1-1-2011

Online Formative Assessments as Predictors of Student Academic Success

Jacqueline L. Croteau
Walden University

<u>Walden University</u>

College of Education

This is to certify that the doctoral dissertation by

Jacqueline Croteau

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Carol Watson, Committee Chairperson, Education Faculty
Dr. Wade Smith, Committee Member, Education Faculty
Dr. Gerald Giraud, University Reviewer, Education Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2014

Abstract

Online Formative Assessments as Predictors of Student Academic Success

by

Jacqueline L. Croteau

MA, University of Phoenix, 1999

BS, Prescott College 1995

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Education Technology

Walden University

August 2014

Abstract

Increasingly, educational reform efforts are turning towards data-driven decision making strategies to help teachers improve instruction through skills-based instruction/content that is both measurable and aligned to common rigorous standards, such as the Common Core State Standards (CCSS). Examining the impact of a formative online assessment system on a summative measurement of student achievement may provide evidence that data-driven instructional platforms can impact student achievement and learning outcomes. Guided by the theoretical frameworks of Vygotsky and Dewey, along with the concepts of multiple intelligence, constructivism, and mastery learning, this study examined the relationship between student scores from an online formative assessment administered quarterly and an end-of-year summative evaluation. A stepwise multiple regression analyzed the predictive power of the iReady formative assessment program towards archived SAT-10 reading and mathematics data among Grades 1-4 students, before and after the iReady program was implemented (N = 339). The results showed a significant relationship between the iReady program and SAT-10, explaining 11.6% of the variance in SAT-10 scores. The study's intended audience is educators, school districts, and policy makers who are using the achievement data produced by formative assessments to improve results on measures of academic achievement, leading to positive social change.

Online Formative Assessments as Predictors of Student Academic Success

by

Jacqueline L. Croteau

MA, University of Phoenix1999

BS, Prescott College 1995

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Education Technology

Walden University

August 2014

Dedication

To my father whose familiar words I will never forget. His words…"You do not live in a vacuum" forever changed me and helped create the person I have become. To my mother who always helped me pick myself up, start over again and shake off the doubt that was strangling me. To my six siblings, thank you for being there for me every step of the way in bad times and good, you are my rock and the foundation of my life. To my Daughter, Ansley and my son, Cameron, I write this with a deep feeling of love for who you were and who you have become; you are truly amazing young people that I know will leave a positive impact on this world. To my granddaughter Kiki Bug, you are the future, and I thank God everyday he brought you into our lives. "I believe in angels, something good in everything I see. I have a dream, a song to sing". I love you all.

Acknowledgments

I thank you all every day for believing in me.

Table of Contents

List of Tables

Section 1: Introduction to the Study

Collecting, analyzing, and getting data into the hands of teachers is essential to promoting dialogue within the education community to improve school efficiency (Popham, 2008). Increased school efficiency ultimately leads to improved student achievement and thus leads to better system-wide performance (Popham, 2008). Teachers use informal formative assessments—one example of collected data—on a daily basis, because much of what teachers and students do in the classroom provides frequent opportunities for collecting evidence of students' understanding (Ruiz-Primo, 2011). Teachers use assessments to diagnose individual learning problems and adjust instructional strategies in order to meet predetermined learning objectives. The assessment process also provides feedback on individual, group, or whole class performance and development while reinforcing learning skills. Formative assessments are an important element in the classroom assessment process. They provide teachers with evidence of student learning. A critical step for teachers in the assessment process is to use the collected data to improve their instruction and move student learning towards expected outcomes (Popham, 2008).

The main purpose of formative assessment is to provide feedback that can be used to increase student content knowledge, skills, and understanding (Shute, 2008). Strategies for gaining meaningful feedback should be directive, response-specific, goal-oriented, ongoing, and delivered immediately (Shute, 2008). A formative assessment process, whether conducted through a computerized system or directly through teacher-driven assessments, provides students with either directive or facilitative feedback. Directive feedback tells students what needs to be fixed or revised and tends to be more precise than facilitative feedback. Facilitative feedback includes comments and suggestions to guide the student through the process of conceptualizing the learning on their own (Shute, 2008). In a technology-assisted instructional environment, feedback is comprised of messages, given either in verbal or written form. Feedback provided in a technology-assisted system provides student assistance, as well as informs the teacher of individual student progress or classroom progress towards established standards or expected results (Shute, 2008).

Additionally, formative assessments are effective measurements for examining instructional practices. Formative assessments are one of the most powerful tools in the teaching toolbox, but only if they are part of a comprehensive assessment system. Formative assessments provide evidence of teaching whereas; summative assessments provide evidence of a longer period of learning, for example, the summation of a particular unit, subject or year.

Technology provides teachers with dynamic articulation of assessment feedback rather than a static form provided through teacher-administered assessments (Pelligrino & Quellmatz, 2011). Technology allows educators to probe deeper into the broader spectrum of human learning (Pelligrino & Quellmatz, 2011). It offers teachers several data points related to instructional strategies and techniques, student learning, and student engagement in the learning process. Most importantly, it offers teachers several data points on the processes of how to apply acquired knowledge. Measuring student learning from computer-driven instruction includes learning results not easily observed during classroom instruction. The computer can capture student input and count the number of attempts needed to produce the correct answer. Computer-driven feedback gives the teacher data on each student's ability to solve problems and to use a variety of techniques and strategies during the learning process.

The knowledge gained from computerized feedback allows for differentiation and scaffolding of instruction that better meets the learning needs of all students (Pelligrino & Quellmatz, 2011). Immediate feedback actively engages learners in the discovery process, which is critical to learning (Pelligrino & Quellmatz, 2011). Active engagement in the content through the learning process promotes retention and the opportunity immediately to correct initial inaccurate response (Brosvic et al., 2002).

Not only do teachers and students benefit from the data-driven feedback, so do school reform efforts that focus on the use of formative and summative assessments to measure student achievement and teacher effectiveness. In this strong push to improve education, policy makers see advantages that computer-driven assessments bring towards accountability. Policy makers view computer-driven formative and summative assessments as a means of determining student achievement and evaluating the effectiveness of both teacher and school.

The process of data-driven decision-making endorses the use of information about the progress of learning for each class in school, or at least some form of measurement focusing on the performance level of individual schools in comparison to other similar schools (Kuiper & Schildkamp, 2010). An advantage for schools making decisions based on data is that data minimizes subjective judgments or interpretation. Data do not provide a sustainable basis for action, but rather, it is the interpretation of that data into actionable information that provides a foundation for the decision-making process (Kuiper & Schildkamp, 2010).

The No Child Left Behind Legislation (NCLB, 2002) is pressing states towards accountability measures. The expectation is that states will need to demonstrate academic progress and achievement in their student populations as a precursor to federal education funding. The theory behind high-stakes assessments suggests that teachers and students will work harder and more effectively and thereby increase academic achievement by tying negative consequences (e.g., public exposure, external takeover) to standardized scores (Berliner, Glass, & Nichols, 2012).

This pressure for high-stakes assessment of student learning is pushing states and districts towards developing a comprehensive assessment system that includes technology-driven formative and summative assessments. The need for accountability in the learning process, combined with the convergence of technological innovations and current developments in the cognitive sciences, is creating a new generation of measurements for assessing the process of teaching and learning (Tucker, 2009). However, he also cautions that for the most part school- and district-level monetary investments in technology have not led to fundamental changes in the process of assessing for learning. Instead, technology has simply made old approaches to testing more efficient (Tucker, 2009).

Another consequence for increased emphasis of accountability in the classroom is the way in which publishing companies and psychometric test makers have reacted with more and more software product all-promising to help improve student achievement. An additional consequence is the development of high-stakes tests that incorporate dynamic, interactive tasks instead of the standard multiple-choice items often used (Tucker, 2009). While the shift towards dynamic, interactive testing is occurring, teachers' reactions to high-stakes tests are shifting focus towards a narrower curriculum. This narrowing places higher emphasis on rote memorization of facts instead of the development of higher-order thinking skills (Tucker, 2009).

The National Assessment for Educational Progress (NAEP) is a nationally administered assessment for measuring academic success in grades K-12. Policy makers commonly refer to it as America's report card for education. NAEP officials asked the Educational Testing Service (ETS) to develop a pilot assessment incorporating interactive computer tasks (ICT) into the NAEP's design (Pelligrino & Quellmatz, 2011). The purpose for incorporating ICT is to support targeted instruction for students by diagnosing their strengths and weaknesses at different points during the problem-solving stage and thus creating a diagnostic profile (Zoanetti, 2010).

The Common Core State Standards (CCSS) represent the cumulative efforts of governors, education policy makers, school administrators, educators, and education stakeholders. Governors worked together with these groups to create a shared road map of what student learning expectations and progress towards academic achievement (CCSS Initiative, 2012). Thus far, 45 states have adopted the CCSS. It provides educators and schools to develop and plan curricula and assessments based on the need for specific and consistent outcomes in the majority of American schools.

Two consortiums, Partnership for Assessment of Readiness for College and Career Standards (PARCC) and the SMARTER Balanced Assessment Consortium (SBAC), are working together to develop common summative assessments that align teaching and learning with the expected results from CCSS. These new summative assessments will replace the high-stakes testing individual states created for NLCB requirements. States in each consortium will use one of the CCSS assessments developed in place of their state-mandated tests and graduation tests (CCSS Initiative, 2012). Addressing education to meet the CCSS will require educators to develop lessons and formative assessments encouraging students to develop a deeper understanding of the content being taught, as well as new skills. Measuring the depth of learning will require the use of formative assessments that reflect what current cognitive research indicates about how people learn (Tucker, 2009).

The successful implementation of CCSS will depend on a balanced use of classroom practices that align assessments to the standards, provide quality feedback, adjust instruction in response to assessment data, and involve students in the process of learning through their own actions and self-reflection (Pelligrino & Quellmatz, 2011). Hwang, McMaken, Porter, and Yang (2011) added that to implement the CCSS to fidelity education must be aware of variables that might limit implementation, such as, types of assessments, alignment of instructional materials to the CCSS, and teacher education. Hwang et al. (2011) also stressed the need to consider the beliefs and values of educators, policy makers and parents throughout implementation and testing of the CCSS.

Therefore, to address the hypothesis of online formative assessments' ability to predict student achievement on summative tests, Section 1 will discuss the purpose for providing evidence of online formative assessments' positive impact on teaching, learning, and measuring progress towards academic achievement. Section 1 includes the background for the study, problem statement, purpose, research questions, hypothesis, theoretical concepts and nature of the study to include definitions, scope and delimitations, limitations and significance of the study.

Background

Assessment, whether formative or summative, is not used solely as an accountability measure or for demonstrating adequate yearly progress (AYP). According to Popham (2008), formative assessments allow teachers to constantly judge instructional effectiveness, attend to learning consequences, and make the necessary changes based on the assessments' evidence. Popham (2008) categorized formative assessments into four distinct levels: (1) Teacher instructional adjustments, (2) Student learning-tactic adjustments, (3) Classroom climate shift, and (4) School-wide implementation.

According to Black and Dylan (2003) the value of assessments is not just in the assessments themselves, but also in the functions they serve. The fundamental value of the assessment process is the alignment of expected learning outcomes to the formative and summative assessment measures. This alignment needs to occur in order to prove overall effectiveness for measuring learning and preventing summative assessments from undermining formative work (Black & Dylan, 2003). It is this aligned relationship between formative and summative assessment that creates the potential for a productive synergy, a synergy that can result in improved student learning (Lam, 2013). Teachers recognize the potential for understanding student learning through assessment activities. It is the teacher's failure to explore the instructional power of assessments for promoting effective learning that is a cause for concern (Lam, 2013).

Summative assessments allow the instructor to drill into the learning process from the perspective of the final evaluation, allowing the teacher to determine if students meet the required learning objectives. Typically, summative evaluations concentrate on the completed learning process or expected learner outcomes. Summative evaluations are commonly used to determine success on completion of a project or process in order for the teacher to produce a final grade or report. As opposed to formative assessments that take place during learning, summative assessment's primary purpose is to provide data on the summation of the learning process or learning outcome.

In contrast to the business world, technology is not ubiquitous within the walls of today's classrooms (Watson, 2001). Even with many years of national policies and investments of time and money, technology is still considered novel as a tool for daily use in school pedagogy (Watson, 2001).

The influence of technology on society is also impacting the role of the classroom teacher. The ability to find an answer to just about any question one could pose is as close as the cell phone in every pocket. The role of teacher, therefore, is moving away from that of deliverer of knowledge to that of active facilitator of the learning process directed by content standards. While technology is dramatically widening the selection of tools available in the teachers' instructional toolbox, the value of these tools is often overshadowed by teachers' anxiety, frustration, and uncertainty about changing educational pedagogies.

Technology is a catalyst for change and is creating changes in teaching styles, instructional strategies, learning approaches, and access to information (Watson, 2001). As Watson (2001) suggested, research indicates that teachers are both threatened by change, and conversely not they are not impressed by changes technology brings to the learning process. However, teachers have a significant impact on whether or not to implement an educational innovation (Biemans, Gulikers, Van der Wel, & Wesselink, 2013).

One of the pedagogical changes occurring, as a result of technology, is the use of computer-based assessments of and for learning. Assessments play an important role in education. Taras (2010) wrote that assessments provide evidence with which to diagnose learning problems, support instruction, provide feedback on learning progress, and reinforce necessary learning skills. According to Taras (2010), formative assessments are an integral piece of a triumvirate of evaluating, learning, and teaching. Without an assessment, wrote Taras (2010), there is no development towards an expert level of learning.

Recent federal education policy has increased focus on accountability by developing national standards and mandating annual assessments to measure student academic progress (Linan-Thompson Murray, Roberts, Vaughn, Wanzek, and Woodruff, 2010). States have either created assessments, normed at the state level or have used commercially prepared, standardized achievement measures normed at the national level, such as the Iowa Test of Basic Skills and the SAT-10 (Linan-Thompson et al., 2010). These high-stakes summative evaluations are used to determine what, if any, measures are required should a student not show progress or meet expected grade-level learning outcomes. Student consequences for failing high-stakes assessments include not progressing to the next grade level or not receiving a diploma. Greene, Trivitt, and Winters (2009) found evidence that threats produced by the sanctions placed on high-stakes accountability measures presented causal factors. These causal factors demonstrated that high stakes testing could have a positive impact on academic achievement.

Assessments are useful in helping to understand the processes that occur during learning (formative) and how best to develop students' ability to move towards expected learning outcomes (summative; Black & Dylan, 2003). This interconnectedness between formative and summative assessments is useful in determining the success or failure of the learning process, project, or outcomes. Formative assessments can produce positive effects on students' mindfulness of the learning process, which translates into overall improved academic performance (Black & Dylan, 2003).

According to Harlen (2005), the main purpose for any measurement of learning is to understand the purpose for collecting evidence through assessments. How is the evidence interpreted and communicated to the end user? Is the data applicable to just for the teacher, student, or both? Harlen (2005) grouped summative assessments into two categories: internal and external. The purposes of internal assessments are to document learning through record keeping, curricular decisions and report learning progress to students and parents. Teacher judgment supports these internal assessments (Harlen, 2005). Internal assessment criteria are as informal as observations of behaviors or the learning process, whereas standardized assessments are more formal, being administered to determine a national, state, or district-wide aggregate of student performance. On the other hand external assessments are used for certification, school-wide progress monitoring, performance, and accountability. The data from these types of assessments are externally produced tests or evaluations (Harlen, 2005).

During the last decade, education has begun to feel the economic pressure and political ramifications of educational reform laws. Federal NCLB funding mandates that states develop accountability assessments. Each state wishing to receive federal funds is required to administer statewide assessments that provide data towards student year-over-year growth and current grade-level learning status. Federal funds are tied to implementation and development of these end-of-year assessments. As a result of the NCLB legislation, governors, state education officials, educators, and policymakers joined together to develop a common set of academic content standards, the Common Core State Standards (CCSS). These standards are intended to bring about changes from clearer common learning objectives, the inclusion of more rigorous content, and the combination of college and career readiness skills. Along with the implementation of the CCSS, comes the expectation that these standards will be measurable and provide a foundation for determining student academic gains.

Therefore, for the implementation of the CCSS to be successful, educators will have to use a variety of formal and informal assessments. They will need to develop formative assessments with more flexibility for diagnosing individual student's strengths and weaknesses. In addition, there will be a need for a variety of assessments that provide even more data on what might or might not be happening in the classroom. Gong, Marion and Perie (2009), suggest that, as a result, of federal policies such as IASA (Improving America's School Act) and NCLB, many school districts are moving resources towards finding interim or formative assessment systems that tie directly to informing instruction. As a result of these federal policies, formative assessments will carry responsibility for the measurement of academic progress towards end-of-year achievement. According to Gong et al., (2009) these measurements must be a formative assessment system that provides teachers with an in-depth understanding of what questions students miss and more importantly, why they missed the questions. Gong et al. (2009) contend that such a formative assessment system constitutes a better use of district resources because of the deeper level of information it provides. Teachers will have access to data not normally available on large-scale summative assessments or through daily classroom activities. The researchers ascertain that resources might be better spent on instructing teachers on how to use formative assessments effectively with a strong emphasis for using data to improve instruction (Gong et al., 2009).

Problem Statement

The legislative push to improve academic achievement is causing educators to look towards using data-driven decision-making rather than using decisions based solely on teacher intuition and instinct (Kuiper & Schildkamp, 2010). The states' adoption of the CCSS adds even more pressure on schools with the expectation that instruction not only includes rigorous content material and skills-based instruction, academic growth must also produce measurable results on a summative end-of-year assessment

Examining the predictability, if any that formative online assessment systems have on end-of-year measurements of learning, may provide evidence that computerized instructional platforms have an inherent ability to predict academic progress towards expected grade-level learning outcomes. The amount of research available on the use of formative assessments indicates there is a potential for new knowledge generated from a study examining this relationship. Such research provides essential information on (a) the impact of summative and online formative assessments on student motivation for learning, (b) summative assessments to drive student learning, (c) the use of online formative assessment data to plan and improve instruction and (d) the ability to harmonize the interconnectedness of summative and formative assessments to improve teaching and learning.

Purpose of the Study

There are studies examining student achievement on end-of-year summative evaluations. However, because of the difficulties of comparing formative assessments to summative assessments, studies examining the impact of one assessment on the other are difficult to find. The purpose of this study was not to compare the two types of assessments to each other, but to consider one's impact on the other. The lack of available research on this topic showed the potential to create new knowledge.

Research Questions and Hypotheses

Based on identified gaps in the research, the following three questions were addressed.

Research Questions

1. Is there a linear relationship between the results from online formative assessments and student achievement as measured on SAT-10 summative assessments?

   $H_0{}^1$: There is no linear relationship between the pre and post SAT-10 scores after the treatment of an online formative assessment system.

   $HA_1{}^1$: There is a significant relationship between pre and post SAT-10 scores after the treatment of on online formative assessment system.

2. Does an online formative assessment system have an effect on student achievement, as measured on SAT-10 summative assessments?

   $H_0{}^2$: Is there a difference from the use of an online formative assessment system on student achievement as measured by SAT-10.

   $HA_2{}^2$: The use of an online formative assessment system has significant effect on student achievement as measured by SAT-10.

3. Can online formative assessments be a predictor of student success on end-of-the year SAT-10 summative assessments?

$H_0{}^3$: The correlation of an online formative assessment system between student achievement is not a reliable predictor of student achievement as measured by SAT-10.

HA_3^3: The correlation of an online formative assessment system is a reliable predictor of student achievement as measured by SAT-10.

The method used to these hypotheses, was multiple linear regression analyses that examined SAT-10 scores from elementary students' enrolled in Grades 1 through 4. Students took the SAT-10 during a 3-year period prior to the start of this research; therefore, the study included these three years of archived data results.

Theoretical Base

Historical and theoretical were the two foundations found to be most relevant to this study. According to Specter (2008), each civilization develops formal methods of educating its youth more efficaciously than what can be learned simply through everyday experiences using trial-and-error methods. During the last 30-35 years, technological innovations have stretched the boundaries of today's educational system to the point where the emphasis is now on incorporating these new technologies, such as integrated learning systems and networked electronic learning environments. The problem posed by technological innovations in education is this: Should the emphasis of educational technology be based on educational goals in order to develop an educational theory of technology? Or should it be more about theorizing educational technology? Specter (2008) writes that it has only been since the early decades of the of the 20th century that individuals and affiliated professional groups began a concerted effort to study technology's impact on knowledge and learning, and to establish educational technology as a field worth studying.

Historically, teachers and schools have not been reluctant to adopt technology innovations; however, millions of dollars are needed to integrate technology into the classroom and instruction (Specter, 2008). It is 13 years into the new century, and the field of education is still trying to figure out how to use these technological tools to help teachers improve teaching and learning. Lam (2000) suggested the problem is not “technophobic teachers” but possibly “technophilic district” policy makers. Too often they purchase the latest and greatest technological innovations without considering the needs and technology training of teachers and students.

Tests administered in the classroom are systematic tools that teachers use for collecting information on learning. The data from these types of measures produce quantified results, or evaluations of the measurements. However, it is the teacher’s judgment about the quality or value of learning experience that determines the learning outcome (Johnson, 2008). Thus, assessments are an important part of an overall process for determining student achievement, modifying instruction, and improving curricula.

The concepts of multiple intelligence, constructivism, and mastery learning all contribute to the discussion of formative and summative assessment practices. Each of these frameworks adds contributions in the development of learning strategies and formative assessments; therefore, it is valuable to delve further into research that examines their role in formative assessments. The contributions these theoretical frameworks provide in the development of formative and summative assessment for teaching and learning are discussed further in Section 2.

Nature of the Study

The study used a stepwise multiple linear regression design. The research sought to answer questions about the predictive power of online formative assessments towards an end-of-year summative assessment (SAT-10). The dependent variable included the 2010 SAT-10 percentile-ranking scores for total reading and total math. The independent variables included (a) the SAT-10 scores for reading/math content during and after the treatment of the iReady system (2010, 2011, 2012, and 2013) and (b) schools participating in the study. The analysis design was to determine if a variation existed in the SAT-10 scores either before or after treatment of the iReady system and which, if any, of the variations are explainable by the independent variables. Both the dependent and independent variables were interval and continuous. The study analyzed archived secondary data sets. The data sets included archived and de-identified SAT-10 scores. The data sets used tests scores from students in Grades 1–4 enrolled in 23 private schools in nine states. The data sets included four years of data collected before and after implementation of the iReady quarterly, online formative assessment program.

Operational Definitions

Curriculum-Based-Measurement: is an assessment tool used in classrooms to measure student progress in basic academic areas such as math, reading, writing, and spelling.

High Stakes/End-of-year Summative Assessments/Tests: For the purpose of this study these types of assessments are formally administered to students to measure end-of-year progress towards predetermined learning objectives which (a) determine students' exit achievement from a course, grade or program of study, and (b) include important consequences for student (i.e., advancement to the next level of education, high school diploma, scholarship, or licensure requirements)

iReady: Curriculum Associates (2011, p. 3) define their product, iReady, "as a robust online platform offering computer-adaptive diagnostic, personalized data-driven instruction on foundation skills, standards-based practice, and a Common Core readiness screener."

Stanford Achievement Test-10 (SAT-10): Pearson, the testing company that owns the assessment defines the Stanford Achievement Test Series as a measurement of student progress toward high academic standards. The SAT-10 is a multiple-choice assessment tool that provides data to evaluate progress toward meeting NCLB federal mandates and state and national standards. The test helps teachers identify children at risk for failure; it also provides documentation for parents to understand what their child knows and can do and what parents can do to help (Pearson, 2012).

Summative Assessments: This study defines summative assessments as final evaluations formally administered to students with the intent to measure if the students met end-of-year standards or learning objectives. Theses assessments are the process through which educators measure the sum of student learning towards a predetermined set of standards or expectations. Summative assessments give educators data based on cumulative measurement of student achievement and the progress of learning in an individual classroom, as well as, provide school-wide data on the quality of curriculum, instruction and their alignment to standards.

Assumptions

Two assumptions were made in this study. First, students that took the SAT-10 assessment before and after the treatment will show improvement in reading and math scores. Second, iReady, the online formative assessment system administered during the treatment period could impact SAT-10 scores in reading and math.

Scope and Delimitations

Today states, districts, schools, and teachers are facing greater pressure on accountability than any time in education's history. The scope of this study was to examine how online formative assessments are impacting student learning. The primary sources used to determine student progress are formative and summative assessments. Online formative assessments provide data that affords teachers the opportunity for immediate instructional adjustments. Knowing the impact an online formative assessment can make on an end-of-year assessment would provide valuable information in how to implement an online formative assessment program to improve student achievement.

The study population consisted of elementary students in Grades 1–4, with students in grades 5-8 excluded because of lower student population numbers and inconsistency in administration of treatment. The students attended private elementary schools located in Arizona, Illinois, Texas, Nevada, North Carolina, Oregon, Pennsylvania, and Washington. Students in each grade level took the SAT-10. The data analyzed came from SAT-10 assessments that each school conducted during the spring of each study year (2010-2013). The data sets consisted of 2 years of SAT-10 scores for Grades 1–4. Additionally the data represented the time before, during, and after implementation of the iReady treatment. Prekindergarten and kindergarten students were not given the iReady assessment during the study period; therefore, they were not included in the study.

The independent variables included the participating schools and the scores for the SAT-10 total reading/math content. The dependent variable was the SAT-10 percentile-ranking scores before and after the iReady treatment for all students in Grades 1–4. The construct validity of the study eliminated threats to internal validity, such as, confounding, selection bias, history and maturation, repeated testing, instrumentation, regression toward the mean, differential attrition, diffusion, and experimenter basis.

Limitations

Limitations of quantitative research such as internal validity, history, maturation, "testing effect" or anxiety were not considered a problem with this study because the examined archived data were collected before the study began. Instrumentation or instrument decay was not an issue because of the administration of identical pre and posttest assessments administered during the treatment time frame. The researcher made a conscientious awareness of and subsequent removal of any bias.

Significance

Technology has tremendous possibilities for improving education. However, it is important to not consider it a strategy for learning. A tool is not effective unless the operator has the knowledge and expertise to maximize the tool to produce positive results. Teachers are still the prime conduits through which learning occurs. Gao and Mager (2013) suggest that there is value in understanding that learning to teach using technology is a multidimensional, developmental process. The study provided an opportunity for the private school corporation participating in the study, to evaluate the strengths or weaknesses of their curriculum. It also provided data to help determine the sustainability of the iReady system. The iReady system is aligned with the CCSS, allowing the company to determine if and how strongly their curriculum integrates with CCSS. Another valid use of the information from the study is the use of the data to determine teacher professional development needs.

A concern for school administrators is the manner in which teachers execute their interpretations of online, formative assessments. In addition, there are the questions of how teachers can align the data with their instruction, and how to formulate instructional responses based upon their analysis of the data. Moving from an analysis of formative assessment results to planning instruction requires teachers to determine what instructional changes (or items of information) to change or re-teach, and for which child (struggling, higher understanding, etcetera). This study provides the field of education a stronger understanding of the impact online formative assessments have on direct classroom instruction and teacher trainings, and also provides information for developing teacher training programs aligned to classroom use of online formative assessments.

This study acknowledges that online formative assessment data is tied directly to teacher instruction and student learning expectations through common standards. As a result, it is important to determine whether or not online formative assessments can be a predictor of success in progress and learning outcomes as measured by an end-of-year summative evaluation. The results from this study provide that knowledge. The examination of online formative assessments impact on student achievement provided by this study presents a research-driven perspective into the importance of further research into online assessments and their impact on classroom instruction, teacher improvement, and student learning.

Summary

Section 1 introduced the concept of measuring and comparing student achievement scores using two different types of assessments. The Section also addressed the purpose for providing evidence of online formative assessments impact on measuring progress towards end-of-year academic performance. Section 1 also introduced the background of the classroom formative and summative assessments and provided the definitions used throughout the study. In addition, the Section addressed research questions, framework, nature of the study, as well as, the scope and delimitations, and the significance of the study.

Section 2 discusses and analyzes relevant research related to instructional platforms and their ability to present evidence of learning. This Section also discusses the purpose for, and feasibility of, determining the ability of online, formative assessments to predict grade-appropriate student mastery measured by an end-of-year, summative assessment (SAT-10). Sections 3, 4 and 5 provide the methodology along with the data analysis and implications for future research.

Section 2: Literature Review

Introduction

The purpose of this study was to examine the correlation between online formative assessments and end-of-year summative assessments. This Section discusses the possibility that online formative assessments may be able to predict appropriate grade-level mastery in reading and mathematics as measured by a nationally normed, end-of-year summative assessment.

Section 2 presents four sub-headings. The first three include a discussion, a review of present trends in formative and summative assessments in an educational environment, and an examination of the current literature on the fundamental concepts of this study. The fourth sub-heading focuses on the methodology used to investigate the hypothesis that online formative assessments can predict student achievement in reading and writing, and correlate with data produced from end-of-year summative assessments.

Literature Search Strategy

Approximately 150 articles on formative and summative assessments, high-stakes testing, computer-aided assessment, learning and, instruction were reviewed. The following databases identified literature within a 5-year timeframe: Educational Resources Information Center (ERIC), Education Research Complete, Academic Search Premier, Google Scholar, Mental Measurements Yearbook, SocINDEX, and Teacher Reference Center. The researcher reviewed the National Center for Educational Statistics (NCES), and the National Assessment for Educational Progress (NAEP) looking for current student achievement data figures. Included in the review are the seminal works of Vygotsky, Dewey, Bloom and Gardner.

Internet searches for current articles used the following keywords: assessment, summative, formative, technology integration, assessment and measurement, accountability, instructional technology, multiple intelligences, cognitive learning objectives, design principles, e-learning, online learning, CBM, computer-aided instruction, learning and technology theories, emerging learning theories, Constructivism, education pedagogy. All searches incorporated a filter to search for relevancy; however, applicable information discovered in seminal works and older articles are used in the discussion.

Theoretical Foundation

Developing, fostering and measuring learning in today's technologically advanced world requires an understanding of the research behind instructional technology. Combining research related to instructional technologies with the learning theories of Vygotsky and Dewey, and the added perspectives of Bloom's Cognitive Taxonomy and Gardner's Multiple Intelligences offers evidence-based and theory-grounded approaches. The works of these four theorists are relevant in the examination of online formative assessments predictability towards student achievement.

The prevalent philosophical orientation of instructional technology in the latter part of the century was instructivism (Anderson & Kanula, 1999). Instuctivists argue that the instructional designer should systematically identify what is to be taught and how it is to be taught. The instructional designer also needs to include an evaluation of the instruction (not learning) to determine if the instructional method was effective or not (Anderson & Kanula, 2007).

In this system, the importance of learning comes from the definition of expected learning outcomes and what must be known prior to any learning transaction. The learning objectives begin with a lower order of understanding and advance to a higher order. Important considerations included in the instructional design is that learning objectives are identified and clearly stated. Activities need to focus on learning expected skills and under conditions that include an ongoing and cyclic process (Anderson & Kanula, 2007). Instuctivists argue that this type of system design allows the teacher to focus on the needs and abilities of the individual learner. Additionally, evaluation tools measure the behavior or learning described in the stated objectives, thereby allowing data from the evaluation to immediately be used to revise instruction so that it is more effective for current students, as well as subsequent students (Anderson & Kanula, 2007).

Constructivism is a polarized view of instructivism. Constructivists theorize that the accomplishment of constructing knowledge is through a social linguistic process that includes gradual advancement of understandings built upon prior knowledge (Anderson & Kanula, 2007). Placing the stress on how learners themselves construct knowledge. Varieties of experiences are necessary for the development and understanding of new knowledge.

The constructivist's philosophy for learning is that learning is an active process that helps learners create meaning from their experiences and interactions with the world by owning their own learning process, and by engaging in learning opportunities experienced in natural settings. Additionally, a constructivist views the learning process as a social activity, involving collaboration, negotiation and participation in authentic practices of communities (Wilson, 2010). Importantly, constructivism also acknowledges the value of reflection, with assessment and feedback being embedded naturally within learning activities themselves (Wilson, 2010). The very nature of constructivism rests on a descriptive base or foundation, but extends to guidelines for instructional design. Constructivism supports engaged learning, between the learner and content, versus sedentary learning in which learning is “done to” the learner.

Although Dewey and Vygotsky differ on their views of the human thought process for learning, their concepts about the outcomes of education are the same. Vygotsky (1978) theorizes that the social interaction plays a fundamental role in the development of human cognition. His theory is that a child’s cultural development appears twice, first on the social level and later on an individual level. Vygotsky’s Zone of Proximal Development suggests a difference between the level of learning that one reaches by studying alone, and the level one can reach if working with a teacher or a more advanced peer (Vygotsky, 1978). In this way, education is the social process of learning. An example of a Vygotsky-themed classroom is one in which students are busy in collaborative group work, including opportunities for peer review along with an effort on the part of the teacher to connect students to the learning.

Vygotsky (1978) saw the teacher as having a greater control over mentoring and creating the activities that will lead the child towards mastery. Teachers in the classroom serve as mediators and coaches whose primary role is encouraging students to formulate their own level of understanding. Students increase their knowledge through practicing what they know and then transitioning into learning something more. Additionally to Vygotsky (1978), social interaction between and among the students, peers, and the teacher is what reinforces the growth of new learning constructs. It is through social interactions that the full cognition develops. Vygotsky's theory does not offer an alternative to Dewey's constructs for the development of knowledge but rather the two approaches complement each other. Both Vygotsky and Dewey condemn an elitist teacher attitude as not being conducive to student learning. Both support the idea that teachers are not the ones who have all the answers, but rather, they are the facilitators for the process of gaining knowledge.

Dewey (1934) saw experience as a means of forming thinking. It is our social interactions that enable or force us to pay attention to contributions made by other participants. Dewey (1934) held that education is our ability to question our reality through experience and is as important for the individual as it is for the human community. Dewey (1934) rejected the idea that schools focus on repetitive and rote memorization. Instead, he proposed a method for directed living in which students engage in real-world practical experiences that demonstrate knowledge through creativity and collaboration. Envisioning Dewey's classroom, one would see the student as a free agent, achieving goals through interest in the activity. Dewey's classroom would include assessments that focus on tasks and learner analysis. Methods and results in a Dewey-structured classroom are not easily measured because open-endedness is difficult to measure and outcomes may not be the same for each learner. Instruction is meant to foster learning not control it, and learning outcomes are not always predictable.

In a technology-enabled classroom following Dewey's education theory, one would see a classroom in which information is analyzed based on tasks broken into smaller chunks of information. The learning progression would follow a simple to complex process, building upon prior schema. In this classroom, learners would determine their own predisposition toward learning with instruction that is more facilitative than prescriptive. This model of the classroom is dependent on students learning how to accomplish a task; however, it may not be well suited to each student or situation. Therefore, learning tasks that require an increased level of intellectual processing should be the predominate method of instruction in this type of learning environment. For Dewey, learning occurs when the possibilities inherent in ordinary experiences are subjected to the tests of intelligent development and direction Progressive organization of subject matter is what allows scaffolding upon existing experiences (Dewey, 1934).

Much of Dewey and Vygotsky's work contains similar ideas about the role of education. Combining these theories produces a learning environment that places an emphasis on learning how to think rather than on rote memorization. Also, this environment would be lead by a guide or facilitator rather than a strict authoritarian figure. Both scholars would view the teacher as being more of an equal to the students, rather than a teacher who is wiser and more knowledgeable than their students. Both theorists would expect to see the interaction between the students and teacher as an integral part of the learning process.

When combined, the connections between these two theorists create a classroom environment that includes priorities in social-emotional learning, student engagement, and is learner-centered. There would be multiple opportunities for students to think and solve problems for themselves. Both Dewey and Vygotsky perceived constructivism as a process conducive to learning.

The constructivist platforms of Benjamin Bloom's Cognitive Taxonomy and Howard Gardner's Multiple Intelligences were also worth examining, in the context of this study. In the constructivist's world, learning is not simply a linear process but an active one. It is a process in which learners are actively constructing and modifying what they know (Proulx, 2006). The works of Bloom and Gardner bridge the theoretical gap between learning and teaching.

Bloom (1956) defined knowledge as one's ability to remember specifics, methods, processes, patterns, structures, or settings. He wrote that for measurement purposes, recall requires little more than bringing to mind particular bits of information. Remembering is the process of providing cues for the information and knowledge already learned. A higher level of knowledge is when intellectual abilities and skills become the process of organizing and reorganizing information and thus achieving a particular purpose.

Bloom's Taxonomy provides educators with a structure for teaching through the creation of objectives that are aligned to the process of learning (Bloom, 1956). The six levels in the taxonomy are remembering, understanding, applying, analyzing, evaluating and creating. Bloom and his group considered their framework to be a work in progress. The increase in modern technology usage has not caused education developers to be stifled by an outdated taxonomy, but rather Bloom's Taxonomy has moved in the direction of becoming more like the dynamic process that Bloom's team envisioned (Anderson & Krathwohl, 2001).

Gardner began his work by examining assessments that claimed to measure intelligence or IQ. Gardner (1987) viewed IQ measurements as a one-dimensional measurement of multi-dimensional levels of intelligence, and, therefore, not measurable using a one-dimensional test. Gardner's (1983) design of multiple intelligences suggests that there are seven distinct forms of intelligence that individuals possess in varying degrees. These forms of intelligence are linguistic, musical, logical-mathematical, spatial, body-kinesthetic, intrapersonal (e.g., insight, metacognition) and interpersonal (e.g., social skills). Moreover, the implication is that learning/teaching should focus on the particular intelligences of each person. For example, if an individual has strong spatial or musical intelligences, he/she should be encouraged to develop these skills (Gardner, 1983). Gardner (1983) wrote that the contrast in intelligences represents not only different content domains, but also different learning modalities. An additional implication is that assessments should measure all forms of intelligence, not just linguistic and logical-mathematical. Gardner (1983) also emphasized the cultural context of multiple intelligences.

The result of the 21st century explosion of knowledge sharing is giving rise to a new dynamic tool in the classroom toolbox: education technology. The definition of education technology had been evolving since 1970. The Commission on Instructional Technology defined it "as a systematic way of designing, carrying out, and evaluating the total process of learning and teaching in terms of specific objectives, based on research in human learning and communication, and employing a combination of human and nonhuman resources to bring about more effective instruction" (Saettler, 2004, p. 6).

The commission added the caveat that the broad definition of education technology belonged to the future. According to Saettler (2004), the definition continues to be a topic of discussion and debate, with a persistent problem being the use of the word technology to refer only to hardware. He cautioned, "because of the current changes in the field, the definition is seriously outmoded and is in need of revision" (Saettler, 2004, pg. 7). For the purpose of this study, the Commission's 2007 definition was used.

Added to teachers' stress in understanding the fluctuating definition of educational technology, is the process of designing, carrying out and evaluating the learning process. Furthermore, technology brings additional stress to the classroom in the form of computer assisted learning environments, mobile technology, digital audio, video and still cameras, LED projectors, and social networking platforms web-based curriculum, online instruction and blended learning. These new technologies provide a plethora of tools that complicate teaching. However, it is innovative teaching tools such as these that will move instruction from the base of static to dynamic inquiry (Pelligrino, 2010).

Another consequence of the speed in which the educational technology field is advancing is the limitation of districts and school administrators to adequately address the frustrations of classroom teachers as they struggle to define and measure what computer programs do for teaching and learning (Harlen, 2005). According to Pelligrino (2010), the abundance of innovative technology is urging the field of education to create and implement new generations of enhanced formative and summative performance evaluations that are able to measure complex forms of learning through computer technology. Saettler (1990) suggests that innovations have been transforming learning throughout history and that many of these transformations occurred without a single tool or piece of hardware. Therefore, it is essential for education to stop equating technology with machines, or suggesting technology as a replacement for classroom teachers (Saettler, 1990).

Key Concepts

A thorough search of current literature showed a limited number of studies comparing the predictability of online formative assessment programs on student achievement as measured by an end-of-year, nationally normed, summative tests, such as the SAT-10. However, there was research indicating educational software designed with the component of computer-based instruction (CBI) computer-based response to intervention (RTI) and curriculum-based measurement (CBM), can provide valid assessments for predicting student achievement (e.g.: Alonzo, Nese, Park, & Tindal, 2011, Hannafin & Foshay, 2008, Keller-Margulis & Hintze, 2008).

Formative Assessments

Formative assessments serve a valuable role in the classroom. Bax, Branford-White, Heugh, and Jacoby (2013) suggested teachers use assessments to estimate or predict student learning; whereas, students see assessments as a motivation for learning. Formative assessments require active participation from the teacher and the student (Bax et al., 2013). Assessments may look very different from one another, but their design requires a framework that encourages the teacher to monitor student progress continually and create a viable feedback loop between the learner and teacher (Bax et al., 2013).

Therefore, examination of the literature addressing both formative and summative assessment concepts added to this conversation on the predictability of online formative assessments towards end-of-year summative assessments. Bennett (2010) purported that in order to fully understanding the role of computer-based assessment systems that measure how what students have performed and how to plan instruction, it is important to develop a theory of action. Bennett (2010) developed a theory of action and an explanation for what elements to included in the assessment system. The elements Bennett (2010) suggested are the intended effects of the assessment system, components of the assessment and a coherent rationale for each component, including backing for the rationale in research and theory. He also included interpretive claims made from the assessment results; action mechanisms implemented to cause the intended effects and lastly, the potential for unintended negative effects and strategies to mitigate theses effects.

Cognitive Based Assessment for Learning (CBAL) is Bennett's initiative to create a K-12 assessment system that documents what students have achieved (of learning) facilitates instructional planning (for learning), and the educational experience in and of itself (as learning). Bennett's research in CBAL was important to consider for this project because his elements provide a working foundation with which to review the similarities between the two assessments, SAT-10 and iReady. It also provided a reasonable procedure for comparison of the two assessments data sets examined in this for analysis.

Curriculum-Based Measures (CBM)

Examination of the assessments used in addressing the needs of special population students is another important dimension to consider when examining formative evaluation data to predict success on high stakes tests (Deno, 2013). The exploration into different education programs with measurement instruments designed to differentiate, diagnose and prescribe learning interventions, added insight to research on measuring instruments for regular classroom instruction (Deno, 2013).

CBM originated to test the effectiveness of special education intervention model. The model uses repeated measurement data that formatively evaluate learning and improve teaching in a special education environment (Deno, 2013). CBM's design is to use a generic set of progress monitoring procedures for measuring performance based upon core tasks, stimulus items, measurement activities, and scoring and decision rules (Deno, 2013). The original purpose of CBM was to enable teachers to evaluate their teaching from formative data.

Hixson and McGlinchey (2004) added that CBM is moving away from an elementary, special education, assessment model towards a more commonly used model for measuring. They claim CBM is a valid program for improving individual instruction programs, predicting performance based on specified criteria, high stakes assessments, and enhancing teacher instructional planning, (Hixson & McGlinchey, 2004). Fuchs and Stecker (2000) describe two features to differentiate CBM from other forms of classroom-based assessments. First, academic performance is frequently assessed on yearlong curriculum through standardized tests with the scores displayed graphically. Second, teachers use the graphic representations to determine instructional adjustments throughout the school year (Fuchs & Stecker, 2000). Their study showed that when teachers adjusted instruction based upon individual progress monitoring data students performed significantly better on a global achievement test than did others who did not receive instructional changes. Hintze, Keller-Marguulis and Shapiro (2008) conducted a study to determine if a CBM program could be a useful tool for identifying students at risk of developing academic problems or in need of potential interventions in reading and mathematics. The researchers examined the relationship between reading and math CBM on performance and growth rates, as well as, the diagnostic accuracy of CBM on statewide and large-scale assessments at one and two years out. They found CBM provided sufficient diagnostic accuracy for screening of performance on statewide and large-scale assessments (Heller et al., 2008)

Hintze et al. (2008) considered a CBM program to determine its ability to provide adequate diagnostic accuracy for performance on statewide and large-scale achievement tests. The diagnostic accuracy of the CBM reading and math measures showed a correlation between the percentages of students who performed better than expected versus students who did not participate in a CBM program (Hintze et al., 2008). They concluded that electronic-based CBM provided a model for integration of instructional technology for monitoring student progress toward instructional objectives (Hintze et al., 2008). The results of the study provided strong evidence of the long-term diagnostic accuracy of the CBM. The practical implications from this study showed the need for further research in the area of formative assessments and their predictability towards student achievement.

Tsuei's (2007) study indicated strong evidence that a class-wide dynamic-growth modeling strategy was more effective for students in mixed-type CBM probes than students in single-type CBM probes. Tseui (2207) observed that in relation to their classmates, students in the dynamic-growth modeling group were more aware of their mathematics performance, which in turn promoted self-expectation (Tsuei, 2007). He hypothesized that it was how teachers implement different CBM probes (math probes and growth models) that created the potential for CBM to be an assessment tool, which can help teachers integrate instructional strategies based on data (Tsuei, 2007).

A study conducted by Alonzo, Nese, Park and Tindal (2011) examined the relationship between the easy Curriculum-Based Measurement (CBM) formative assessment program and a statewide, large-scale reading test. The researchers sought to establish evidence of predictability of CBM measures on reading fluency in elementary school students. Using multiple regression analyses, Alonzo et al. (2011) studied 3,600 students in fourth grade. Their data analysis produced evidence that CBM has potential but would benefit from additional studies in two areas. First, the potential for CBM reading measures to provide beneficial insights for teachers preparing students for state mandated reading assessments. Second, there is a potential for CBMs to be used as a screening assessment to identify students who may be at risk for failing the end-of-year high stakes assessment (Alonzo et al., 2011).

Forster and Souvignier (2011) added to the understanding of using Curriculum-Based Measurement (CBM) to monitor student growth. A hierarchical model of reading comprehension was used to predict student growth. The study design utilized a pre- and posttest measurement of both reading comprehension and mathematics scores from a CBM assessment and state administered end-of-year assessment. The results indicated an overall increasing pattern for reading rate and comprehension. The findings also demonstrate technical soundness for using CBM models to monitor student growth in reading achievement. Forster and Souvignier (2011) suggested that future research should include exploring reading accuracy and comprehension measures to determine the reading progress of poor readers and the potential to predict reading disabilities.

Computer-Based Assessments and Instruction

The technology era and its accelerating changes on society are redefining the skills and competencies needed for success in today's workforce. Even the skills needed in everyday life are being stretched by the technological innovations of the 21st century. Johannsen and Redecker (2013) suggested that education should focus on the development of transversal and basic skills. They suggested entrepreneurial and IT skills are a necessary addition to classroom instruction (Johannsen & Redecker, 2013). However, because assessment is an essential component for influencing practices and affecting learning, bringing IT skills into the learning process requires that the assessment processes be improved as well (Johannsen & Redecker, 2013). The researchers suggested that, for either formative or summative assessments to be valid, students would have to be developed into self-directed learners capable of monitoring their own work (Johannsen & Redecker, 2013). The authors noted that assessments would have to go beyond the mere testing of facts to being able to assess the intangible themes and underlying learning competencies that computer enhanced assessment tools are able to recognize (Johannsen & Redecker, 2013).

Institutions are beginning to examine the use of computer-assisted assessments as a means of formally measuring academic progress. The formative process increases the student learning time, familiarity with the materials, and introduces the student to the summative information that they may encounter on an end-of-year or high stakes assessment (Bax et al., 2013). Electronic assessments afford opportunities for immediate and ongoing feedback, which have the potential to enhance student learning, as well as providing reflection on the outcomes of that learning (Bax et al., 2013).

Foshay and Hannafin (2008) conducted a study examining high school students' end-of-year summative MCAS (Massachusetts high stakes assessments) scores in order to determine if computer–based instruction (CBI) played an integral role in the overall remediation strategy that prepared students for the end-of-year state mandated test. The evidence they presented suggests that students do benefit when challenged with computer supported learning environments that demand critical and complex thinking (Foshay & Hannafin, 2008). However, they reiterated that it was the efforts of teachers and students, along with the CBI, that produced the positive results (Foshay & Hannafin, 2008). The researchers concluded that any continuing discussion should focus on the benefits of challenging student teaching and learning through well designed, direct instruction delivered via computers (Foshay & Hannafin, 2008).

Burns, Klingbeil and Ysseldyke's (2010) looked at Technology-Enhanced Formative Evaluation (TEFE) and its effect on student performance on state accountability tests measuring math competencies. The TEFE system uses a framework to administer computer adaptive assessments to students. The data from the TEFE program that was examined helped teachers determine appropriate instructional targets while also allowing them to monitor student progress (Burns et al., 2010). The researchers compared a TEFE computer software program in which students completed curriculum objectives using individualized, software generated assignments, to students using only a computer screen to complete math assignments. Burns et al. (2010) wanted to determine if schools using a TEFE program would have a higher percentage of students classified as proficient on statewide assessments than schools that did not employ any TEFE program.

Burns et al., (2010) discovered schools that used a TEFE program saw a higher percentage of student scores reaching the proficiency range. In addition, results from a five-year period showed that schools participating in the TEFE program had a slightly higher percentage of students scoring at the proficient level than those schools that only used the program for one to four years. Burns et al. (2010) concluded that TEFE programs could have a positive effect on the students' scoring at the proficiency level on statewide assessments. They encouraged further investigation into how TEFE programs might improve student academic proficiency on summative assessments. In addition, Burns et al. (2010) suggested research delving into the effect of using of a TEFE program on the academic progress of children that had been referred to an RTI (response to intervention) program as a result of a TEFE referral.

Baker, Goldstein and Hefferman (2011) considered an intelligent tutoring systems (ITS) model and their ability to detect accurately that given moment when a student learns a particular skill or knowledge content (KC). Their study examined a computer-based model designed to discover at what particular point in a problem solving process the learning occurs (Baker et al., 2011). The researchers studied two tutor-type software programs, 1) a middle school cognitive tutor called Bridge to Algebra, and 2) ASSISTment Tutoring (Baker et al., 2011). Both programs included multiple components that supported students that encountered difficulties with specific knowledge content. Additionally, both systems included tailored feedback for particular common misconceptions detected within student behaviors and a multi-level, on-demand hint system.

With the Middle School Cognitive Tutor, hints are provided automatically; whereas, with ASSITments, students had to either incorrectly answer the problem or request help. Baker et al. (2011) discussed the potential uses for moment-by-moment learning detection. The team of researchers began with the hypothesis that these types of computer-based intelligent tutoring models improve the effectiveness of cognitive mastery learning. They analyzed data from 4187 middle and high school students. The results of the analysis showed that overall, the tutoring models were successful in predicting established training levels (Baker et al., 2011). The power to predict the moment when student learning might occur on a computer-based tutor is important to note when discussing computer-based formative assessments and their ability to evaluate student learning effectively (Baker et al., 2011).

Eggen, Timmers, Van der Kleij and Veldkamp (2012) conceded that feedback is an essential element for assessment when integrated into the learning process and that it is key to moving learning forward. They sought to determine if the type of feedback received from a CBA program affected students' attention to feedback differently than if the feedback came from the teacher. Eggen et al. (2012) suggested that students pay more attention to immediate feedback versus delayed feedback. They concluded that the time spent on reading feedback positively influenced students' attitude and motivation (Eggen et al., 2012). The researchers acknowledged that since the study only considered written feedback, further research is needed to examine different types of feedback within a larger study population in order to increase statistical power, as well as establish if any significant effects occur (Eggen et al., 2012). Additionally, they indicated further research could examine the correlations between various variables influencing student attention towards feedback (Eggen et al., 2012).

Parallel to examining student learning and students' understanding of feedback, one should also consider students' acceptance of CBA as something that enhances their learning experiences. To that end, Economides and Terzis (2011) examined how the development and use of the CBA program are dependent upon student acceptance. The researchers probed into the measurement and structural model of a CBA program through the lens of student perception, for ease of use and perceived playfulness (Economides & Terzis. 2011). Their results showed student understanding had a direct effect on the development and use of the CBA program. They rationalized that CBA assists educators by moving beyond the typical methods of testing security. CBA decreases cost and time, increases the speed of results, allows for automatic record keeping, and improves time analysis. They also stipulate that CBA expands the potential for using technological innovations in testing and assessment (Economides & Terzis. 2011).

Summative and High Stakes End-of-Year Assessments

Education reform efforts are pressing for assessment tools that take advantage of technology, but still remain guided by cognitive models of progression towards competence (Haertig & Nehm, 2012). The use of these technologically enhanced assessments is important for revealing critical junctures towards student conceptual understanding and for measuring instructional efficacy (Haertig & Nehm, 2012). Learners throughout the educational hierarchy are assessed on a daily basis, whether it is to measure learning progress or to determine the sum of learning. Society's progress towards adaptation of 21st century standards of information and communication technologies is reshaping education. Johannessen and Redecker (2013) described this re-shaping of society as increasing expectations from schools to produce learning environments with higher levels of ability in problem solving, reflection, creativity, critical thinking, innovation, risk-taking and entrepreneurship.

As much as technology is reshaping teaching and learning, so are reform movements that are calling for end-of-year testing of all students. Federal funding requirements have placed pressure on schools for accountability testing to the point that end-of-year testing is acceptable as a substantive structure of change (Supovitz, 2009). Investigation into the relationship between online formative assessments and high-stakes/end-of-year summative evaluations is essential to understanding the consequences for predicting success on these widely-used assessments for student achievement.

It is important to identify the fact that teachers' concepts of formative and summative assessments assists in understanding the alignment or misalignment of formative practices and the conceptual changes that occur because of outcome-based summative assessments (Biemans, Gulikers, van der Wel & Wesselink, 2013). They contended that teachers are critical in determining whether or not an educational innovation is implemented to the point of sustainability; therefore, teachers' conceptions can hinder adoption and fidelity of implementation. Biemans et al. (2013) discovered a misalignment between teachers' perceptions of assessments and contemporary views of assessment. In conjunction with the misalignment, the researchers found that teachers do not differentiate between formative and summative assessments (Biemans et al., 2013). They found teachers' preference, when grading and certifying the end of the learning process, was to use both formative and summative assessments. This misalignment is incongruent with current research indicating that formative assessments are most helpful in improving instruction and the learning processes. The teachers in the study understood the assessment process to be an isolated activity. This view stands apart from the view that an assessment framework is a grounding component that is important in the educational system. The researchers suggested additional research on the effectiveness of formative assessment practices in relation to summative assessments (Biemans et al., 2013).

Lam (2013) recognized the need for placing emphasis on the benefits of formative assessments, although he does suggest that there is a lack of research on how to utilize summative assessments in a formative capacity. Lam's (2013) theoretical framework focused on the relationship of functions between formative and summative assessments. The key issue was whether the functions of these two assessments are incompatible or whether they were synergized (Lam, 2013). His research examined effective use of formative test preparation strategies and any resulting productive synergies displayed by the students on the summative assessment (Lam, 2013). Lam's (2013) findings demonstrate that formative test preparation strategies have the potential for actively involving students in the assessment process, thus motivating them towards constructing knowledge and mastery of expected content (Lam, 2013).

Summary and Conclusions

In summary, the research presented throughout Section 2 provided evidence that technology is a useful tool for evaluating student learning and in guiding instructional strategies. Additionally, research supports the need for examining the educational effects of computer-aided formative assessments towards end-of-year summative assessments and the role both assessments play in developing instructional strategies and improving student achievement. Increasingly, educational reform efforts are turning towards data-driven decision-making strategies that improve instruction and align content to rigorous standards. Collecting, analyzing and placing instructional data into teachers' hands is essential to promoting school systems while increasing efficiency that ultimately leads to improved student achievement and better system-wide performance.

Section 3: Research Method

Introduction

The purpose of this Section is to provide an overview of the selected research design along with a choice of methodology, setting, population, structure, instrumentation and additional resources employed in the process of answering the study research questions. The purpose of the study was to determine if an online formative assessment had any impact on an end-of-year summative evaluation, used to determine student achievement in reading and total math over the course of one year.

Research Design and Rationale

This study sought to determine if there was a relationship between an online, formative assessment program (iReady) and the SAT-10, a summative assessment measuring end-of the-year progress in math and reading. The study used a stepwise, multiple-regression model. The design was most appropriate when offering a hypothesis to explain O1- O2 difference. Maintaining internal validity was a concern with this design because of deficiencies that might occur as a result of maturation, testing effect/reactivity, instrumentation, and regression. This study addressed the following methodological weaknesses:

- The pretest data already existed without students knowing they were participants
- The posttest SAT-10 data included archived data sets from the treatment years.

Methodology

Population

The study population consisted of students in Grades 1–4, enrolled in a corporately owned, private school setting. Students participated in a quarterly online formative assessment system (iReady) and an end-of-year summative assessment (SAT-10). These assessments measured student progress and grade-level achievement in reading and mathematics. This study examined the archived data from these assessments.

Sampling and Procedure

The sample population was similar to a small suburban school with a student population of 300 elementary students in Grades 1–4. However, by combining the data from all of the private schools and their students, the sampling number increased to such a degree that the sample population represented a large suburban elementary district of more than 22,0000 students enrolled in Grades 1–4.

The SAT-10 assessment is administered every spring to all students; however this study considered only the data students in grades 1-4 in the years from 2010 to 2013, SAT-10 assessment results for the years 2010, 2011, 2012 and 2013 were analyzed. The sample population began participating quarterly in the iReady online formative assessment system during the same years as the SAT-10 (2010, 2011, and 2013). A stratified random sample of students from each grade level was used to ensure inclusion of different groups of the population. This sampling strategy allowed for homogeneous samples based on the variables (SAT scores).

Procedures for Recruitment, Participation, and Data Collection

The N corporation owns and operates 35 elementary schools in Arizona, California, Texas, Nevada, Oregon, and Washington, with students enrolled in Kindergarten through Grade 8. The schools in Florida, Illinois, North Carolina, Pennsylvania, and South Carolina enroll Kindergarten through Grade 4. The SAT-10 student achievement scores came from archived data stored electronically by the testing company, Pearson.

Archival Data Procedures

Data collection involved gathering data from the archival storage databases located at the Person publishing company's online test results website. Data collection began once the study received IRB approval (02-20-14-0196906).

Information collected and formatted from the private databases was input into an SPSS database.

Instrumentation and Operationalization of Constructs

SPSS analytical software created a stepwise multiple regression with two ordinal variables (school numbers and SAT-10 total reading/math content) and one dependent variable (SAT-10-percentile ranking scores in math and reading). The study adhered to all assumptions related to multiple regression models in order to provide accurate predictive results, test how well the regression model fitted the data, and to determine the variation in the SAT-10 scores. The analyses addressed and adhered to all assumptions throughout. Section 4 includes a description of the elements.

Threats to Validity

There were no concerns related to internal validity, such as history, because the study looked at a specific period between the administration of the pre-test and posttest. Maturation was not a problem, even though students experienced changes during the time between the pre and posttests. However, that was an expected occurrence because the study was looking at the cumulative effect of the learning process between the pre and posttests. Since the study used a non-reactive, passive measurement of student behavior, the “testing” effect did not present a concern. Because the data already existed, there was no concern that students might or might not experience a reaction to testing anxiety. There was no instrument decay because of the administration of identical pre and posttest assessments (SAT-10) during the treatment time frame.

Ethical Procedures

The CEO of the school corporation granted the researcher access to all student data pertinent to the study, with the caveat that the researcher would share and/or distribute this information only for purposes of development, presentation and/or dissertation review. In addition, the researcher was prohibited from using student names and all data results were to remain anonymous with respect to both student and school (Appendix A).

The collection of the data prior to the research eliminated any student-researcher interaction. The archived data came from archived SAT-10 data sets, beginning with the year before iReady (2010), including two years of SAT-10 data during iReady (2011, 2012, and 2013) implementation, and SAT-10 data from the year following (2013) implementation of iReady.

Reports examined from Pearson's SAT-10 results for students included the Combined Summary Report, Roster Reports, and Subtest Summary Reports. Research began with the downloading of the necessary SAT-10 data reports to the researcher's computer, which were stored in password-protected documents and folders. No conflicts of interest existed. Neither the School Corporation nor Pearson employed the researcher during the study period. The researcher did not receive funding for the project.

Summary

Section 3 discussed the research design and justification for the methodology, experimental design with a one-group pretest-posttest method of observation. The Section discussed the sampling methods along with recruitment, participation and data collection procedures. Section 3 demonstrated the researcher's consideration of constructs with the study variables, the data analysis plan, threats to validity and ethical concerns. Section 4 will discuss the overall fit of the methodology chosen, and how the evaluated data addresses the discussion and the hypothesis described in the study.

Section 4: Research

Introduction

The analyses presented in this Section sought to provide definitive data supporting or negating the use of online formative assessments for measuring student progress towards academic achievement as measured by an end-of-year summative evaluation. Section 4 discusses the results of a stepwise multiple linear regression analysis intended to establish the predictability of an online formative online assessment system (iReady) and determine if there is any calculable power of online formative assessments on end-of-year summative assessment.

Research Questions and Hypotheses:

1. Is there a linear relationship between the results from online formative assessments and student achievement as measured on SAT-10 summative assessments?

   H_0^1: There is no linear relationship between the pre and post SAT-10 scores after the treatment of an online formative assessment system.

   HA_1: There is a significant relationship between pre and post SAT-10 scores after the treatment of on online formative assessment system.

2. Does an online formative assessment system have an effect on student achievement, as measured on SAT-10 summative assessments?

   H_0^2: Is there a difference from the use of an online formative assessment system on student achievement as measured by SAT-10.

HA_2: The use of an online formative assessment system has significant effect on student achievement as measured by SAT-10.

3. Can online formative assessments be a predictor of student success on end-of-the year SAT-10 summative assessments?

H_0^3: The correlation of an online formative assessment system to student achievement is not a reliable predictor of student achievement as measured by SAT-10.

HA_3: The correlation of an online formative assessment system is a reliable predictor of student achievement as measured by SAT-10.

Data Collection

The time frame for the data collection was one week. The data was collected over the treatment period and stored in an online database accessible to the researcher through the publishing company's (Pearson) website, Results online. Only Pearson and private school company officials could access the database. Walden University's IRB (02-20-14-0196906) approved the study on February 14, 2014. The process of downloading the data began after the IRB approval date. The analysis began with importing and cleaning the data from Excel spreadsheets; followed by the conversion of the data into an SPSS database. SPSS software analyzed the data using stepwise multiple regression analyses. SAT-10 percentile ranking was the dependent variable. School number identification and SAT-10 total reading/math content scores represented the independent variables. The data analysis showed no discrepancies, unexpected changes or impacts. The baseline data included SAT-10 percentile rank scores for total reading and total math scores for the years 2010 and 2013. Included in the analysis was data from 24 of school corporation schools; located in 9 states with 339-student data entries from grades 1 through 4. The bulk of the data came from the 19 schools located in California. The remaining data sets spread throughout schools located in Arizona, Illinois, Texas, Nevada, North Carolina, Oregon, Pennsylvania, and Washington.

Treatment

All schools received the treatment during the 2010 through 2013 school years. The treatment was Curriculum and Associates' product, iReady, an online assessment, and diagnostic, intervention software program. According to the iReady Diagnostic and Instruction: Administrator's Guide, iReady provides an adaptive diagnostic that can quickly identify individual student learning weaknesses and strengths in both reading and mathematics (Curriculum Associates: iReady, 2011). The system provides administrative reports on performance by school; grade and class, class and student profiles, needs analysis by grade, instructional grouping profiles, diagnostic and instructional data, class and student response to instruction/intervention, parent reports, state standards and progress towards CCSS performance. The iReady diagnostic portion allows teachers to determine individual students' needs quickly in key strands or domains. According to the iReady Administrator's Guide, combining the diagnostic with the systems comprehensive reports, allow teachers to adjust their instruction while continually monitoring student progress over time.

iReady domains in reading include phonological awareness, phonics, high-frequency words, vocabulary and informational and literature reading comprehension. In mathematics, the domains include number and operations, algebra and algebraic thinking, measurement and data, and geometry. SAT-10's content clusters for total reading scores included word study skills, reading vocabulary, reading comprehension. In addition, SAT-10 included content clusters in spelling and language mechanics and expressions. In mathematics, the content clusters for SAT-10 total math included math problem solving and math procedures.

Results

The results from the study showed that a significant relationship is present between the dependent and independent variables. However, there was small effect size because very little variation is accounted for within the R^2 values. The analysis suggested that the predictive value is in the standardized coefficients, which indicate a low correlation between independent variables of the 2010-pretest school year and 2013-posttest year. Multiple regression stepwise analyses were run to predict SAT-10 percentile ranking from total reading and total math scores, and participating schools. The study results met the assumptions of linearity, independence of errors, homoscedasticity, unusual points and normality of residuals. The independent variables accounted for .116 or 11.6% of the variance (see Table 1).

Table 1

Summary of Stepwise Multiple Regression Analysis

Model	Variable	*B*	SE^{B*}	SE^{C}	β
1	Schools	-.293	.083	.086	-0.25
2	Schools, Reading/Math	.189	.116	.036	0.068

Note. $\rho < .05$; *B = Standardized regression coefficient,* SE^B *= Adjusted* R^2*;* SE^c *= Changed* R^2*, β = Relationship*

The independent variables, schools and reading/math content, were able to statistically and significantly predict the dependent variable, SAT-10 percentile rank scores.

Summary

The results from the data analyses and evaluation demonstrated a linear relationship between the results from online formative assessments and summative assessments. Additionally, the results suggest that there is a possibility that online formative assessment system, such as the iReady, can predict student success on end-of-the year summative assessments, such as the SAT-10. However, the independent variables only accounted for 11.6% of the variance in summative results, which indicated that there are additional variables that would need to be considered to better determine the relationship between online quarterly formative assessments and end-of-year administered summative assessments.

Section 5 delves further into the study's results and evaluates the potential for additional studies into the formative/summative relationships and the impact for driving teaching and learning toward improved student achievement.

Section 5: Results

Introduction

The purpose of this study was to examine the impact of online formative assessments on end-of-year summative assessments, provide research-based evidence on whether online formative assessments have a linear relationship with summative measurements, and determine the potential to predict student achievement. There are studies that examine student achievement and end-of-year summative evaluations; however, as the literature review revealed, few studies examine the relationship between online formative assessments and end-of-year summative assessments on student achievement. The approach for the study was not to compare formative to summative assessments, but to consider if an online formative assessment, iReady, might affect SAT-10 end-of-year scores in reading and math.

This Section includes discussions on the interpretation of the findings, the limitations of the study, recommendations for further research, implications and conclusion.

Interpretation of the Findings

This study used an experimental research design to examine three academic years' worth of archived data for SAT-10 percentile rank scores in total reading and total math for students in Grades 1–4. The data sets were coded and analyzed using SPSS. In order to place the resulting outcomes into a discernable perspective and allow for a deeper understanding of the data, inductive reasoning was employed until an integrated picture of what the data showed became evident.

The first research question sought to determine if a linear relationship existed between the iReady implementation years (2010-2013) and SAT-10 scores during the same testing timeframes. The second question asked if the online formative assessment (iReady) had on effect on the end-of-year, summative assessment (SAT-10). Finally, the third question addressed the predictability of the formative assessment toward student achievement. The findings addressed in Section 4 answer each of the research questions. The data suggested a linear relationship existed between the two types of assessments. The data also showed the significance of a possible interconnection between the two assessments. The findings suggest that research into this symbiotic relationship of formative and end-of-year summative assessments should continue.

The data analysis exposed an occurrence of some significance between the pre and post-treatment years. The SAT-10 data showed an increase in student percentile rankings for total reading and total math between the pre and post administration of the iReady assessment. The analytical findings of these variables point towards a significant relationship. The independent variables, total math/reading content and school numbers, accounted for 11.6% of the variance against the dependent variable, SAT-10-percentile ranking. The study examined an aggregate of the SAT-10 schools' scores instead of reviewing accumulated scores for individual students or individual grade level scores over the course of the iReady treatment. As a result, the research data produced a horizontal continuum of the total study schools' pre and post treatment scores from SAT-10. Another variable considered in the interpretation of the findings was that N School Corporation, before and during the study timeframe, expected all of their schools to achieve a yearly goal of 80% or higher score on SAT-10- percentile ranking. Since the pretest administration occurred with this pre-existing administrative caveat and the data showing improvement in percentile rankings, it is likely that the implementation of the formative assessment program (iReady) added to the overall improvement of the schools' SAT-10-percentile ranking.

Limitations of the Study

The variables presented limitations. The results showed that outside variables contributed to 88.4% of the variance, so an examination of these variables might have provided a different or deeper perspective on the interrelationship between formative and summative assessments. In addition, the study drew its data from an aggregated horizontal continuum of all schools' data between the treatment years; rather than an accrued vertical continuum, reviewing a cohort of student scores as they progressed from first-grade to fourth-grade.

The exclusion of iReady data is a final limitation to consider with the investigation and analyses. Examination of this data might have shown to what extent the SAT-10 data would show if we knew the amount of time students spent in the reading or math intervention portion. Also, a consideration looking into what, if any, extent teachers adjusted their instruction after receiving the results from the diagnostics and interventions portion of the system. A study analyzing the impact of these variables could provide additional insight into the relationship between iReady and similar types of online programs, end-of-year summative assessments or high-stakes tests.

Besides the limitations, it is important to contemplate what the next steps might be. Recommendations resulting from this research are important to consider as they can provide educators with a sense of codification to all of the data actualized by online formative and end-of-year summative assessments.

Recommendations

Although the results of this study contained several limitations and provided no definitive answer as to the impact and specific linear relationship of online formative to summative end-of-year assessments, it does provide some direction for future studies. The results of the study contribute new information about the role that online assessments play on student achievement. Additional studies into teachers' perceptions of online assessments could provide further insight. Chien, Wu, and Hsu (2014) explored this construct and found that 85% of their teacher participants perceived online formative assessments as useful tools; however, nearly 40% of the participants indicated there were difficulties in implementing the assessments. The study also revealed that teacher perceptions about the ease of implementing the online assessment program were mainly negative. In the same study, they ascertained, through the teachers' control beliefs, that the negative feelings focused more towards social and uncontrollable external forces such as time, classroom support and IT infrastructure (Chien et al., 2014). In relation to the teachers' normative beliefs, the researchers noticed that teachers focused more on school policies and parent constraints, even though the teachers appeared to grasp the benefits of the online assessment program (Chien et al., 2014).

Variables to consider in future studies include student and teacher responses to the diagnostic data and resultant interventions, as well as parent and student expectations and attitudes about the assessments and the outcomes produced. Professional development activities related to fidelity of implementation of the online system will be necessary to assist teachers and school leaders in learning how to interpret the results from the programs' diagnostic assessments, formative/summative assessment results, and any indices of classroom instructional adjustments. The type of data proffered by this computer driven assessment could allow schools to set goals, prioritize resources, and make intervention plans. The data can also provide schools with a measurable platform for the development of or improvement towards an efficient decision-making process.

Decisions based on data could inform districts on how to proceed with conventional technology concerns such as, financial commitment, IT infrastructure, implementation and sustainability plans, and professional development needs (Slavin et al., 2012). An advantage to the implementation of an online formative assessment system is that such a system provides a measurable, ongoing examination of student progress towards expected learning outcomes and state or federal standards. Reports generated from such programs provide relevant and timely information on the district, and individual school level improvement can help districts grapple with budget issues and concerns that impact FTE, sectioning and staffing, as well as administrative costs.

Additional research along this vein could contribute to an evaluation framework for the development, implementation and sustainability of online formative and diagnostic software programs. Such a framework could provide a data-initiated checks and balance system to assist policy makers, state education agencies; district level administrators make informed decisions before purchasing or implementing an online formative/diagnostic/intervention program. Technology and data accrued from programs such as these cannot replace teachers; however, they have the potential to redefine teaching and learning and move educators towards becoming facilitators of knowledge in an educational system that fosters critical thinking while encouraging creativity.

Implications for Change

As the literature review noted, studies related to the connections between online formative and summative assessments are limited. Nonetheless, with the increase in legislative pressure (i.e., No Child Left Behind, NCLB) for schools to improve teaching and learning through the use of data, it is surprising that studies continue to show that schools and districts have been slow to find, adopt, and implement online assessment programs (Slavin et al., 2013). Altering schools, along with the educational community in general, will require a shift towards a culture of collecting, interpreting and disseminating data. This data needs to come from measures of student learning, teacher perception, demographic needs and examination of individuals and school culture and beliefs about the teaching and learning process itself (Slavin et al., 2012).

Conclusion

In conclusion, the results of this study supported the need for future research on online formative assessment programs and their impact on end-of-year assessments. Additionally, those in the educational arena should recognize that such programs play an integral role in a comprehensive assessment system.

It is important for districts and policy makers to understand that disseminating and delivering timely, thorough, and useful information from any data sources relating to teaching and learning directly to the classroom teacher, can and should influence student achievement. However, continued improvement in student achievement requires districts and policy makers to move towards capitalizing on evidence-based research to find, adopt, and implement online formative assessment programs that provide targeted data. However, this will require that administrators quickly distribute that data into the hands of teachers, in order to help them sustain long-term improvements in every aspect of the teaching and learning process

Therefore, more research on the relationship of formative online assessments and end-of-year summative assessments is needed, along with the analysis and dissemination of information. In addition, research to develop an evaluation framework that could illustrate a pattern of strong evidence on the effectiveness of any online formative assessment program towards improving student achievement is needed.

References

Alonzo, J., Nese, J., Park, B., & Tindal, G. (2011). Applied curriculum-based measurement as a predictor of high-stakes assessment. *Elementary School Journal*, *111*(4), 608-624.

Anderson, T., & Kanuka, H. (2007, June 25). Using constructivism in technology-mediated learning: Constructing order out of the chaos in the literature. *Radical Pedagogy (1999)*. Retrieved from http://auspace.athabascau.ca/bitstream/2149/728/1/Using%20Constructivism%20in%20Technology-Mediated%20Learning_%20_br_Constructing%20Order.pdf

Baker, R. D., Goldstein, A. B., & Heffernan, N. T. (2011). Detecting learning moment-by-moment. *International Journal of Artificial Intelligence in Education*, *21*(1-2), 5-25.

Bax, C., Branford-White, C. Heugh, S. ,& Jacoby, J., (2013). Enhancing learning through formative assessment. *Innovations in Education and Teaching International*. doi:10.1080/14703297.2013.771970

Biemans, H., Gulikers, J., Van der Wel, M., & Wesselink, R. (2013). Aligning formative and summative assessments: A collaborative action research challenging teacher conceptions. *Studies in Educational Evaluation*, *39*, 116-124. doi: http://dx. doi. org.ezp.waldenulibrary.org/10.1016/j.stueduc.2013.03.001

Bennett, R. (2010). Cognitively based assessment of, for, and as learning (CBAL): A preliminary theory of action for summative and formative assessment. *Measurement*, *8*(2/3), 70-91. doi:10.1080/15366367.2010.508686

Bennett, R., Persky, H., Weiss, A. R., Jenkins, F., National Center for Education Statistics (ED), W. C., Educational Testing Service, P. J…. Westat, I. D. (2007). *Problem solving in technology-rich environments. A Report from the NAEP technology-based assessment project, research and development series.* NCES 2007-466. National Center For Education Statistics Washington, DC.

Bernhardt, V. L. (2003). No schools left behind. Educational Leadership, (5), 26–30.

Black, P., & William, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation & Accountability*, *21*(1), 5-31. doi:10.1007/s11092-008-9068-5

Black, P., & William, D. (2003). In praise of educational research: Formative assessment. *British Educational Research Journal*, *29*(5), 623.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (Eds.) (1971). Handbook on the formative and summative evaluation of student learning. New York: McGraw-Hill.

Bloom, B., Engelhart, M., Furst, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: David McKay Company.

Brosvic, G., Calvano, T., Hendel, R., Epstein, B., Epstein, M., Lazarus, A., Mathews, K., & (2002). Immediate feedback assessment technique promotes learning and corrects inaccurate first responses. *The Psychological Record*, *52*(2), 187-201.

Burns, M. K., Klingbeil, D. A., & Ysseldyke, J. (2010). The effects of technology-enhanced formative evaluation on student performance on state accountability math tests. *Psychology in the Schools*, *47*(6), 582-591.

Carlson, D., Borman, G. D., & Robinson, M. (2011). A multistate district-level cluster randomized trial of the impact of data-driven reform on reading and mathematics achievement. *Educational Evaluation and Policy Analysis, 33*(3), 378-398. doi: : 10.3102/0162373711412765

Chappuis, S., Stiggins, R., Arter, J., & Chappuis, J. (2005). Part 1 Building the foundation. In *Assessment for learning: An action guide for school leaders* (2nd ed., p. 3). Portland, OR: Educational Testing Service.

Chien, S., Wu, H., & Hsu, Y. (2014). An investigation of teachers' beliefs and their use of technology-based assessments. *Computers in Human Behavior, 31*, 198-210. doi: 10.1016/j.chb.2013.10.037

Common Core State Standards Initiative. (2012). In the States. *Common Core State Standards Initiative*. Retrieved from http://www.corestandards.org/in-the-states

Curriculum Associates (2011). Add title of what was retrieved. Retrieved from http://www.i-ready.com/

Deno, S. L. (2003). Developments in curriculum-based measurement. *Journal of Special Education, 37*(3), 184-192.

Dewey, J. (1934). *Experience and education*. New York: Simon & Schuster.

Dron, J. (2012). The pedagogical-technological divide and the elephant in the room. *International Journal on E-Learning, 11*(1), 23-38. Chesapeake, VA: AACE.

Duffy, T. M., & Jonassen, D. H. (1992). *Constructivism and the technology of instruction: A conversation*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Eggen, T., Timmers, C., Van der Kleij, M., & Veldkamp, P. (2012). Effects of feedback in a computer-based assessment for learning. *Computers & Education*, *58*(1), 263-272. doi: 10.1016/j.compedu.2011.07.020

Eisner, E. (2000). Benjamin Bloom. *Prospects: The Quarterly Review of Comparative Education*, *30*(3),1-7. Retrieved from http://www.ibe.unesco.org/publications /ThinkersPdf/bloome.pdf

Forster, N., & Souvignier, E. (2011). Curriculum-based measurement: Developing a computer-based assessment instrument for measuring student reading progress on multiple indicators. *Learning Disabilities: A Contemporary Journal*, *9*(2), 65-88.

Fournier, H., Sui Fai Mak, J., & Kop, R. (2011). A pedagogy of abundance or a pedagogy to support human beings? Participant support on massive open online courses. *The International Review of Research in Open and Distance Learning*, *12*(7), 74-93.

Fuchs, L. S. & Stecker, P. M., (2000). Effecting superior achievement using curriculum-based measurement: The importance of individual progress monitoring. *Learning Disabilities Research & Practice*, *15*(3), 128-134.

Gao, P., & Mager, G. (2013). Constructing embodied understanding of technological pedagogical content Knowledge: Preservice teachers' learning to teach with information technology. *International Journal of Social Media and Interactive Learning Environments*, *1*(1), 74-92. doi: 10.1504/IJSMILE.2013.051654

Gardner, H. (1983, Summer). Beyond the IQ: Education and human development. *Harvard Education Review*, *57*(2), 187-196.

Gardner, H. (1977) *Frames of mind.* New York,: Basic Books.

Gong, B, Perie, M., & Marion, S. (2009). Moving toward a comprehensive assessment system: A framework for considering interim assessments. *Educational Measurement: Issues & Practice*, *28*(3), 5-13. doi: 10.1111/j.1745-3992.2009.00149.x

Greene, J. P., Trivitt, J. R., Winters, M. A., & (2009). The impact of high-stakes testing on student proficiency in low-stakes subjects: Evidence from Florida's elementary science exam. *Economics of Education Review*, *29*, 138-146. doi: 10.1016/j.econedurev.2009.07.004

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The Effect of School Resources on Student Achievement. *Review of Educational Research, 66*(3), 361-396. doi: 10.3102/00346543066003361

Hannafin, R. & Foshay, W. (2008). Computer-based instruction's (CBI) rediscovered role in K-12: An evaluation case study of one high school's use of CBI to improve pass rates on high-stakes tests. *Educational Technology Research and Development*, *56*(2), 147-160. doi: 10.1007/s11423-006-9007-4

Harlen, W. (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *Curriculum Journal*, *16*(2), 207-223. doi: 10.1080/09585170500136093

Hintze, J. M., Keller-Margulis, M.A., & Shapiro, E.S., (2008). Long-term accuracy of curriculum-based measures in reading and mathematics. *School Psychology Review, 37*(3), 374-390.

Hixson, M. D., & McGlinchey, M. T. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review, 33*(2), 193-203.

Hwang, J., McMaken, J., Porter, A. & Yang, R. (2011). Assessing the common core standards: Opportunities for improving measures of instruction. *Educational Researcher*, *40*(4), 186-188. doi: 10.3102/0013189X11410232

Jimoyiannis, A., & Komis, V. (2007). Examining teachers' beliefs about ICT in education: Implications of a teacher preparation programme. *Teacher Development*, *11*(2), 149-173. doi: 10.1080/13664530701414779

Johannessen, O., & Redecker, C. (2013). Changing assessment — Towards a new assessment paradigm using ICT. *European Journal of Education Research, Development and Policy*, *48*(1), 79-96. doi: 10.1111/ejed.12018

Johnson, D. C. (2008). Thinking critically about assessing online learning. *International Journal of Learning, 14*(12), 125-130.

Kuiper, W., & Schildkamp, K., (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education*, *26*(3), 482-496. doi: 10.1016/j.tate.2009.06.007

Lam, R. (2013). Formative use of summative tests: Using test preparation to promote performance and self-Regulation. *Asia-Pacific Education Researcher*, *22*(1), 69-78. doi: 10.1007/s40299-012-0026-0

Lam, Y. (2000). Technophilia vs. Technophobia: A preliminary look at why second-language teachers do or do not. *Canadian Modern Language Review*, *56*(3), 389-420.

Linan-Thompson, S., Murray, C., Roberts, G., Vaughn, S., Wanzek, J., & Woodruff, A. (2010). Differences in the relationship of oral reading fluency and high-stakes measures of reading comprehension. *Assessment for Effective Intervention*, *35*(2), 67-77. doi: 10.1177/1534508409339917

Pearson (2014. Stanford Achievement Tests Series, 10th edition. http://www. pearsonassessments.com/haiweb/cultures/en-us/productdetail.htm?pid=SAT10C

Popham, W. J. (2008). *Transformative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

Proulx, J. (2006). Constructivism: A re-equilibration and clarification of the concepts, and some potential implications for teaching and pedagogy. *Radical Pedagogy*, *8*(1), 1-5.

Ruiz-Primo, M. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation*, *37*(1), 15-24.

Saettler, L. P. (1990). *The evolution of American educational technology*. Englewood, CO: Libraries Unlimited.

Spector, J. M., Merrill, M. D., Merrienboer, J. V., & Driscoll, M. P. (Eds.). (2008). *Handbook of research on educational communications and technology* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Slavin, R. E., Cheung, A., Holmes, G., Madden, N. A., & Chamberlain, A. (2013). Effects of a data-driven district reform model on state assessment outcomes. *American Educational Research Journal, 50*(2), 371-396. doi: 10.3102/0002831212466909

Smith, P. A., & Hoy, W. K. (2007). Academic optimism and student achievement in urban elementary schools. *Journal of Educational Administration, 45*(5), 556-568. doi: 10.1108/09578230710778196

Stiggins, R. J., Arter, J. A., Chappuis, J., & Chappuis, S. (2006). Classroom assessment for student learning: doing it right – using it well. Princeton, NJ: Educational Testing Service.

Supovitz, J. (2006). Uses and misuses of data for educational accountability and improvement. *Journal of Educational Change*, *7*(4), 351-354. doi: 10.1007/s10833-006-9005-7

Supovitz, J. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform. *Journal of Educational Change*, *10*(2/3), 211-227. doi: 10.1007/s10833-009-9105-2

Taras, M. (2010). Back to basic: Definitions and process of assessments. *Revista Práxis Educativa*, *5*(2), 123-130.

Taras, M. (2009). Summative assessment: The missing link for formative assessment. *Journal of Further and Higher Education*, *33*(1), 57-69. doi: 10.1080/03098770802638671

Tsuei, M. P. (2007). A web-based curriculum-based measurement system for class-wide ongoing assessment. *Journal of Computer Assisted Learning*, *24*, 47-60. doi: 10.1111/j.1365-2729.2007.00242.x

Tucker, B. (2009). The next generation of testing. *Educational Leadership*, *67*(3), 48-53.

U.S. Department of Education, Institute of Education Services, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 and 2011 Mathematics Assessments.

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

Watson, D. M. (2001). Pedagogy before technology: Re-thinking the relationship between ICT and teaching. *Education and Information Technologies, 6*(4), 251-266. doi: http://dx.doi.org/10.1023/A:1012976702296

Wilson, B. (February 2010). Constructivism in practical and historical context. (Draft chapter) for inclusion in Bob Reiser & Jack Dempsey (Editors), Current Trends in Instructional Design and Technology (third edition) Retrieved from http://carbon.ucdenver.edu/~bwilson/Constructivism.pdf

Zoanetti, N. (2010). Interactive computer based assessment tasks: How problem-solving process data can inform instruction. *Australasian Journal of Educational Technology*, *26*(5), 585-606.

Appendix A: Written Confirmation for Sharing of Data

Date: October 11, 2012 7:47:31 AM MST
Cc: Jacqueline Croteau <jacquelinecroteau@me.com>
Subject: **Written Confirmation on sharing SAT10 Data**

Jodi,

This email will serve as written approval and permission to share NLCI SAT 10 and i-Ready data with Jackie Croteau in support of her Doctoral Thesis and for no other purposes.

Please make certain that Jackie knows that she may only share and/or distribute this information for purposes of development, presentation and/or review of her dissertation. In addition, Jackie is prohibited from using individual student names or associating any specific data or assessments with any identified or identifiable students. All data utilized must remain anonymous as to both students and specific school.

Please let me know if you need any additional approvals or permissions.

Appendix B: SAT-10 Correlations

Numerous unsuccessful attempts were made to obtain copies of the SAT-10 Correlations. The only provision made was to view the Correlations documents at the Philadelphia corporate headquarters for Nobel Learning Communities. In addition, Nobel was told they would have to pay for the documents and oversee my access to the information. Since SAT-10 has been proven to be a valid measurement for reading and math, it was decided that the missing SAT-10 Correlations information would not hinder the proposed study.

Appendix C: iReady Correlations

Complete copies of iReady correlations in Reading and Mathematics were made available to me; however the file is too large to attach to this document. The files were received from Woody Paik | Vice President Curriculum Associates LLC 978-495-6858 | wpaik@CAinc.com http://www.CurriculumAssociates.com

Curriculum Vitae

Contact Information

Jacqueline L. Croteau,

3982 S. Alexandrite Ave.

Tucson, AZ 85735

Education

Walden University (GPA 3.79)

University of Phoenix, 1999 (GPA 3.85)

Prescott College, 1995 (GPA 3.5)

Education and Certifications

PhD. Candidate In Educational Technology

MAED Administration and Supervision

BA Secondary Education

English

English as a Second Language

AZ Principal Certificate

AZ Teaching License

AZ SEI Certified Instructor

Dissertation Topic

Online Formative Assessments as Predictors of Student Academic Success

Technical & Professional Skills

Microsoft Certified Systems Engineer Certification (MCSE)

Educational technology tools

Research

Data analysis

SPSS analysis

SharePoint

Professional development: develop, create and present

Microsoft Office Suite

Online Learning Management Systems

Interactive online communication technology

Digital media