Marzano, 2004).

Under NCLB, content standards are determined by the educational system. In most cases today, this means that each state determines their own set of content standards. Content standards, therefore, are the educational vision of the state. They represent what a society understands and agrees upon to be of importance to the education of its students. The academic content standards are the centerpiece on which all future changes and reforms are

based upon. They define the way states approach education issues. The selection of textbooks, the development of professional in-services, the creation of statewide initiatives and programs as well as the allocation of financial resources, the alignment of curriculum and academic assessment are all decisions that are based on the state's academic content standards.

At this time, there is no consistent agreement on the use of terms that describe content standards. Different states use different terms to communicate the same ideas. Other states use the same terms, but use different definitions. This has lead to great confusion among states and educators. There is a great need to standardize terms and definitions in order to facilitate the creation and use of academic content standards (Kendall, Richardson, & Ryan, 2005).

The development of content standards is a long and complicated process with many different organizational levels. There is no set consensus on the best way to determine the content standards. However, some techniques have proven successful (Kendall & Marzano, 2004). One process of determining a state's academic content standards begins by establishing "... a set of common expectations for what all students should know and be able

to do upon completion of high school" (Ohio Department of Education, 2007, para. 6). The content standards are then developed based on what skills and knowledge are necessary to meet those "common expectations" of a high school graduate.

The first organizational level is general. The content is first categorized into content areas and then organized and sequenced by grade level. The content areas are often described as broad statements that provide a general framework for further organization and specificity. These statements organize a subject area into a manageable number of general goals. There are usually between five and twelve statements in any given content area. At this organizational level, the content standard is used to clarify a general goal within a content area and to help the teacher find the standard within a large document when searching for a particular standard. They are, however, too broad to be used to grade students or plan lessons (Kendall, Richardson, & Ryan, 2005). These general objectives only indicate the general type of desired learning that is expected. They are not specific. They are not measurable and, they cannot be tested (Stewart, 1975).

The general objectives are then narrowed down into strands or topics. This is the organizational level between general objective and the benchmark. The purpose of the strand or topic is to help teachers find a benchmark, to make connections easier for cross-curricular planning, and to facilitate communication between the teacher and parents.

The strand or topic is then organized into benchmarks. A benchmark is an explanation of the exact knowledge or skill that students should learn by a particular grade level. It is also referred to as an indicator or learning expectation. It is specific enough so that all stakeholders understand what should be learned in the classroom.

Benchmarks can be further narrowed down to the specific knowledge or skills needed to attain the benchmark. These are the specific objectives. Specific objectives are defined and precise. They are measurable and can be tested. Most importantly, they specify exactly what the learning is supposed to be. The specific objectives explain to teachers the exact learning that should occur in the classroom. They are specific enough to be used to guide assessment, to grade students and plan lessons.

Content standards, therefore, are written as general statements, but "... represent what education policy makers want students to know and be able to do" (Barton, 2004, p. 17). They are statements that describe the knowledge students should acquire, and the skills that they should develop and attain through the course of a K-12 education. Content standards also represent what students are expected to have mastered by the end of each grade level. They are the ultimate objectives of educational institutions. It is extremely important to remember that the academic content objectives are classified as the minimum learning that a student should be able to do by the end of a specified grade level (Stewart, 1975).

The purpose of the content standards is to shape overall instruction to meet the ultimate goal of a well-educated high school graduate. The content standards specify what should be taught at each grade level (Crocker & Zieky, 1995). Therefore, the classroom curriculum must be aligned with the content standards. However, the content standards do not specify how the specific knowledge or skills have to be taught.

The academic content standards also provide guidance for assessment. It is imperative that content standards be

aligned with the state's assessment so that students can be prepared for the state's yearly standardized tests as well as the High School Exit Exam. The content standards are what will be tested by the state on the standardized test.

As a nation, the development of strong content standards has not occurred in most states, leaving more than half of America's students with mediocre standards (Finn, Julian, & Petrilli, 2006). While many states have revised their standards, in most cases, the states' standards do not offer clarity, rigor or strong content description. On the other hand, content standards have been developed by some states to be very rigorous in nature. These rigorous content standards are based on the idea that students can learn far more than what is currently being expected in today's schools. Some of the strongest and most rigorous academic content standards in the nation are now found in California (Finn, Julian, & Petrilli, 2006).

Performance Standards

The second aspect in the NCLB accountability system is called the performance standards. Often overlooked and easily misunderstood, performance standards have become

very important and one of the most controversial topics of the NCLB Act. States are now required to report how many students have tested at the "proficient" level in reading, writing and science. Each state must also set goals so that an ever increasing amount of students test at the proficient level each year until all of the students test at the proficient level by the year 2014.

Performance standards, also called proficiency standards or performance levels, are the expected level of achievement at a designated level. They indicate how well a student has learned the content standards that were tested. Performance standards designate what students must do to demonstrate various levels of proficiency with respect to specific content (Crocker & Zieky, 1995). In other words, "... performance standards indicate how much of the content the students have mastered" (Barton, 2004, p. 20).

States are given great latitude and flexibility to determine their own performance levels. However, these levels are usually termed Basic, Proficient and Advanced. While the law allows states to determine their own performance standards and to define what a student can do at each of the performance levels, the law does not say how the states must set the performance levels. Therefore,

one state's definition of proficient can be substantially different from another state's definition. Because of these widely varying definitions in determining what is considered student proficiency, research has shown that a student can be deemed proficient by one state's performance standards, but labeled not proficient in another state (Kingsbury et al., 2003). In fact, there is a wide range in the cut-scores that states use to deem proficiency (U.S. Department of Education, 2007). In addition, the state's idea of proficiency differs significantly from what the National Assessment of Educational Progress (NAEP) has determined to be proficient. Therefore, it is not reasonable to punish failing schools in one state where the performance standards have been set high, such as in California, and reward schools in other states who have met Adequate Yearly Progress (AYP) where the performance standards have been set quite low.

In order to establish performance standards a scale on which to set them is needed (Green, 1996). The scale is usually divided into levels and labeled Basic, Proficient, and Advanced. In California, however, there are five performance levels for reporting the individual results of a student's standardized test. They are labeled Far below

Basic, Below Basic, Basic, Proficient and Advanced. It is important to understand that these performance levels are based on a continuum of development. The lines that are drawn between the levels do not mark clear distinct student levels of development.

The scale is composed of both a written description and a numeric score. The written description of the scale requires just a short paragraph to describe what it means to be at any one of the achievement levels. The written descriptions are often vague and require extensive work on the part of the teacher or parent to understand. For example, the Proficient level can be described as "being of satisfactory level" or "indicative of solid understanding". On the other hand, the description for a Basic level is described as "an academic performance that is marginal", "that it displays partial understanding" or "that there is a need for further work to achieve the proficient level". The NAEP's national standards are not much more descriptive. These standards describe a fourth grade student in reading at the Basic level as "one who should demonstrate an understanding of the overall meaning of what they read". A fourth grade student when reading at the Proficient level however, should be able to "...demonstrate an overall understanding of the text..."

(National Assessment of Educational Progress, 2005, para. 3).

Along with a written description of the performance level, there is a numeric scale score. Numeric scores are used when multiple-choice standardized tests are the measure of a student's reading, mathematic or science knowledge and skills. In California, the scale scores range from a low of 150 to a high of 600. The numeric score corresponds to the number or percent of questions that a student must answer correctly at each performance level. This numeric score is called the cutpoint. The cutpoint on a scale is the delineator between two predetermined performance levels (Barton, 2004). Each state sets its own cutpoints, and they vary considerably. Unfortunately, the cutpoint scores and the ranges of the levels are not consistent from grade to grade or from subject to subject. This makes grade level comparisons difficult. A student could actually have a higher numeric score than earned the previous year and still drop a performance level. On the other hand, a student can earn a lower numeric score than what was achieved the previous year and rise up to the next performance level.

Therefore, the development of a valid and acceptable way to establish performance standards is still in debate.

Congress has required that performance standards or levels be labeled developmental until the debate is settled (Barton, 2004).

There is a variety of traditional methods for setting cutpoints on tests. The Angoff method looks at the test questions of a borderline proficient student. By identifying and analyzing the questions the borderline student missed, the test analyzer can then write a description of what it means to be proficient. The Bookmarking method sorts all the test questions in order from least difficult to most difficult. Members of a panel of qualified people then go up the list until reaching the point where they feel a certain level of performance needs to be attained (Barton, 2004). Other common methods used to set performance levels are the Kentucky Synthesis method, the Contrasting Group method and the Haegar-Mills method.

These methods are standard setting processes. They do not tell us whether the breadth or depth of the content standards is being achieved or what proportion is being achieved. The methods stipulate rigorous procedures, but they still rely on personal judgment. Currently, there is no agreement as to which is the best method (Barton, 2004). In addition, research shows that each of the

different methods produces a different cutoff score on the same state assessment.

The performance standards need to be aligned with the standardized test. The performance standards must also be aligned with the content standards. The performance standards have to reflect the content standards for which they represent. As Green (1996) states:

> Logically, it would seem preferable for the judges to set standards just on the content domain. They could identify what parts of the domain are basic, what parts go with proficient persons, and what parts would be mastered by advanced students. It is not at all clear how to do this, but a way might be found. (p. 18)

Curriculum

The third aspect in the NCLB accountability system is called curriculum. The NCLB Act has brought significant changes in regards to the state curriculum. The core curriculum of many states has been revised. The development and implementation of a standards based curriculum has been the result for others.

The curriculum is "... what is actually being taught in the classroom" (Barton, 2004, p. 18). It is what

teachers do and what they are expected to do in the classroom. The curriculum is based on how teachers interpret the state standards and implement learning activities to achieve those standards. Thus, curriculum is shaped by the adopted content standards.

Therefore, there is a need for the curriculum to be aligned with the content standards and the standardized test. In order for America's schools to better prepare students for the challenges they will face in the rising global economy, the curriculum in the nation's schools need to be aligned to the content standards so that what is taught in the classroom matches what is considered to be the important skills and knowledge identified in the content standards. The curriculum also needs to be aligned with the state's standardized test. This alignment of curriculum with the state content standards and the state standardized test is important in all grades from Kindergarten through the twelfth grade. It is essential that the curriculum, instruction and the state assessment be aligned and consistent through all grade levels to promote student achievement to the fullest possible extent.

With the adoption of NCLB, states are creating and implementing a more rigorous curriculum. Schools are

holding students to higher standards. Not only does the curriculum expect students to learn more content and skills, there is a new emphasis on the mastery of these skills and knowledge. In many cases, the curriculum has become more challenging so that students can be better prepared for college and other post secondary educational opportunities. Students who are involved in a more rigorous curriculum are better prepared for college (Learning Point Associates, 2005). In addition, research shows that the same skills that are needed for college are now the same skills that are expected in the workplace (Camacho & Cook, 2007).

The level of expectation and depth of coverage of the content standards varies widely among schools and classes. Inner city schools have tended to have lower expectations of there students. This has been repeatedly revealed in the standardized test scores. Teachers have complained that it is difficult to fit the entire curriculum into a school day and still achieve adequate standardized test scores in the subject areas of reading, mathematics and science that are tested for accountability (Barton, 2004).

High-stakes standardized tests for accountability have also affected teaching and the shape of instruction in two other ways. "One [way] is what happens to the

teaching of the subjects being tested" (Barton, 2004, p. 33). The core academic subjects have become the central focus under NCLB (Petterway & Kritsonis, 2006). Language arts, writing, math and science have garnered greater attention and increases in instructional time because they are tested. Some teachers hold views that progress to proficiency is so important in these basic subjects that there is good reason for redirecting instructional time (Barton, 2004). They hold the view that a student needs to be able to read, write and perform mathematics skills in order for the student to have a chance to perform well at the next level of education. They feel that if a student does not have these skills then teachers have the duty and responsibility to help students in these core academic areas rather than sacrifice instructional time on art, music and physical education. This narrowed curriculum is often referred to as "teaching to the test."

The other way NCLB and high-stakes standardized tests for accountability has affected the curriculum is what happens to the teaching of the subjects that are not being tested. If a subject is not tested, it is considered less important and often eliminated from the instructional day. The curriculum is essentially limited to subjects that appear on the standardized test. Schools cut non-core

curriculum subjects. Foreign languages are cut because the test does not include them. Art, music and physical education have had a decreased emphasis because they are not tested. Teachers also report deemphasizing untested areas in favor of tested ones. The subject of social studies is not emphasized because it is not currently tested. Barton (2004) found the following:

> Despite statements about the importance of geography, history, civics/social studies, health, art and music, if these subjects are not tested, the time spent on them in the classroom appears likely to be reduced in favor of subjects that are assessed - particularly those with high stakes standardized testing attached to the results. (p. 41)

Testing

Testing is the fourth component of the No Child Left Behind accountability system. Under the No Child Left Behind Act, all students must now be tested in reading, mathematics, and science. It is the most controversial component and the most scrutinized. While the main objective of testing in accountability is the final result, this is just one of the many ways to use testing

to improve schools and the academic achievement of the students' they serve.

The purpose of testing is to collect and provide information. Traditionally in education, a test is a way for a teacher to determine if a student had learned the required content or developed the needed skill. "Given that there is something that should be learned by the students, evaluation of the achievement of that something should provide two important facts: how much of that something has been learned and how much of it hasn't been learned" (Stewart, 1975, p. 457). This information can then be used to make decisions on what the students still need to learn and how to help them learn it effectively and efficiently.

An assessment is made up of a set of questions or items that a student is expected to answer. It can be referred to as an examination, exam, a test, a quiz, or a pop-quiz. The testing process has three distinct parts: the administration of the test, the scoring of the test and the use of the scored results (Fremer & Wall, 2003).

Traditionally tests can be described in two main ways. These two ways are designated as such based on the evaluation processes of the test. The evaluation of a test

may be primarily based on a subjective basis, or on an objective basis.

A subjective test is a vague measurement of students' performance based on feelings, attitudes or opinions that commonly result in wide a disparity in grading by different teachers for the same performance (Stewart, 1975). The evaluation of essays, term papers and projects are subjective.

An objective test, on the other hand, is a more specific measurement of students' performance based on achievement of specific measurable objectives which results in uniformity in grading by different teachers for the same performance (Stewart, 1975).

Tests have been used in a variety of ways. Tests in education have been used extensively in the United States for measuring IQ; for placement or promotion into a grade level; for determining if a young student is ready to start Kindergarten; and for diagnosing whether a student has special needs. Tests are also used for comparing students to other students; for judging the progress of students in developing their skills as they proceed through a class and their schooling; and for sorting students; for self-discovery such as determining the

attitudes and motivations of students (Barton, 2004; Fremer & Wall, 2003).

Now with the adoption of the No Child Left Behind Act testing will be used for determining if the content standards are being mastered and for the accountability evaluations of schools and school districts.

Depending on the intended use of the test, there are two main types of test evaluation: formative and summative. Formative evaluation is used to evaluate instruction and to make changes so that the students will learn more. These types of tests are used during the school year. They can help teachers by showing the student's strengths and weaknesses. These tests also include many test questions that help in diagnosing how the student's make their errors. Formative tests are usually criterion-referenced tests and mastery oriented (Stewart, 1975). Formative assessment increases student achievement with an effect size between 0.4 and 0.7, and can help low achievers more than average and high achieving students (Black & Wiliam, 1998). Research has also shown that frequent use of formative assessment (every week or two) and the subsequent use of the information provided by these tests to identify and address the students' learning needs have been linked to

improved student academic performance and self esteem (Barton, 2004).

Summative evaluation, on the other hand, is primarily used at the end of a course or school year, and can be used to assign grades. These tests are general purpose or survey tests that have a few questions in each content area or skill. Scoring is essentially based on a predetermined curve of expected scores. Summative tests can be either norm-referenced tests or criterion-referenced tests. A summative test is the standard accountability test used by most states for the NCLB assessment. These summative evaluations are called standardized tests (Stewart, 1975).

Standardized Testing

The use of standardized tests has become popular. According to Richard P. Phelps (2006), "... those in favor of high-stakes standardized testing outnumber those opposed at ratios as high as twelve to one" (p. 1).

A standardized test is a test that is administered and scored in a standard manner. They are developed by professionals and administered under regular, ordered circumstances. The test questions are based on the state content standards for each grade level in each subject

area, such as mathematics, language arts and science. Consequently, the tests are aligned with the state content standards. All students answer the same questions, usually in multiple-choice format, and each question has only one correct answer.

A standardized test is a tool designed to measure student performance relative to all others taking the same test. It compares the performance of every individual subject with a norm or criterion. The scores can then be used to evaluate a student, a school or a school district.

"The purpose of the test is to measure the degree of achievement of the standards" (Barton, 2004, p. 14). Barton (2004) states the following:

> Tests are a collection of questions and tasks that represent a sampling of a 'domain' of knowledge, or a subject area such as eighth grade math. The tests should provide an estimate of how much the student has mastered the domain of knowledge. (p. 33)

In other words, testing is conducted to measure how much of the content standards the students have learned or mastered.

While the standardized test does determine whether the student has met the state standards and how well the student has performed in each subject area, the

standardized test results are also used for a variety of other measurements. The results indicate how students compare to a variety of different socio-economic and ethnic groups. They also indicate how well teachers are preparing their students. Finally, the results of the standardized tests indicate how schools and school districts have performed at preparing their students. As required by the NCLB Act, all of these results must be reported. Therefore, by simply using the internet or reading a newspaper, any concerned individual can see how any school or school district is performing. This has put pressure on and caused great stress to those teachers and schools that are deemed low performing or in need of improvement.

For this reason, standardized testing is often referred to as "high-stakes" testing.

> A test is a high-stakes test if the results have perceived or real consequences for staff, students, or schools. Consequences of high-stakes tests may include grade promotion, high school graduation, academic probation, allocation of resources and financial incentives for schools and teachers. (Togut, 2004, p. 12)

Standardized test scores are in some states used as the main or only condition for graduation. For example, in addition to having earned enough credits, the state of California requires the High School Exit Examination. Passing scores on these exit exams are required by a student to graduate high school and to receive a high school diploma. Though many organizations recommend that major educational decisions not be based solely on a standardized test score, the truth is there are many high-school seniors not receiving a diploma because the cannot pass this test.

High-stakes standardized testing has had a wide range of both positive and negative effects. The main benefit is the new data and information that can be used to guide instruction. Other benefits include more efficient, aligned instructional practices and motivation to reach the more demanding set goals. High-stakes standardized testing has also promoted better professional development for teachers, focused classroom assessment, created an environment of individualized instruction and developed teachers with a stronger understanding of their subject matter.

On the other hand, standardized tests have had significant negative impact on teaching. Some research

reveals that standardized testing emphasizes low level thinking skills, narrows the curriculum and encourages teachers to spend more time with test preparation (McMillan, 2005). Another negative aspect of standardized testing is called testing bias. Testing bias includes the idea that prejudice can exist within standardized tests. Typically, test makers are from an upper middle class, white background. Standardized testing often matches the values, previous experiences and language used by the test makers. Tests are also written in Standard English. Students who speak a second language or something other than Standard English have greater difficulty with the reading, vocabulary and comprehension of a text.

There are many other disadvantages to standardized testing. One is that not all tests are well written. In addition, standardized tests by their very nature have poor coverage of the desired curriculum. In addition, some standardized tests include essay questions, and there have been recent criticism of the effectiveness of the grading procedures of these essay questions. Perhaps, the most controversial aspect of standardized testing is the computerized grading of tests in which the tests have repeatedly been found to be inaccurately corrected.

Finally, some standardized tests contain multiple-choice questions with ambiguous answers.

Multiple Choice Tests

Multiple-choice tests are the primary format for most standardized tests, including state exams and commercial achievements tests. They are used because they can be easily, inexpensively and quickly scored by computers. Although multiple-choice tests have the reputation of only testing for low level thinking skills such as recall, many new multiple-choice tests require higher-level thinking skills.

The first multiple-choice test was developed at the University of Kansas in 1914 by Frederick J. Kelly.

Multiple-choice tests are sometimes called selected-response tests. These tests are usually made up of many test questions. These test questions are referred to as test items. For each question, the test-taker is required to select the best choice from a set of four or five possible answers. The student then marks an answer on an answer sheet by filling in a bubble that corresponds to the letter of the answer. The answer sheets are then graded by machine. Each test question is given one point

for a correct answer, and correct answers are added together to determine an overall score.

There are several different versions of the multiple-choice format. The most common versions come with three, four or five choices. Matching test items are multiple-choice test items with many choices. True-false test items are considered multiple-choice test items with just two choices (Stewart, 1975).

A standardized multiple-choice test item is made up of two essential parts. The first part is the problem. The problem is referred to as the stem (Burton, Sudweeks, Merrill, & Wood, 1991). The stem can be a question or an incomplete statement.

The second required part of a multiple-choice test question is the list of suggested answers or options. The options consist of two basic parts: the correct answer and the incorrect but tempting distractors (Kehoe, 1995). Distractors are false or inferior options. "The purpose of the distractors is to appear as plausible solutions to the problem for those students who have not achieved the objective being measured by the test item" (Burton, Sudweeks, Merrill, & Wood, 1991, p. 5). Therefore, when a teacher gives students multiple-choice tests, they are not only testing to see what the student knows, but also if

the student can take a multiple choice test by identifying correct answers, and discerning the wrong answer choices (Stewart, 1975).

Test-makers often promote multiple-choice tests as objective. This is because there is no human judgment involved with the scoring. Although multiple choice test items are referred to as objective type test items, they are only objective from the point of view of scoring or grading (Stewart, 1975). The actual writing of a so-called objective test item, however, is very subjective.

Norm-Referenced Tests

There are two main types of standardized tests: norm-referenced tests and criterion-referenced tests. The two tests differ in how they are scored, their intended purpose, how there content is chosen and in a variety of other ways (Bond, 1996).

Many of the standardized tests taken in school are norm-referenced tests. A normative or norm-referenced test, which is used throughout a state, region or the nation, is referred to as a standardized test (Stewart, 1975).

Norm-referenced tests (NRTs) are based on the belief that there is an average achievement level, and that some

students will score higher and some students will score below this average level.

> Basically, normative or norm-referenced tests are tests that measure a student's achievement with reference to a standard of achievement under conditions where the standard of achievement is further defined in terms of a predetermined curve of achievement scores on that particular test of a representative sample of students at the local, district, state, regional or national levels. (Stewart, 1975, p. 505)

In other words, NRTs compare a student's score against the scores of a group of people who have already taken the same test and set a standard curve.

Norm-referenced tests scores are reported as scale scores, percentile ranks, grade-level equivalents, stanines, or normal-curve equivalents. In order to understand a norm-referenced test score two scores are needed: the test score in question and the average score of the normative group. The test score can then be compared against the average to determine whether it is a good, bad or average score. For example, a student's score may be reported as being in the 60th percentile. This means that the student scored higher than 60 percent of

all the students who took the test, but 40 percent of the students who took the test scored higher than that student did.

Norm-referenced tests measure success by ordering students by rank. The purpose of the NRT is to rank students from high to low. The norm-referenced approach was designed to rank students by percentiles and sort students into tracks or ability groups in K-12 classes. Schools and school districts can then determine whether to place the student in a remedial class or a more advanced program. A teacher can use the results to group students into different ability level groups for reading and mathematics.

The content of an NRT test is selected according to how well it ranks students from high achievers to low (Bond, 1996). Selection of material, questions or items for an NRT is based on how well it discriminates among students rather than how well it correlates to the established objectives or standards. According to Stewart (1975), "A good test item is one that 50 percent of the students miss ..." (p. 507).

Criterion-Referenced Tests

Criterion-referenced tests (CRTs) are based on the belief that at every grade level and in every subject there are basic skills that need to be mastered before the student progresses on to the next more difficult level. It also believes that all students can become proficient in the content standards (Stewart, 1975). Standards-based assessments are based on the belief that all students can succeed if they are tested against high standards, which all students are required to master regardless of ability or economic background.

Criterion-referenced tests measure factual knowledge of a well-defined body of knowledge. They tell what the test taker can do and what they cannot do.

Criterion-referenced scores compare test-takers to a criterion. The tests measure students against a fixed standard of achievement called a cut score or criterion. The cut score is often expressed as a percentile. A score above this percentile is considered a good or passing score, while scores below the percentile are considered bad, failing or not proficient.

The purpose of the CRT is "... to see how well students have learned the knowledge and skills which they are expected to have mastered" (Bond, 1996, p. 2). From

this, a teacher can determine what the student does not know or where the student is having difficulty and correct the problem.

The content of a CRT test is determined by the objectives of a class, school, district or state and its significance to the curriculum. A good test item is one that matches the content standards and tests for what was explicitly expressed.

Summary

The literature important to the project was presented in Chapter 2. Though there are many different learning theories that have elaborate research behind them, most organized instruction has been based on one model: Behaviorism. Its influence on education can be seen in the first developments of Classical Conditioning where subjects are taught to react automatically and involuntarily to a stimulus to Operant Conditioning where the concepts of reinforcement and Programmed Instruction have been shown to facilitate the learning of content and skills. It can also be noted how behaviorism has had a great effect upon the development of teaching machines in the first half of the 20th century and on Computer-Assisted Instruction in the second half.

Reading and reading comprehension skills are the most important academic subject matter in America's public schools today. The importance of reading and reading comprehension is paramount in today's modern society. A person who cannot read with comprehension faces a difficult economic future.

Reading and reading comprehension have a variety of different components. Each, however, is vitally important for the student to have developed and mastered in order to reach the goal of independent reading and reading to learn. Students must have phonemic awareness where they can identify, manipulate and work with the individual sounds of a spoken language. They must have sufficient phonics abilities where they understand how letters are linked to sounds to form letter-sound correspondences and spelling patterns. Students must also develop fluency so that they can read automatically and effortlessly so that they can concentrate on the message within the words and comprehend or understand what they read. Unfortunately, students who do not learn to read by the third grade will most likely not graduate from high school or contribute to the future of society.

Educational testing is a fact for today's student. Under the No Child Left Behind Act, all students must now

test in reading, mathematics, and science. It is a high-stakes test with real consequences for students and schools. Students who do not perform well on them will be retained in the same grade or not allowed to graduate from high school. Schools that do not show considerable growth in standardized test score face

NCLB has created a system that has many different aspects that work together to measure the students' mastery of the skills needed to survive in today's economic world. This system must first establish a challenging set of academic content standards for the core academic subjects such as Reading/Language Arts, Mathematics, and Science. It must define student proficiency and create performance standards that specify the minimum expected levels of attainment necessary for a student to be deemed proficient within a specific academic subject. This system requires that schools annually measure the students' progress in Reading/Language Arts, Mathematics and Science. It also sets a timeline that expects all school districts and schools to have their entire student population test at or above the proficient level by 2014.

In today's educational setting, in order to raise the academic achievement of the student, the instructor needs

to have a full understanding of the issues and problems that face education, as well as what work that has been done and discovered to be effective in making gains in education. The three concepts of Behaviorism, Reading and Reading Comprehension, and Educational Testing under No Child Left Behind can be seen as interrelated. If these concepts are not fully understood and new and revolutionary methods to help students learn and achieve are not put into place, then teachers will not be able to help fully prepare our students for the demands and rigors they will face in the years to come.

CHAPTER THREE

METHODOLOGY

Introduction

The purpose of the study was to investigate whether a web-based Standardized Test Preparation Intervention for Reading Comprehension was a more effective and efficient intervention than a paper-based version at improving, to some degree, the low reading comprehension levels of students within a fourth grade classroom of elementary school students in a large urban elementary school district in Southern California. In this study, causality between the successful use of the computer comprehension software and improved standardized test scores in Language Arts was desired.

For the purposes of this study, quantitative methods were used. An experimental quantitative research design was used in which the student's reading comprehension levels were measured by the Woodcock-McGrew-Werder Mini-Battery of Achievement before and after the use of both the web-based and paper-based versions of the Standardized Test Preparation Intervention for Reading Comprehension. In order to analyze the data collected, SPSS for Windows software was used.

Statistical analysis using UPSS was conducted to determine if the dependent variable, reading comprehension level of fourth grade students, was affected by the independent variables, web-based intervention and paper-based intervention. It was also used to determine whether there was a statistical significance between the two, or whether growth in reading comprehension could be attributed to chance or some other unknown cause.

The dependent variable, reading comprehension level, was defined as the Standard Score as determined by the Woodcock-McGrew-Werder Mini-Battery of Achievement. The independent variable was defined as the type of intervention used for the improvement of reading comprehension intervention, either web-based or paper-based.

Many factors contributed to the need for further exploration into the subject of improving reading comprehension levels in students. Raising reading comprehension levels through the use of computer technology has been judged to be of great importance to state and national interest. In addition, society has placed more and more emphasis on standardized testing of reading comprehension in education. Finally, very little research had been conducted in raising standardized test

scores in reading comprehension with the use of technology.

This study focused on the central question regarding the effectiveness and efficiency of a web-based Standardized Test Preparation Intervention for Reading Comprehension in an elementary school context. It was the purpose of this study to explore the following central question: How will the use of a web-based Standardized Test Preparation Intervention for Reading Comprehension by 4th grade elementary students affect their academic achievement in Reading Comprehension as measured by the Woodcock-McGrew-Werder Mini-Battery of Achievement vs. students who work with a paper-based version of the Standardized Test Preparation Intervention for Reading Comprehension?

Chapter 3 presents the following sections to support this quantitative study: (a) population (b) instrumentation (c) data collection (d) and data analysis. The chapter concludes with a summary to review the main points of the study concerning whether a web-based reading comprehension intervention or a paper-based reading comprehension intervention was more effective in raising reading comprehension levels in fourth grade students.

Population Served

This study utilized a class of 18 fourth grade students at Highland Pacific Elementary School in the San Bernardino City Unified School District. Nine students (50.03%) were female and nine (50.0%) were male.

These students came from a variety of different backgrounds. The students were primarily regular education students from various racial and ethnic backgrounds. There were three African American students (16.6%), one Asian student (0.5%), six Hispanic or Latino students (33.3%), and eight White students (44.4%).

All of the students were in good physical condition. The class did not have a significant population of Special Education students. Only one student, 0.5%, had been diagnosed with a learning problem. None of these students were labeled Severely Emotionally Disturbed. None of the students had any auditory or visual problems. None of the students took any doctor prescribed medications.

Though the school was located within a median-low income suburban community, the majority of the students were from socio-economically disadvantaged homes and some from neighborhoods that had a strong gang influence. Therefore, the students were served by many special programs. The school was identified as a Title I school

and received federal funds to help support all students. Seventy-seven percent of the targeted population received Free/Reduced Lunch. Five students, 26.3%, participated in the English Language Learner program.

Many of the students came from low-income families that received free or reduced lunches. These low-income students comprised 88% of the student population.

The students' attendance rate was also very poor. Student mobility was very high, and many of the students' parents have poor education and/or do not value education.

While some students were behind academically, some students also had low motivation and/or a disconnection from school. A significant number of the students were unmotivated to learn. Many students demonstrated a lack of vocabulary or limited background knowledge on a variety of subjects.

Instrumentation

The Standardized Test Preparation Intervention for Reading Comprehension was designed and developed by the author to help students perform better on standardized tests such as the California Achievement Test (CAT6). It was located at http://readingcomprehension.freehostia.com.

It was a web-based activity in which students signed in, read a passage of text and answered ten comprehension questions. The web-based intervention was designed to self-correct with a click of a button.

The main test page was designed to look like the California Achievement Test (CAT6). When first opening an individual reading passage, a short auditory cue was played. This was designed to gain the student's attention and to illustrate the topic to be read. Examples include the hoot of an owl, the roar of the crowd after the crack of a baseball bat, or Ed Sullivan introducing the Beatles (see Figure 1).

At the top of the page was a gray bar that displayed the reading level of the reading passage. For example, 3.5 represented a reading passage from the third grade, fifth month. Located just under this gray, reading level bar was the title of the reading passage.

The reading passages were previously released texts from the California Achievement Test (CAT6) released tests. This ensured acceptable content for elementary students to read. There are sixty reading passages arranged from the first to the sixth grade level. Each grade level contained ten reading passages. The passages were made up of narrative stories, expository text and

poetry. In general, the text ranged in length from about 100 words for first grade to approximately 600 words for sixth grade.

In the top, right corner of the reading passage was a picture or animated picture that illustrated the reading passage.

Following the reading passage were ten multiple-choice comprehension questions each with four possible answers.

At the bottom of the page was a Check Button that the students clicked to grade the reading passage after they had answered the questions.

A reading passage was considered complete when a student correctly answered nine out of ten (90%) comprehension questions. Upon completion of a reading passage, the students were automatically redirected to a results-positive page. This page served as a behavioral reinforcement. The page contained a Rock and Roll song in midi format (singing of lyrics was not included), the name of the band or artist, the song title, chart position and year, a collage picture of the band or artist, and a congratulatory statement. The students had the option of printing the page as a certificate of completion for that particular reading passage. At the bottom of the page were

links to continue on to the next reading passage or to return to the Table of Contents (see Figure 2).

If the student did not reach the desired 90% correct for a reading passage, the students was automatically redirected to a results-negative page. This page contained an encouragement to try again and two links to retake the previous reading passage or to return to the Table of Contents (see Figure 3).

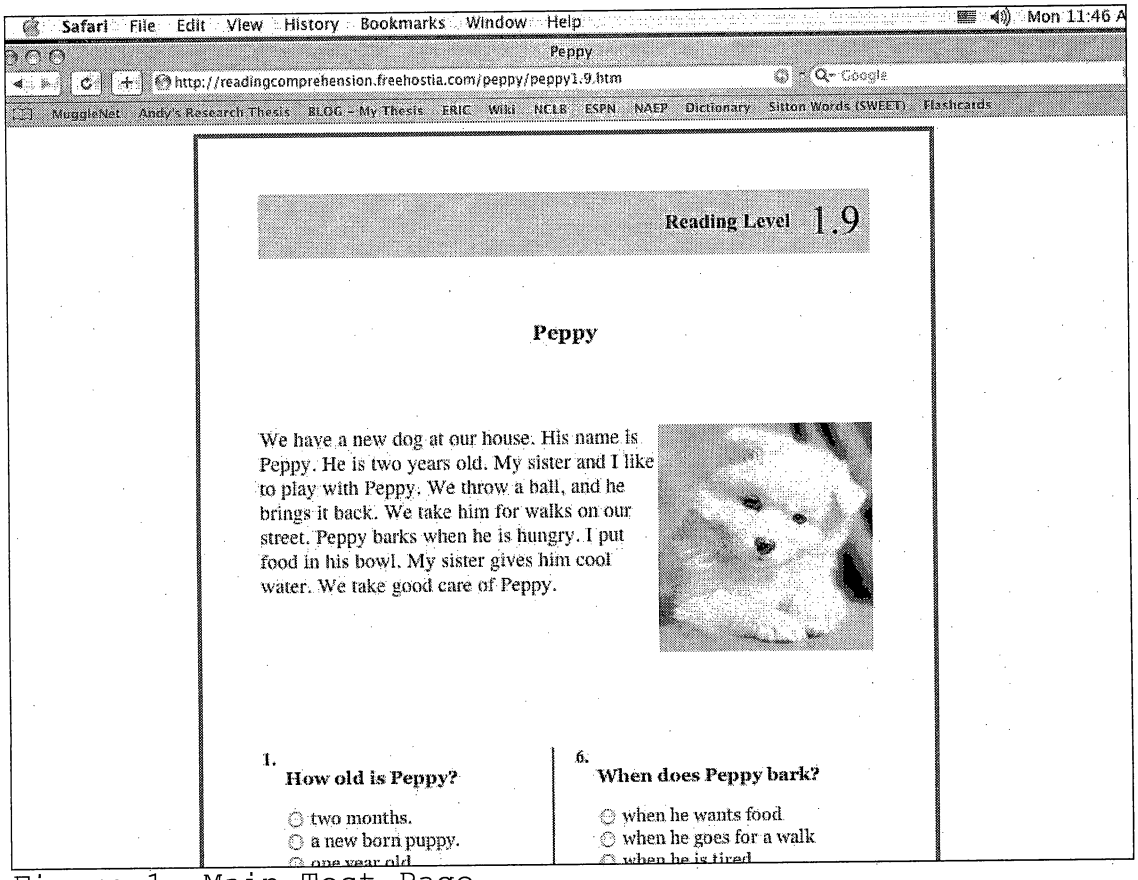

Figure 1. Main Test Page

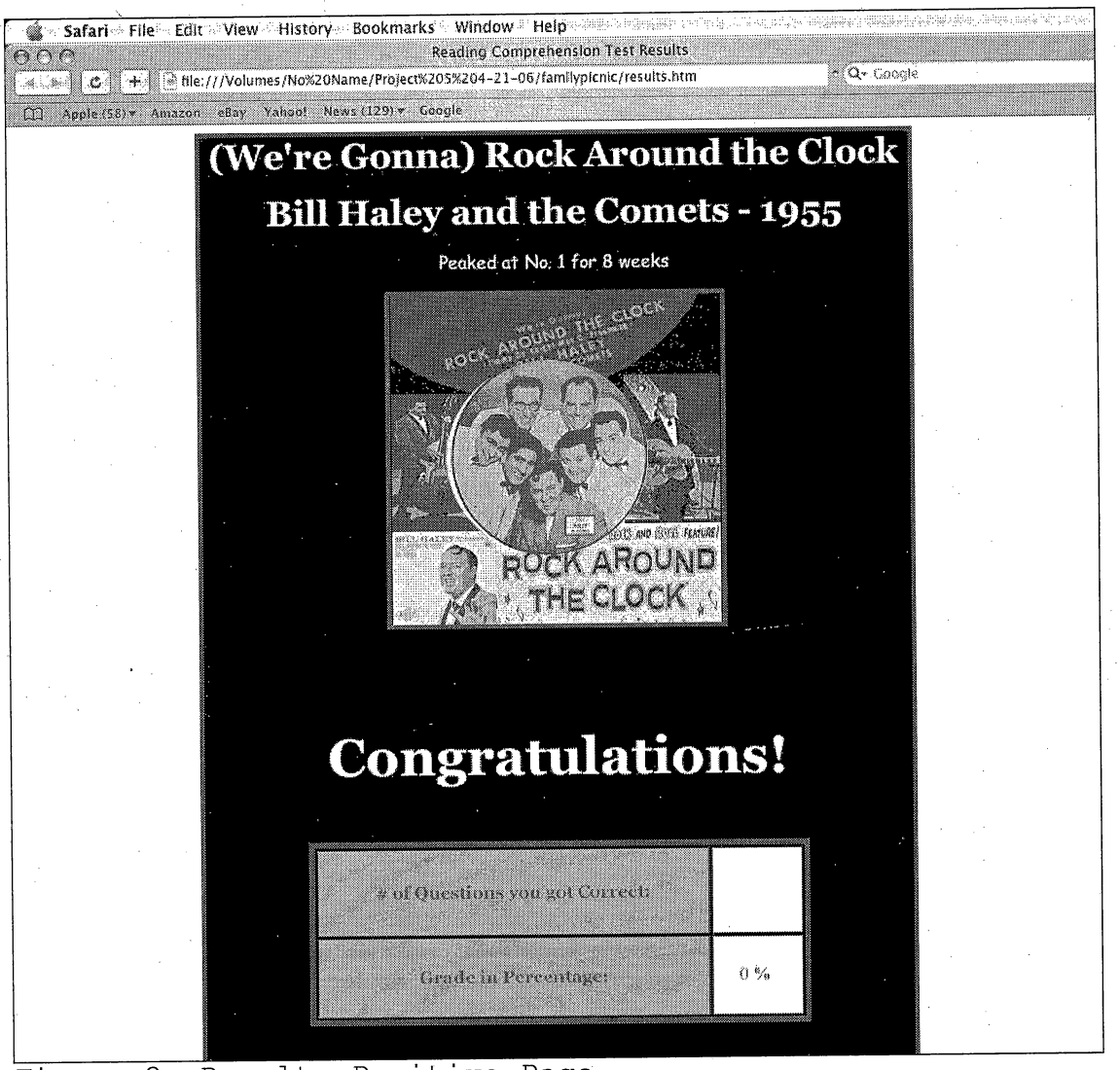

Figure 2. Results Positive Page

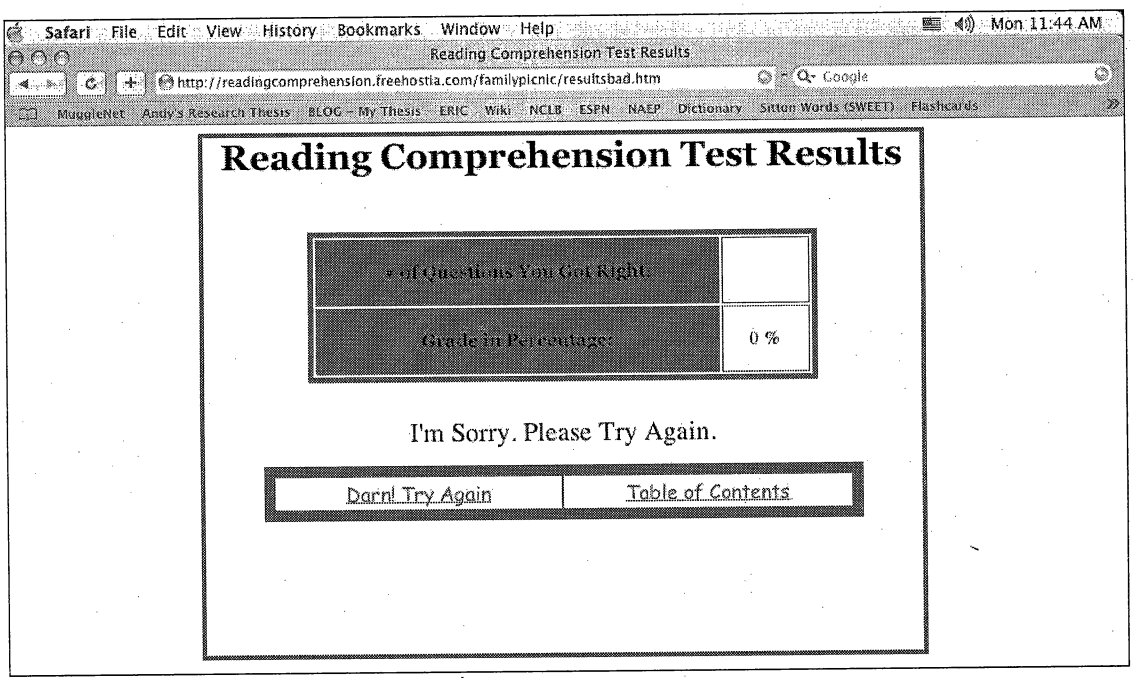

Figure 3. Results Negative Page

Data Collection

The original research question was developed among a community of inquirers. Feedback created a few adjustments and assurance that the question was valid and possible to answer, and offered no bias. After the original question was developed and accepted, the quantitative design tool was selected based on its use and acceptance within the educational community.

Approval of the request to conduct research at Highland-Pacific Elementary School was secured through the San Bernardino City Unified School District designee Dr. Paul Shirk and Principal Brad McDuffee. The participants

were made aware of this project through their regular classroom teacher: Mr. Andy Kubitza. The participants each received in writing, a guarantee of confidentiality, clarification that participation in this study was strictly voluntary, notification that students may discontinue participation at any time, instruction dates and times, the URL of the website, and a thank you letter for participation (See Appendix C & D). The program was conducted from April 1, 2007 though May 30, 2007.

At the inception of the study, an initial assessment of the students' reading levels was conducted on April 1, 2007. Following the protocols as described in the test directions, the Woodcock-McGrew-Werder Mini-Battery of Achievement was administered. The test asked students to identify letters and words, name antonyms for selected vocabulary words, and to complete a sentence with the correct word. After completion of the test, the correct responses were tallied by the teacher. This raw data was then input into the Woodcock-McGrew-Werder Mini-Battery of Achievement diagnostic software program. The diagnostic software program analyzed the data and determined the students' reading levels. Each student's reading level was then recorded as a Standard Score (See Appendix A). This data served as a baseline from which to evaluate the

results after having implemented the Standardized Test Preparation Intervention for Reading Comprehension.

After the pre-test was completed, the students were then given an overview of how to interact with the web-based intervention. The students were directed on how to navigate to the web site, how to register their name, how to sign in, and then to print a hard copy of their progress report. The students were then given a brief introduction to the directions for the reading intervention. The students were directed (a) to read the passage; (b) to read and then answer each question by clicking the radio button next to the answer they thought correct; (c) to click the Check button at the bottom of the page to correct their answers. An example of the first reading passage was modeled for the students to show the students both the positive and negative results pages and what their options were when at each page.

Two groups, a control group and an experimental group, were formed. The subjects were chosen at random from the classroom population. The experimental group practiced with the web-based Standardized Test Preparation Intervention for Reading Comprehension for 30 to 40 minutes a day, five days a week for the 3 weeks. The students used EMac computers with a high-speed internet

connection in a computer lab situation. Each day, the students turned on the computers and self-navigated to the Standardized Test Preparation Intervention for Reading Comprehension at http://readingcomprehension.freehostia.com. The students practiced at the same time together everyday between 9:00 AM and 10:30 AM. This time was at the beginning of the students' normal instructional day.

While the experimental group worked at the computer, the control group practiced with the paper-based version of the Standardized Test Preparation Intervention for Reading Comprehension. At the beginning of each day, the paper-based versions of the test were arranged in order on a table in the back of the computer lab. Each student retrieved the appropriate reading passage, sat down at a center table and began to work. The students used blank 4" X 5" pieces of paper to record their answers. The students were directed to write their answers as A, B, C, and D. When the student completed a reading passage, the teacher acted as facilitator and as the evaluator for each of the reading passages. The teacher corrected the student's answers and determined if the student had achieved a passing score. If the student did, he/she was directed to move on to the next reading passage. If the student did not pass, the student was directed to repeat

the reading passage until they did achieve a passing score.

All students worked silently and uninterrupted. Students were not asked to complete book reports or other types of projects or assessments. The instructor actively monitored students through general observations.

After the initial three week time period in which the Standardized Test Preparation Intervention for Reading Comprehension was implemented, the students were assessed a second time. After completion of this second assessment, the correct responses were tallied and this data was used to determine any growth in the students' reading levels.

The two groups were then reversed. For the next three weeks, the control group became the experimental group and practiced with the web-based intervention while the experimental group became the control group and practiced with the paper-based intervention.

At the end of the second three-week time of implementation, a third and final reading assessment was given to measure the students' reading level and determine whether the students had improved their reading comprehension as a result of practicing with the intervention. The students were again given the Woodcock-McGrew-Werder Mini-Battery of Achievement in the

exact same manner as previously described. A second Standard Score was determined by the Woodcock-McGrew-Werder Mini-Battery of Achievement diagnostic software program (See Appendix B).

All student information and data have been secured and will be maintained for five years after the completion of this quantitative study.

Summary

Chapter 3 described the methodology and quantitative methods that were used for this study. The purpose of the study was to investigate whether a web-based Standardized Test Preparation Intervention for Reading Comprehension was an effective and efficient intervention to remedy the low reading comprehension levels of students within a Fourth grade classroom of elementary school students in a large urban elementary school district in Southern California. This study focused on the effectiveness and efficiency of a web-based Standardized Test Preparation Intervention for Reading Comprehension in an elementary school context. It was the purpose of this study to explore the central question: How will the use of a web-based Standardized Test Preparation intervention by 4th grade elementary students affect their academic

achievement in Reading Comprehension as measured by the Woodcock-McGrew-Werder Mini-Battery of Achievement vs. students who work with a paper-based version of the Standardized Test Preparation intervention?

Students' reading comprehension levels were measured by the Woodcock-McGrew-Werder Mini-Battery of Achievement. Reading comprehension was defined as the Standard Score as determined by the Woodcock-McGrew-Werder Mini-Battery of Achievement.

Participants were 18 fourth grade students from Highland-Pacific Elementary school in the San Bernardino City Unified School District, who completed both the pre-test and post-test assessments. The data analysis was conducted using SPSS for Windows software.

CHAPTER FOUR

RESULTS AND DISCUSSION

Introduction

This study was designed to determine if using a web-based reading comprehension intervention was more effective in raising standardized test scores in reading than a traditional paper-based version. Nineteen students originally participated in the study. One student that was involved in the study moved and did not take the posttest. Therefore, this student's scored was removed from the analysis. Eighteen students did complete all parts of the study. There were eighteen students in both the experimental and control groups.

The Woodcock-McGrew-Werder Mini-Battery of Achievement was used to examine the study's research question. SPSS for Windows was used to analyze the data collected. An independent samples t-test was conducted to investigate the research question.

Presentation of the Findings

Data Analysis

Once the data had been collected from the Woodcock-McGrew-Werder Mini-Battery of Achievement, it was analyzed. The pre-test Standard Scores were compared to

the post-test Standard Scores to evaluate progress in student reading comprehension as a result of implementation of either the web-based or the paper-based Standardized Test Preparation Intervention for Reading Comprehension. The difference of the pre-test and post-test Standard Scores was then calculated for each student.

For purposes of addressing the research question, the calculated differences of the Standard Scores were then divided according to the dependent variables of web-based intervention and paper-based intervention.

Appendix A shows the experimental group's "value added" growth scores determined from the pretest and posttest Standard Scores of the Woodcock-McGrew-Werder Mini-Battery of Achievement. Appendix B shows the control group's "value added" growth scores determined from the pretest and posttest Standard Scores.

An independent samples t-test was conducted to compare the test scores of the experimental group, those who worked with the web-based reading comprehension intervention, with the control group in order to determine if the two groups' test score means were significantly different. Table 1 shows the comparison of means of the experimental and control scores. The test results were

t = .543, df = 34. As can be seen in the table, there was not a significant difference between the independent samples from the experimental group and the control group at the end of the study.

Table 1. Comparison of Means of Experimental and Control Scores

Groups	N	Mean	df	Mean Difference	t	Sig.2-tailed
Control	18	5.1111	34			
				1.38889	.543	.591*
Experimental	18	3.7222	33.919			

*Not Significant

Hypothesis Rejected

An independent samples t-test was conducted to evaluate the hypothesis that there would be a difference between learning as measured by pre-post tests between the treatment and control groups. The test results were not significant. The results were t = .543, df = 34. The students in the two study groups did not perform differently.

The hypothesis was that 4th grade elementary students would perform better in reading comprehension as measured by the Woodcock-McGrew-Werder Mini-Battery of Achievement

Reading Test using a web-based Standardized Test Preparation intervention than students who work with a paper-based version of the Standardized Test Preparation intervention.

After completing the t-test, the data of the study showed that there was not a significant difference in the test scores between the students who used the web-based reading comprehension intervention and the students who used a paper-based version of the same intervention at a 0.05 level of significance. The hypothesis was rejected (see Table 1).

In this research, the two reading comprehension intervention methods, the web-based and the paper-based reading comprehension interventions, were equally effective in helping 4th grade elementary students perform better in reading comprehension. There is a variety of possibilities for this phenomenon. First, the students in both cases were using an intervention that directly targeted the development of reading comprehension skills. Having extra practice and more time in reading comprehension can have a great affect on the development of the student's skills in reading comprehension.

Another possible reason is that the students were excited to participate with this type of intervention. The

students wanted to work with the intervention regardless of whether it was the paper-based version or the web-based version. It was different. It was out of the normal regimen of the classroom day.

Another possible reason for both methods being equally effective in developing reading comprehension skills is the built in use of scaffolding. In both cases, all the students started working with stories and comprehension questions that were at a first grade reading level. This allowed the student to have success regardless of what reading level they were currently at in the classroom.

Summary

This study was designed to determine if using a web-based reading comprehension intervention was more effective in raising standardized test scores in reading comprehension than a traditional paper-based version. An independent samples t-test was conducted to evaluate the hypothesis that there would be a difference between learning as measured by pretests and posttests between the treatment and control groups. The data of the study showed that there was not a significant difference in the test scores at a 0.05 level of significance. Based on these

results, it can be determined that there was not a significant difference in increased reading comprehension performance between students who used a web-based Standardized Test Preparation intervention and the students who worked with a paper-based version of the Standardized Test Preparation intervention. These results refute the hypothesis.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

Introduction

The development of reading comprehension skills has become a primary concern at the local, state and federal levels. The federal program No Child Left Behind (NCLB) has made student academic achievement in reading a top priority. In an attempt to provide a resource to practice the multiple-choice test format, to help students improve their reading comprehension skills, and to help students perform better and raise their standardized test scores, a web-based Standardized Test Preparation Intervention for Reading Comprehension was designed and developed to an initial stage of completion. The research project was an attempt to evaluate this intervention in terms of how effective the intervention helped the students develop reading comprehension skills and to perform better on standardized tests.

The purpose of this study was to investigate whether the web-based Standardized Test Preparation Intervention for Reading Comprehension was more effective than a paper-based version of the intervention to remedy, to some degree, the low reading comprehension levels of students

within a target classroom of elementary school students in the San Bernardino City Unified School District.

The study was conducted at Highland-Pacific Elementary during the 2006-2007 school year. The participants consisted of 18 fourth grade students. These students were randomly selected for control and experimental groups. The control group worked with the paper-based version of the Standardized Test Preparation Intervention for Reading Comprehension. The experimental group worked with the web-based Standardized Test Preparation Intervention for Reading Comprehension.

After an initial reading comprehension assessment was given, the experimental group practiced with the web-based Standardized Test Preparation Intervention for Reading Comprehension program while the control group practiced with a paper based version of the same intervention. After a three-week implementation period, the students were assessed again. The two groups were then reversed. For the next three weeks, the control group became the experimental group and practiced with the web-based intervention while the experimental group became the control group and practiced with the paper-based intervention. A final assessment was then given to determine whether students had improved their reading

comprehension because of participating with the intervention.

The results of these tests showed, at the 0.05 level, there was not a significant statistical difference in the value added means as related to the pretest and posttest scores. This suggests that the web-based reading comprehension intervention was not more effective at raising standardized test scores than the paper-based version. The hypothesis, therefore, was rejected.

Conclusions

Based on an analysis of the study findings, the following conclusions can be extracted from the project:

1. There was not a significant difference in reading comprehension achievement of the control group and the experimental group. The students who used the web-based intervention did not perform better in reading comprehension than the students who used the paper-based version of the same intervention.
2. One unintended, yet positive consequence was that both methods, the web-based and the paper-based reading comprehension interventions, were equally effective in helping 4th grade

elementary students perform better in reading comprehension. The value added means of the control and experimental groups showed that using the reading comprehension intervention produced higher scores in both groups. The control group improved by an average of 3.722 points, while the experimental group improved by an average of 5.111 points. This is equivalent to nearly half a school year of growth.

3. Anecdotal observations during the research period showed that the web-based intervention was extremely teacher friendly. The students were constantly on task and never had to be prompted to get back to work. The paper-based version, on the other hand, was much less teacher friendly. The teacher had to correct each of the tests by hand. While the teacher was correcting the students' tests, the students were waiting and aimless. This led to talking and the drawing of other students off task.

Recommendations

Further research is needed on the use of computers and the Internet to improve reading comprehension skills.

It is important to determine the best methods of using computers to help students succeed in reading comprehension, as well as achieve higher standardized test scores. The following recommendations resulting from the project are made for future studies.

1. In future studies, it is recommended that a more appropriate multiple-choice, standardized reading comprehension test such as the California Achievement Test (CAT6) be used to evaluate and determine the students' growth rather than the Woodcock-McGrew-Werder Mini-Battery of Achievement (MBA) that was used in this study.
2. It is recommended that further research should be conducted over a longer period of time. Three weeks was a very short amount of time to validate the use of the intervention. A longer research period would improve the accuracy of the results.
3. Further studies might address the impact of the duration of the reading block. For example, a sixty or ninety minute block of uninterrupted reading time would possibly have a greater impact on the improvement of reading

comprehension levels than the 30 to 40 minutes a day used in this study.

4. It is recommended that future research should be conducted using a larger sample size. A larger sample size would generate results that are more general.
5. Further research should be conducted to determine how the web-based reading comprehension intervention affects students' attitudes toward reading comprehension.

Summary

Chapter 5 discussed the conclusions that were determined from this study and recommendations for future research studies. One idea seems reasonably clear; more information is needed. While there is an implication that the Standardized Test Preparation Intervention for Reading Comprehension may increase reading comprehension levels, the compounding variables need to be considered.

It is unknown if the limited sample size or the short duration of the study may have skewed the data one way or another. In addition, the use of the Woodcock-McGrew-Werder Mini-Battery of Achievement (MBA)

to determine growth in reading may not be the best indicator of growth for this study.

Therefore, while this study provides some information as to the effectiveness of the Standardized Test Preparation Intervention for Reading Comprehension, further research is needed. In order to answer the question as to whether the intervention can be an effective means to improve reading comprehension skills, the stated recommendations need to be considered when future research is conducted.

It is important to determine the best methods of using computers to help students succeed in reading comprehension, as well as achieve higher standardized test scores. These new ideas must be studied in order to better prepare teachers for what they will experience in the classroom. As society moves toward standardized testing as high-stakes testing, researchers and teachers need to identify new and greater resources to help students meet the ever-growing demands of standardized testing. If these resources are not put into place, then teachers will not be able to fully prepare their students for the ever-increasing demands and rigors that they will face in the years to come.

APPENDIX A

STANDARD SCORES AND "VALUE ADDED" SCORES

EXPERIMENTAL GROUP

Student	Pretest	Post Test	Value Added
1	77	94	+17
2	94	105	+9
3	95	95	0
4	105	105	0
5	107	120	+13
6	108	112	+4
7	113	116	+3
8	115	134	+18
9	120	111	-9
10	128	123	-5
11	120	120	0
12	119	118	-1
13	112	120	+8
14	109	122	+13
15	100	99	-1
16	100	107	+7
17	99	104	+5
18	79	90	+11

APPENDIX B

STANDARD SCORES AND "VALUE ADDED" SCORES

CONTROL GROUP

Student	Pretest	Post Test	Value Added
1	79	79	0
2	105	119	+14
3	92	100	+8
4	95	99	+4
5	98	120	+22
6	99	100	+1
7	105	109	+4
8	108	112	+4
9	111	128	+17
10	134	122	-12
11	120	119	-1
12	116	115	-1
13	112	112	0
14	111	110	-1
15	105	101	-4
16	105	107	+2
17	95	101	+6
18	94	98	+4

APPENDIX C

INFORMED CONSENT

CALIFORNIA STATE UNIVERSITY
SAN BERNARDINO
5500 University Parkway, San Bernardino, CA 92407-2397

COLLEGE OF EDUCATION
Department of Science, Mathematics, and Technology Education
(909) 880-5290
fax: (909) 880-7522

Study of Using Comprehension Software to Improve Student Academic Achievement as Measured by Standardized Tests
INFORMED CONSENT

The study in which your child is being asked to participate is designed to investigate how a web-based reading comprehension intervention could improve student academic achievement on standardized tests. This study is being conducted by Mr. Andy Kubitza under the supervision of Dr. Brian Newberry, PROFESSOR OF COLLEGE OF EDUCATION, DEPARTMENT OF SCIENCE, MATH, AND TECHNOLOGY. This study has been approved by the Institutional Review Board, California State University, San Bernardino.

In this study your child will be asked to practice with a web-based intervention for Reading Comprehension. Your child will begin by reading a passage and then answer ten comprehension questions that go along with the story. When the child has completed answering these questions they will select a submit button to have the quiz graded. The task should take about 30 to 40 minutes to complete. All of your responses will be held in the strictest of confidence by the researchers. Your child's name will not be reported. All data will be reported in group form only. You may receive the group results of this study upon completion at January 31, 2008 at the following location Highland Pacific Elementary School, 3340 Pacific St. Highland, Ca. 92346.

Your child's participation in this study is totally voluntary. He/She is free to withdraw at any time during this study without penalty. There is not any foreseeable immediate or long range risks involved in this study. Students will not be asked to do anything outside of the normal Language Arts curriculum and teaching practices. The intended benefit is improved reading comprehension skills and better preparation to take a standardized test.

If you have any questions or concerns about this study, please feel free to contact Professor Brian Newberry Ph.D at (909) 880-7630.

I acknowledge that I have been informed of, and that I understand, the nature and purpose of this study, and I freely give consent to my minor child to participate. I also acknowledge that I am at least 18 YEARS OF AGE.

Signature: _____ Date: _____
Participant/Parent/Guardian

CALIFORNIA STATE UNIVERSITY, SAN BERNARDINO
INSTITUTIONAL REVIEW BOARD COMMITTEE
APPROVED 03/08/07 VOID AFTER 03/07/08
IRB# 06075 CHAIR Jamal A Kisher

APPENDIX D

CHILD ASSENT

CALIFORNIA STATE UNIVERSITY
SAN BERNARDINO
5500 University Parkway, San Bernardino, CA 92407-2397

COLLEGE OF EDUCATION
Department of Science,
Mathematics, and
Technology Education
(909) 880-5290
fax: (909) 880-7522

Study of Using Comprehension Software to Improve Student Academic Achievement as Measured by Standardized Tests

Child Assent

You are being asked to be part of a study. I want to see if I can help you do better in reading comprehension. You will read stories on paper and answer questions. You will also get to use a computer to read the same stories with the same questions. I want to see which one can help you do better and score higher on a reading test.
In this study you will get to practice with a computer. You will start by reading a story. Then you will answer ten questions. These check to see if you understand the story. You will click the button next to the answer you think is right. When you are done, you will click on the on the Submit button at the bottom of the page. This will grade the quiz. The work should take about 30 to 40 minutes to do. Nobody will ever see your answers. Nobody will ever see your name. I am going to write how the class did. You can get the results when I am done. I will be done January 31, 2008.

You do not have to be in the study. You do not have to answer any questions. You can quit the study at any time. You will not get into trouble. Please don't tell anybody the answers because they may get to do it next. There is nothing unsafe to be worried about. In fact, I hope you will have fun. You do not have to do anything that we don't normally do in class. You are going to read. Then do the questions just like we do everyday. My goal is to help you better understand what you read. I also want you ready to score well on the big CAT6 test at the end of the year.

Signature: _____ Date: _____
Student/Child

CALIFORNIA STATE UNIVERSITY, SAN BERNARDINO
INSTITUTIONAL REVIEW BOARD COMMITTEE
APPROVED 03/08/07 VOID AFTER 03/07/08
IRB# 06075 CHAIR Samuel A. Kleiner

REFERENCES

Allington, R.L. (2001). *What really matters for struggling readers: Designing research-based programs*. New York: Longman.

Alyousef, H.S. (2005). Teaching reading comprehension to ESL/EFL learners. *The Reading Matrix, 5*(2), 143-154. Retrieved June 30, 2007, from http://www.readingmatrix.com/articles/alyousef/article.pdf

Armbruster, B.B., Lehr, F., & Osborn, J. (2003). Put reading first: The research building blocks of reading instruction. Retrieved May 30, 2007 from The National Institute for Literacy Web site: http://www.nifl.gov/partnershipforreading/publications/Cierra.pdf

Ashby, N., Ed. (2007, February). Spellings celebrates law's fifth anniversary, Displays new logo. *The Achiever, 6*(2), 2.. Retrieved July 15, 2007, from U.S. Department of Education Web site: http://www.ed.gov/news/newsletters/achiever/2007/0207.pdf

Atkinson, R. (1969). Computer-assisted learning in action. *Proceedings of the National Academy of Sciences , 63*, 588-594. Retrieved June 28, 2007, from http://www.pnas.org/cgi/reprint/63/3/588.pdf

Barton, Paul E. (2004). Unfinished business: More measured approaches in standards-based reform. Retrieved April 28, 2007, from Educational Testing Service Web site: http://www.ets.org/Media/Education_Topics/pdf/unfinbusiness.pdf

Benjamin, L.T. (1988). A history of teaching machines. *The American Psychologist, 43*(9), 703-712. Retrieved May 7, 2005 from http://campus.dyc.edu/~drwaltz/FoundLearnTheory/FLT_readings/HistTeachMach.htm

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. Retrieved April 29, 2007, from Phi Delta Kappa International Web site: http://www.pdkintl.org/kappan/kbla9810.htm

Bond, L. A. (1996). Norm- and criterion-referenced testing. *Practical Assessment, Research & Evaluation, 5*(2). Retrieved May 7, 2007, from http://pareonline.net/getvn.asp?v=5&n=2

Briggs, L.J. (1958). Two self-instructional devices. In A.A. Lumsdaine & Robert Glaser, *Teaching Machines and Programmed Learning: A source book* (pp. 299-304). Washington, DC: National Education Association. Retrieved June 16, 2007, from http://www.openlibrary.org/details/teachingmachines009134mbp

Buck, G.H. (2000). The Y1K situation: Gerbert's instructional devices, their influence, and possible parallels to the present [Electronic version]. (ERIC Document Reproduction Service No.ED443403).

Burton, S.J., Sudweeks, R.R., Merrill, P.F., & Wood, B. (1991). How to prepare better multiple-choice test items: Guidelines for university faculty. Retrieved May 7, 2007 from http://testing.byu.edu/info/handbooks/betteritems.pdf

California Department of Education. (2006). Expectations and examples of classroom practices. Retrieved June 30, 2007, from http://www.cde.ca.gov/pd/ca/rl/teachreadpart1a.asp

California Department of Education. (2007a). California standardized testing and reporting (STAR) program. Retrieved June 30, 2007, from http://star.cde.ca.gov/star2007/viewreport.asp

California Department of Education. (2007b). English-language arts content standards. Retrieved June 30, 2007, from http://www.cde.ca.gov/be/st/ss/engmain.asp

Camacho, D., & Cook, V. (2007). Standardized testing: Does it measure student preparation for college & work? [Electronic version]. (ERIC Document Reproduction Service No.ED495251).

Casas, M. (2002). The use of Skinnerian teaching machines and programmed instruction in the United States 1960-1970 [Electronic version]. (ERIC Document Reproduction Service No.ED469942)

Casas, M. (2003). Crisis for social efficiency in the pedagogical arenas of the early twentieth century and the early 1960s [Electronic version]. (ERIC Document Reproduction Service No.ED479761).

Clark, D. (1999). B.F. Skinner. Retrieved on May 6, 2005. http://www.nwlink.com/~donclark/hrd/history/skinner.html

Cope, E.W., & Suppes, P. (2002). Gifted students' individual differences in distance learning computer-based calculus and linear algebra [Electronic version]. *Instructional Science, 30*(2): 79-110. Retrieved June 29, 2007, from Stanford University, Collected Works of Patrick Suppes Web site: http://suppes-corpus.stanford.edu/webpages/article.html?id=390

Cotton, K. (1991). Computer assisted instruction (SIRS Close-up No. 10). Portland, OR: Northwest Regional Educational Laboratory. Retrieved June 24, 2007, from http://www.nwrel.org/scpd/sirs/5/cu10.html

Crocker, L., & Zieky, M. (1995). Joint conference on standard setting for large scale assessments (Washington D.C., October 5-7, 1994) [Electronic version]. (ERIC Document Reproduction Service No.ED403325).

Diaz-Rico, L., & Sandlin, R. (1995). *Study guide for Woolfolk educational psychology, 6th Edition.* Boston: Allyn & Bacon.

Duke, N.K. (2004). Strategies for building comprehension of informational text. Retrieved July 3, 2007, from http://www.prel.org/programs/rel/comprehensionforum/duke.pdf

Duke, N.K., & Pearson, P.D. (2001). Developing comprehension in the primary grades. Retrieved July 3, 2007, from http://www.ciera.org/library/presos/2001/2001IRA/ira01ddp.pdf

Ediger, M. (1999). Psychological foundations in teaching mathematics [Electronic version]. (ERIC Document Reproduction Service No.ED431606).

Ediger, M. (2000). Assessing student progress in spelling. [Electronic version]. (ERIC Document Reproduction Service No. ED439441).

Fernald, P.S., & Jordan, E. A., (1991). Programmed instruction versus standard text in introductory psychology. *Teaching of Psycology, 18*(4), 205-211.

Finn Jr., C.E., Julian, L., & Petrilli, M.J. (2006). The state of state standards 2006. Retrieved July 16, 2007, from Thomas B. Fordham Foundation Web site: http://www.edexcellence.net/doc/State%20of%20State%20Standards2006FINAL.pdf

Fremer, J., & Wall, J.E. (2003). Why use tests and assessments? [Electronic version]. (ERIC Document Reproduction Service No. ED480036).

Glaser, R. (1960). Christmas past, present and future: a review and preview. In A.A. Lumsdaine and Robert Glaser (Eds.), *Teaching Machines and Programmed Learning: A source book* (pp. 23-31). Washington, DC: National Education Association. Retrieved June 16, 2007, from http://www.openlibrary.org/details/teachingmachines009134mbp

Gleason, M.S. (1928). Educational game apparatus [Electronic version]. Patent No. 1,660,501. Retrieved June 16, 2007, from United States Patent and Trademark Office Web site: http://patimg1.uspto.gov/.piw?PageNum=3&docid=US001660501&IDKey=BF7CB01F0295%0D%0A&HomeUrl=http%3A%2F%2Fpatft.uspto.gov%2Fnetacgi%2Fnph-Parser%3FSect1%3DPTO1%2526Sect2%3DHITOFF%2526d%3DPALL%2526p%3D1%2526u%3D%25252Fnetahtml%25252FPTO%25252Fsrchnum.htm%2526r%3D1%2526f%3DG%2526l%3D50%2526s1%3D1%2C660%2C501.PN.%2526OS%3DPN%2F1%2C660%2C501%2526RS%3DPN%2F1%2C660%2C501

Green, B. (1996). Setting performance standards: Content, goals and individual differences. Retrieved July 5, 2007, from Educational Testing Service Web site: http://www.ets.org/Media/Research/pdf/PICANG2.pdf

Hanson, A.M. (2006). No Child Left Behind: High stakes testing and teacher burnout in urban elementary schools [Electronic version]. (ERIC Document Reproduction Service No. ED493443).

Hayes, B.G. (1999). Where's the data: Is multimedia instruction effective in training counselors? *Journal of Technology in Counseling 1*(1). Retrieved June 28, 2007, from http://jtc.colstate.edu/vol1_1/multimedia.htm

Hess, F. M. (2004, August). Reforming reform: An excerpt from Common Sense Reform. *Harvard Graduate School of Education, HGSE News*. Retrieved July 15, 2007, from http://gseweb.harvard.edu/news/features/hess08012004.html

Johnson, R.A., & Howard, C.A. (2003). The effects of the accelerated reader program on the reading comprehension of pupils in grades three, four, and five. *The Reading Matrix 3*(3), 87-96. Retrieved June 30, 2007, from http://www.readingmatrix.com/articles/johnson_howard/article.pdf

Kehoe, J. (1995). Writing multiple-choice test items. *Practical Assessment, Research & Evaluation*, 4(9). Retrieved May 7, 2007 from, http://pareonline.net/getvn.asp?v=4&n=9

Kendall, J.S., & Marzano, R.J. (2004). *Content knowledge: A compendium of standards and benchmarks for K-12 education*. Aurora, CO: Mid-continent Research for Education and Learning. Online database: http://www.mcrel.org/standards-benchmarks/

Kendall, J.S., Richardson, A.T., & Ryan, S.E. (2005). *The systematic identification of performance standards*. Aurora, CO: Mid-continent Research for Education and Learning. Retrieved July 16, 2007, from http://www.mcrel.org/PDF/Standards/5041TG_PerfStnds.pdf

Kim, J. (1992). Three approaches for the integration of teaching, testing and learning [Electronic version]. (ERIC Document Reproduction Service No. ED354250).

Kingsbury, G.G., Olson, A., Cronin, J., Hauser, C., & Houser, R. (2003). The state of state standards: Research investigating proficiency levels in fourteen states. Retrieved July 16, 2007, from Northwest Evaluation Association Web site: http://www.nwea.org/research/getreport.asp?ReportID=5

Learning Point Associates. (2005). Implementing the No Child Left Behind Act: Strategies to improve high schools. NCREL Quick Key Number 9. Action Guide [Electronic version]. (ERIC Document Reproduction Service No. ED489529).

Lumsdaine, A.A. (1959). Teaching machines: An introductory overview. In A.A. Lumsdaine and Robert Glaser (Eds.), *Teaching Machines and Programmed Learning: A source book* (pp. 5-22). Washington, DC: National Education Association. Retrieved June 16, 2007, from http://www.openlibrary.org/details/teachingmachines009134mbp

McMillan, James H. (2005). The impact of high-stakes test results on teachers' instructional and classroom practices [Electronic version]. (ERIC Document Reproduction Service No. ED490648).

McNeil, S. (2004) A hypertext history of instructional design - the 1920s. Retrieved May 7, 2005, from the University of Houston, College of Education Web site: http://www.coe.uh.edu/courses/cuin6373/idhistory/pressey.html

Morra, J. & Tracey, D.H. (2006). The impact of explicit fluency instruction [Electronic version]. (ERIC Document Reproduction Service No. ED491520).

National Assessment of Educational Progress. (2005) The NAEP reading achievement levels by grade. Retrieved April 28, 2007, from The National Center for Education Statistics Web site: http://nces.ed.gov/nationsreportcard/reading/achieveall.asp

National Center on Education and the Economy. (2007). Tough choices or tough times: The report of the new commission on the skills of the American workforce - executive summary. Retrieved July 15, 2007, from National Center on Education and the Economy Web site: http://www.skillscommission.org/pdf/exec_sum/ToughChoices_EXECSUM.pdf

National Reading Panel. (2000a). Report of the national reading panel: Teaching children to read. Retrieved July 1, 2007, from National Institute of Child Health and Human Development Web site: http://www.nichd.nih.gov/publications/nrp/upload/smallbook_pdf.pdf

National Reading Panel. (2000b). Report of the national reading panel: Teaching children to read. Reports of the Subgroups. Retrieved July 1, 2007, from National Institute of Child Health and Human Development Web site: http://www.nichd.nih.gov/publications/nrp/upload/report_pdf.pdf

Ohio Department of Education. (2007). Academic content standards frequently asked questions. Retrieved July 15, 2007, from http://www.ode.state.oh.us/GD/Templates/Pages/ODE/ODEDetail.aspx?Page=3&TopicRelationID=305&ContentID=1929&Content=22326

Paige, R. (2003). *Education in America: The complacency must end.* Retrieved July 15, 2007, from U.S. Department of Education Web site: http://www.ed.gov/news/speeches/2003/09/09242003.html

Petterway, A.L., & Kritsonis, W.A. (2006). A national perspective: A mixed-methods analysis of the impact of high stakes testing on English language learners in major urban high schools in Texas [Electronic version]. National Journal for Publishing and Mentoring Doctoral Student Research v1 n1 2006. (ERIC Document Reproduction Service No. ED491980).

Phelps, R. P. (2006). Characteristics of an effective student testing system [Electronic version]. Educational Horizons, 85(1), 19-29. (ERIC Document Reproduction Service No. EJ750639).

Pressey, S.L. (1926). A simple apparatus which gives tests and scores - and teaches. In A.A. Lumsdaine and Robert Glaser (Eds.), *Teaching Machines and Programmed Learning: A source book.* (pp. 35-41). Washington, DC: National Education Association. Retrieved June 16, 2007, from http://www.openlibrary.org/details/teachingmachines009134mbp

Pressey, S.L. (1927). A machine for automatic teaching of drill material. In A.A. Lumsdaine and Robert Glaser (Eds.), *Teaching Machines and Programmed Learning: A source book.* (pp. 42-46). Washington, DC: National Education Association. Retrieved June 16, 2007, from http://www.openlibrary.org/details/teachingmachines009134mbp

Pressley, M. (2001). Comprehension instruction: What makes sense now, what might make sense soon. *Reading Online, 5*(2). Retrieved July 3, 2007, from http://www.readingonline.org/articles/handbook/pressley/index.html

Rasinski, T.V. (2003). *The fluent reader: Oral reading startegies for building word recognition, fluency and comprehension.* New York, NY: Scholastic.

Reisman, S., & Carr, W.A. (1991). Perspectives on multimedia systems in education. *IBM Systems Journal, 30*(3), 280-295. Retrieved June 24, 2007 from http://www.research.ibm.com/journal/sj/303/ibmsj3003E.pdf

Rosenthal, T., & Suppes, P. (2002). Gifted students' individual differences in computer-based C programming course. Retrieved June 29, 2007, from Stanford University, Collected Works of Patrick Suppes Web site: http://suppes-corpus.stanford.edu/article.html?id=402

Rothkopf, E.Z. (1958). Some research problems in the design of materials and devices for automated teaching. In A.A. Lumsdaine and Robert Glaser (Eds.), *Teaching Machines and Programmed Learning: A source book.* (pp. 318-328). Washington, DC: National Education Association. Retrieved June 16, 2007, from http://www.openlibrary.org/details/teachingmachines009134mbp

Shaffer, D.R. (1994). *Social and personality development - 3rd ed.* California: Brooks/Cole Publishing Company.

Singhal, M. (2001). Reading proficiency, reading strategies, metacognitive awareness and L2 readers. *The Reading Matrix, 1*(1). Retrieved July 4, 2007, from http://www.readingmatrix.com/articles/singhal/article.pdf

Skinner, B.F. (1948). 'Superstition' in the pigeon. Retrieved May 7, 2005 from Classics in the History of Psychology Web site: http://psychclassics.yorku.ca/Skinner/Pigeon

Skinner, B.F. (1955). Teaching machine [Electronic version]. Patent No. 2,846,779. Patented Aug. 12, 1958. Retrieved June, 16, 2007, from United States Patent and Trademark Office Web site: http://patimg2.uspto.gov/.piw?PageNum=5&docid=US002846779&IDKey=E812387C1010%0D%0A&HomeUrl=http%3A%2F%2Fpatft.uspto.gov%2Fnetacgi%2Fnph-Parser%3FSect1%3DPTO1%2526Sect2%3DHITOFF%2526d%3DPALL%2526p%3D1%2526u%3D%25252Fnetahtml%25252FPTO%25252Fsrchnum.htm%2526r%3D1%2526f%3DG%25261%3D50%2526s1%3D2%2C846%2C779.PN.%2526OS%3DPN%2F2%2C846%2C779%2526RS%3DPN%2F2%2C846%2C779

Skinner, B.F. (1958). Teaching machines. Retrieved May 6, 2005, from B. F. Skinner Foundation Web site: http://www.bfskinner.org/teachingmachines1958.pdf

Skinner, B.F. (1963). A brief survey of operant behavior. Retrieved May 6, 2005 from B.F. Skinner Foundation Web site: http://www.bfskinner.org/Operant.asp

Slavin, R. E. (1991). *Educational psychology: Theory into practice 3rd Ed.* Boston, MA: Allyn and Bacon.

Smith, R. (1999). Behaviorism. Retrieved May 4, 2005 from Georgia State University Web site: http://www2.gsu.edu/~mstswh/courses/it7000/papers/behavior.htm

Snow, C., Burns, S., & Griffen, P. (1998). *Preventing reading difficulties in young children*. Washington, D.C., National Academy Press.

Soe, K., Koki, S., & Chang, J.M. (2000). Effect of computer-assisted instruction (CAI) on reading achievement: A meta-analysis. Retrieved June 24, 2007, from Pacific Resources in Education and Learning Web site: http://www.prel.org/products/Products/effect-cai.htm

Stephens, A.L. (1953). Certain Special Factors Involved in the Law of Effect. In A.A. Lumsdaine and Robert Glaser (Eds.), *Teaching Machines and Programmed Learning: A source book.* (pp. 89-93). Washington, DC: National Education Association. Retrieved June 16, 2007, from http://www.openlibrary.org/details/teachingmachines009134mbp

Sternberg, B.J., Kaplan, K.A., & Borck, J.E. (2007). New directions in research: Enhancing adolescent literacy achievement through integration of technology in the classroom. *Reading Research Quarterly, 42*(3), 416-420. Retrieved July 4, 2007, from http://www.reading.org/Library/Retrieve.cfm?D=10.1598/RRQ.42.3.6&F=RRQ-42-3-Sternberg.html

Stewart, D. (1975). *Instruction as a humanizing science: Volume II - a behavioral learning systems approach to instruction: Analysis and synthesis*. California: SLATE Services.

Sulaiman, J., & Dwyer, F. (2002). The effect of varied instructional text design on the achievement of different educational objectives. *International Journal of Instructional Media, 29*(2) 215-223.

Suppes, P. (1979). Current trends in computer-assisted instruction. In M.C. Yovits (Ed.), *Advances in Computers, 18*, 173-229. New York: Academic Press. Retrieved June 28, 2007, from Stanford University, Collected Works of Patrick Suppes Web site: http://suppes-corpus.stanford.edu/article.html?id=203

Suppes, P. (1988). Computer-assisted instruction. Derick Unwin and Ray McAleese (Eds.), *The Encyclopaedia of Educational Media Communications and Technology (2nd Edition)*, New York: Greenwood Press. Retrieved June 28, 2007, from Stanford University, Collected Works of Patrick Suppes Web site: http://suppes-corpus.stanford.edu/article.html?id=278

Suppes, P., & Fortune, R.F. (1985). Computer assisted instruction: Possibilities and problems. *NASSP Bulletin, 69*, pp. 30-34. Retrieved June 24, 2007, from Stanford University, Collected Works of Patrick Suppes Web site: http://suppes-corpus.stanford.edu/article.html?id=257-1

Togut, T.D. (2004). High stakes testing: Educational barometer for success, or false prognosticator for failure. *The Beacon - Journal of Special Education Law and Practice, 2*(3). Retrieved April 28, 2007, from http://www.harborhouselaw.com/articles/highstakes.togut.htm

Traynor, P.L. (2003). Effects of computer-assisted instruction on different learners. *Journal of Instructional Psychology, 30*. Retrieved June 24, 2007, from http://findarticles.com/p/articles/mi_m0FCG/is_2_30/ai_105478983

Troutner, J.J. (1991). The historical evolution of educational software [Electronic version]. (ERIC Document Reproduction Service No. ED349936).

U.S. Department of Education. (2007). *Building on results: A blueprint for strengthening the "No Child Left Behind Act"* [Electronic version]. (ERIC Document Reproduction Service No. ED495309).

Vargas, J.S. (2005). Brief biography of B.F. Skinner. Retrieved May 5, 2005, from the B.F. Skinner Foundation Web site: http://www.bfskinner.org/bio.asp

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Watson, J.B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20,* 158-177. Retrieved May 6, 2005 from Classics in the History of Psychology Web site: http://psychclassics.yorku.ca/Watson/views.htm

Wenning, R., Herdman, P.A., Smith, N., McMahon, N., & Washington, K. (2003). No Child Left Behind: Testing, reporting, and accountability [Electronic version]. (ERIC Document Reproduction Service No. ED480994).

Woolfolk, A.E. (1995). *Educational psychology* 6th Ed. Boston, MA: Allyn and Bacon.