To the Graduate Council:

I am submitting a dissertation written by Wendy L. Oliver entitled, "Investigating Whether a Value-Added Teaching Effectiveness Model Designed for Traditional Classrooms Can be Used to Measure Online Teaching Quality." I have examined the electronic copy of this dissertation and recommend that it be accepted in partial fulfillment of the requirements for the Doctor of Education, with a major in Learning and Leadership.

_____
Dr. M. D. Roblyer, Chairperson

We have read this dissertation
and recommend its acceptance.

_____
Dr. Hinsdale Bernard

_____
Dr. Kirk Kelly

_____
Dr. Deborah McAllister

Accepted for the Graduate Council:

_____
Dr. Stephanie Bellar
Interim Dean of the Graduate School

INVESTIGATING WHETHER A VALUE-ADDED TEACHING EFFECTIVENESS MODEL DESIGNED FOR TRADITIONAL CLASSROOMS CAN BE USED TO MEASURE ONLINE TEACHING QUALITY

A Dissertation
Presented for the
Doctor of Education Degree
The University of Tennessee at Chattanooga

Wendy L. Oliver
April 2010

Dedication

To the students, administrators, and teachers of Tennessee who participate in nontraditional learning environments. I have been honored to have the opportunity to work with you during the creation of Tennessee's online learning program.

Acknowledgements

I am overwhelmingly grateful to Dr. M.D. Roblyer for serving as my dissertation chair and dedicating countless hours supporting and advising me throughout the dissertation process. I have been very fortunate to apprentice under someone who is an active researcher in the field of online learning. Dr. M.D. Roblyer's perseverance and expertise in the field of online learning have made my study possible.

My dissertation committee, Dr. M.D. Roblyer, Dr. Hinsdale Bernard, Dr. Kirk Kelly, and Dr. Deborah McAllister, have been supportive as ideas and data challenges extended their efforts and commitments, necessitating unconventional thinking and utmost flexibility.

I am thankful for guidance, patience, and, most importantly, support and understanding from the EdD faculty, as I discovered how to overcome a very difficult personal challenge while balancing doctoral level work and a blossoming career.

I am grateful to Mr. Dan Long for his support. I am amazingly fortunate to have a mentor who allows me the opportunity to construct leadership and advises me of guidance while allowing me to learn. Without his support and leadership online learning would not be at the advanced stage it is in Tennessee.

I am appreciative of the e^{4}TN staff, my family, and my friends who have carried on and provided support while I have focused time and energy on the experience of constructing this dissertation. It has been a remarkable learning experience.

Abstract

The study reported here focused on assessing teacher quality in online environments. The purpose of the study was to explore the feasibility of using the same method Tennessee currently uses to gauge teaching quality of traditionally-delivered courses to determine teaching quality in the online environment. Research questions were:

1) Is there a significant difference in program effects of traditional classrooms (as measured by end-of-course scores for a sample of traditionally-taught students in a Tennessee school district) and online classrooms (as measured by end-of-course scores for a sample of Tennessee's online students)?

2) Do program effects between traditional and online environments vary significantly by subject area (i.e., Algebra I, Biology, and English 1)?

3) Do Tennessee educators perceive that Tennessee's model for teacher-effect scores can be used equally well in both traditional and online environments?

4) What factors and strategies do educators perceive should be considered in determining teaching quality in the traditional and online environments?

Research questions 1 and 2 were addressed by comparing EOC scores from students in each program. EOC scores from 162 students in a Tennessee online program were compared with a sample of 162 students from a Tennessee school district that were systematically selected to match the online sample in several important characteristics (e.g., socio-economic levels, indicators of prior achievement). A regression analysis was used to identify variables that contributed significantly to students' EOC scores, and effects of the two programs were compared by using an Analysis of Covariance

(ANCOVA) to control for contributions of these variables. With these variables controlled, no significant differences existed between online and traditional programs in any content areas, except in Algebra I when only NCE scores were considered as a covariate.

To address research questions 3 and 4, 68 Tennessee educators completed a survey with open-ended and Likert-scale items. Survey data indicated a lack of understanding for Tennessee's teacher-effect model and a general perception that traditional teacher quality indicators cannot be used to assess teachers in the online environment. Implications of these findings and directions for future research are discussed.

Table of Contents

List of Tables

Tables **Page**

Chapter 1: Introduction to the Study

Statement of the Problem

Distance learning, in particular online learning, is a growing medium for secondary education in American high schools (Roblyer, 2006; Tucker, 2007; Watson & Ryan, 2007; Zandberg & Lewis, 2008). Online learning is an interactive learning environment in which the curriculum is delivered via the Internet with the instructor communicating through a variety of methods such as email, telephone, chat, text, discussion boards, or virtual classrooms (Tucker 2007; Watson & Ryan, 2007). As more students and schools venture into online learning in order to provide equitable access to courses and to meet the needs of students, quality of online instruction is of great concern to educators (SREB, 2006).

With the reauthorization of the 1965 Elementary and Secondary Education Act (commonly called the No Child Left Behind Act of 2001 or NCLB), the focus on what constitutes effective instruction, already a controversial debate, has been reframed (Crane, 2002). The NCLB Commission changed the law to read "highly-effective, qualified teachers" rather than "highly-qualified teachers." This represents a shift in thinking about measuring quality of instruction, moving the emphasis from teacher qualifications (an input variable) to teacher products and outcomes (an outcome variable). It also means that a measurement for effective instructors must be designed and put into place (Hammond & Prince, 2007).

Much attention has been paid to various measures of teacher effectiveness. A value-added model is one such measure. Value-added testing models allow administrators to

measure student gains by year. In other words, they are a statistical measure of what a particular teacher has contributed to a student's learning value based on predicted test scores and data from previous years of testing (Education Week, 2004).

The Tennessee version of a value-added model is the Tennessee Value-Added Assessment Systems or TVAAS. The TVAAS model, created by Sanders (Hammond & Prince, 2007; Olson, 2005) provides a score that indicates how much of an effect teachers have had on students. Teacher-effect scores seek to provide a quantifiable way to identify high-quality teachers, serving as a basis for selecting teachers or offering a way to implement an evaluation tool to identify high-performing teachers (i.e., master teachers) and allowing low-performing teachers to receive further training or professional development specific to their needs for the purpose of improving student achievement (Hammond & Prince, 2007; Olson, 2005).

Using the longitudinal, general linear model, which is unavailable because it is now copyrighted to SAS Institute Inc., SAS calculates Tennessee TVAAS scores. For example, if the mean predicted TVAAS score using Gateway exams for high school English 10 in X district is 525.0, but the actual mean TVAAS score students earned is 522.1, then the teacher-effect score is -2.9, which means that the teacher did not add to increased learning. Rather, the students lost academic performance based on longitudinal data on predicted scores (K. Kelly personal interview, July 23, 2009).

Tennessee data suggest that high-quality teachers produce considerably increased learning gains in their students as compared with weaker instructors. This conclusion is based on value-added assessments in the traditional classroom (Deubel, 2008; Education Week, 2004). Determining what makes a teacher “high-quality” is highly debated. Trends

in data seem to indicate the following factors produce high quality teachers: an educational focus in the content area, especially in science and math; length of teaching experience; strong ability to communicate in both the written and verbal format; understanding of how students learn (pedagogical knowledge); passion for the content being taught; and basic abilities in reading, writing, math (Deubel, 2008; Education Week, 2004; Hammond & Prince, 2007; Pierce, 2008; Roblyer & McKenzie, 2000; Where we stand on teacher quality, 2004; Yang, 2005). TVAAS predicts reliably with the extreme scorers, but there may be disparities within the mid-range scorers, as statistical measurements always have a margin of error. TVAAS refers to the mid-range scorers as "non-detectable difference" (Crane, 2002). Additionally, there is no guarantee that extreme high scorers are not teaching to the test. However, it appears to be the best statistical test currently available to measure objective student gains, which is the outward, measurable indicator of an effective or high-quality instructor (Crane, 2002).

Research shows that high-quality online instructors have at least some of the same characteristics as high-quality traditional online instructors. For example, both must present a positive attitude and have very high verbal and written communication skills; both must know each student's learning style and have a strong knowledge of content (Deubel, 2008; Education Week, 2004; Hammond & Prince, 2007; Pierce, 2008; Roblyer & McKenzie, 2000; Where we stand on teacher quality, 2004; Yang, 2005).

However, it is generally assumed that traditional teaching quality does not necessarily predict online teaching quality (Wood, 2008). Standards for online teaching quality have been prepared by various organizations (SREB, 2006; NACOL, 2008; Trotter, 2008), but currently there are no studies that compare teaching quality in the two

environments. Consequently, there is no way to confirm or refute the commonly-held assumption that teachers' abilities in traditional classes cannot predict their abilities in online courses. The purpose of this study is to explore the feasibility of using a value-added method to gauge the quality of online instruction by comparing student mean scores in traditionally-delivered courses with courses delivered in the online environment. If a significant correlation exists in student performance, then value-added findings for teachers should be consistent in either area.

Background on the Problem

There are many ways in which teacher ability and effectiveness are currently measured including, but not limited to: observation, interview, portfolio, and standardized tests. All of these have inherent limitations. Qualitative measures leave opportunities for administrative bias, while standardized testing scores have margins for error. Standardized testing, however, provides a quantifiable and measurable result that can be tied directly to student gains (Alliance for Excellent Education, 2008; Braun, 2005; Elmore, 2002; Gore, 2007). Most educators agree that student improvement is an important indicator of effective instruction; however, determining how to measure a teacher's effectiveness in terms of student improvement is an area upon which experts rarely agree (Alliance for Excellent Education, 2008; Darling-Hammond, 2007; Gordon, Kane, & Staiger, 2006).

Using its Tennessee Value Added Assessment Score (TVAAS), the State of Tennessee has found a correlation between student achievement and teacher effectiveness, asserting that " … students given the most effective teacher for three years in a row made over twice the gains of comparable students assigned the least effective

teachers" (Alliance for Excellent Education, 2008, p. 2; Sanders & Rivers, 1996). Therefore, the State of Tennessee defines teacher effectiveness in terms of student gains.

In application to the K-12 online learning environment, there is no formal model to assess online instructors. However, with the continued teacher shortage, budget cuts and a growth of 30% per year nationally in online learning, a formal evaluation method for online teachers is necessary (NACOL, 2008). Although there is significant concern that online learning is less rigorous than traditional learning, online students are required to take the same end-of-course examinations as they do in other courses, according to NCLB. The fact that there is no way to determine equivalence in any meaningful way between the effectiveness of online and traditional teachers is a problem, especially when so many brick-and-mortar schools are turning to online learning.

Definition of Terms

The following terms will be used throughout this dissertation. They are given here to clarify meaning of words and phrases related to education and online learning.

1. Asynchronous - A learning situation where the learner and instructor are not interacting at the same time. For example, the student may post messages at 5:00 p.m., and the teacher may respond to messages in the same online area at 10:00 a.m.
2. AYP - Adequate Yearly Progress - With the implementation of the No Child Left Behind (NCLB) Act of 2001, all students are required to make Adequate Yearly Progress based on standardized test scores, with increased requirements each year (Kupermintz, 2003).
3. Distance education – According to the U. S. Distance Learning Association (USDLA), "The organizational framework and process of providing instruction at a

distance. Distance education takes place when a teacher and student(s) are physically separated, and technology (i.e., voice, video, data, or print) is used to bridge the instructional gap." (USDLA, n.d.)

4. ELL – English Language Learners are students who speak a native language other than English, and are acquiring English as their second, third, and sometimes fourth language (Eckes & Law, 2000).
5. Effective instruction - Instruction that allows students to improve by showing gain in the intended curriculum as defined by Eisner (2002).
6. NCE – According to the Educational Consumer's Foundation, "A test score reported on a scale that ranges from 1 to 99 with an average of 50. NCE's are approximately equal to percentiles. For example, an NCE of 70 is approximately equal to or greater than 70% of its reference group. Assuming a normally distributed population, plotting the distribution of scores will result in a bell shape commonly known as a bell curve." (Glossary of Terms and Abbreviations, n.d., p.1).
7. NCLB - The No Child Left Behind Act of 2001 is federal legislation, which became effective in 2002, requires all states to report student academic progress using standardized achievement tests (Kupermintz, 2003).
8. Online learning - "Education in which instruction and content are delivered primarily via the Internet. Online learning is a form of distance learning" (Ryan & Watson, 2006, p. 134).
9. Synchronous learning - A learning situation that takes place in real time with one or more instructors. Participants are logged on and interact at the same time (Dwyer, 2008).

10. Teacher effect - The average class effect, which is determined by: (a) district average for that specific school year and specific grade; (b) class or teacher effect for the specific grade or specific year; (c) system or unsystematic variations for that specific class and specific year; and (d) teacher effect for the previous year (Braun, 2004).
11. TVAAS - Tennessee Value-Added Assessment Score. An algorithm is used to calculate the Tennessee Value-added Assessment Score, which measures student achievement gain from year to year. This score is used to calculate the teacher's effectiveness in courses with standardized end of course assessments (Sanders & Rivers, 1996).
12. Value-added Model (VAM) - A statistical analysis to measure student academic gain over one year rather than student achievement that takes into consideration factors other than the isolation of student gain (Ballou, 2002; Braun, 2005).
13. Virtual school – "Instruction in which (K-12) students and teachers are separated by time and/or location and interact via computers and/or telecommunications technologies" (National Forum on Educational Statistics, 2006, p. 1).

Significance of the Study

This study has significance for the field of virtual learning and for the education system in which virtual schooling plays an increasingly greater role. Findings will shed light on whether or not educational systems can use traditional indicators of teacher quality such as value-added scores to select effective online instructors. This information could be useful to both Tennessee and other states that are searching for such measures to guide selection of online teachers.

If the program effects between traditional and online programs are not significantly different, then this study provides support for online programs to use the same value-added scores for assessing teacher quality in online courses as they now do for traditional courses. States might also consider using their own, specific, current measures of teaching quality to select their online teachers. However, if the program effects are significantly different between traditional and online courses, this is an indication that another method for assessing teacher quality in online courses must be found. Differing effect scores between traditional and online courses would fail to provide needed evidence for or against the fact that online teaching effectiveness is different from traditional teaching effectiveness. Instead, it would indicate that differing conditions between the two environments call for different ways of measuring teaching quality in online courses and traditional courses.

If results between the two programs are similar, the study will provide data to support a more standardized, easy-to-calculate way of measuring teacher quality, a practice that has the potential to increase the quality of online programs and student instruction. Teacher quality data will allow administrators to make data-driven decisions to drive professional development plans and hiring decisions. For example, better methods of evaluating teacher quality will assist administrators in deciding who will be the most effective teacher mentors and curriculum or pedagogy specialists.

Research Questions

As online learning continues to grow, it is necessary to find a method for evaluating instruction to comply with NCLB requirements for highly-qualified, effective teachers (Crane, 2002; SREB, 2006). Tennessee has found a correlation between student

achievement and teacher effectiveness, asserting that " … students given the most effective teacher for three years in a row made over twice the gains of comparable students assigned the least effective teachers" (Alliance for Excellent Education, 2008; Sanders & Rivers, 1996). This study will provide evidence to determine whether such a relationship can be established in the same way with Tennessee's online instructors. The following research questions focus on the feasibility and practicality of using teacher quality indicators as evidence of online teaching quality.

1. Is there a significant difference in program effects of traditional classrooms (as measured by end-of-course scores for a sample of traditionally-taught students in a Tennessee school district) and online classrooms (as measured by end-of-course scores for a sample of Tennessee's online students)? This question seeks to address whether or not online teachers can be assessed using the same value-added scores as teachers in traditional classrooms or if they need alternative ones. Significant differences in program effects would indicate the presence of differing learning environment conditions (e.g., a different variety of learners in online classrooms vs. traditional classrooms, differing pedagogical requirements for effective online teaching than in a traditional classroom). For example, students in an online classroom are often there because they have not succeeded in traditional classrooms and, thus, may have a lower-than-average expectation of gain, based on their past performance. Consequently, their program-effect scores would be higher online than average students in traditional environments.
2. Do program effects between traditional and online environments vary significantly by subject area (i.e., Algebra I, Biology, and English 1)? According to data from past

evaluations, students seem to do better online in some subject areas than others (Florida Tax Watch Report, 2008). This study will provide evidence on whether or not this trend is reflected in data from a Tennessee online program and, consequently, if it can be expected to be reflected in program effect data.

3. Do Tennessee educators perceive that Tennessee's model for teacher-effect scores can be used equally well in both traditional and online environments? If a value-added model is planned for use in online courses, it would be helpful to have the support of Tennessee educators for its use. Negative perceptions could affect recruitment of online teachers. For example, they may feel that they should not be assessed in the same way online as they are in the traditional environment and, therefore, would not be as likely to apply for such teaching opportunities.
4. What factors and strategies do teachers perceive should be considered in determining teaching quality in the traditional and online environments? Legislation related to online learning is in the infancy stages in Tennessee. The first piece was passed on August 22, 2008. As the State of Tennessee becomes more aware of both the challenges and opportunities offered by online learning, this study will help identify educator perceptions of problems and potential concerns prior to the model's implementation, should the state choose to implement teacher effect scores with online instruction.

Chapter Summary

As educators and policymakers realize the benefits of the flexibility of online learning, student enrollments increase daily. As more states develop state-level virtual

schools and district-level programs increase, there is a growing need for accountability and quality indicators for online instruction (Watson, Gemin, Ryan & Wicks, 2009).

There are a variety of ways to measure the effectiveness of instruction including observation, portfolio, student-feedback and peer-feedback. A quantitative method for measuring the effectiveness of a teacher in the traditional classroom is commonly referred to as teacher-effect scores or value added. Teacher effect scores are based on several factors, and ultimately they indicate whether a student gains, maintains or loses knowledge from year to year as a result of what teachers do (Alliance for Excellent Education, 2008; Braun, 2005; Elmore, 2002; Gore, 2007). With such rapid growth and continued expansion by states and school districts in online learning, a formal model for measuring teacher effectiveness in the online medium is necessary (NACOL, 2008).

The purpose of this study is to explore the feasibility of using the same value-added scores currently used to gauge the quality of traditional instruction to assess quality of online instruction by comparing programs effects. If samples of students have equivalent performance in traditional and online classrooms (Algebra I, Biology and English I), then programs are similar in effects and, therefore, similar measurements to assess teacher effectiveness could be used. If program effects differ, then variables outside those usually used to measure teacher effect are having an effect on students.

Four research questions were investigated to determine if a relationship between traditional and online teachers indicates that the same value-added methodologies used for assessing quality of traditional teachers can be used for online teachers. Research question 1 examines if there is a significant difference in program effects of traditional classrooms and online classrooms in Algebra I, Biology and English I. Research question

2 investigates whether program effects between traditional and online environments vary significantly by subject area (i.e., Algebra I, Biology, and English 1). Research questions 3 and 4 assess educators' perceptions of value added as a model to measure teacher effectiveness online and what factors they perceive should be considered in determining teaching quality in the traditional and online environments.

If results indicate that the programs are similar, it would suggest that the same methods can be used to assess both traditional and online teachers. However, significantly-different results in program effect for online and traditional classrooms would indicate that teacher quality indicators in the online environment should be assessed using a different model than those used with traditional teachers.

Chapter 2: Review of Related Literature

Background on Virtual Schooling

What is virtual schooling? Prior to the emergence of a graphic interface for the Internet, commonly known as the World Wide Web (WWW), distance learning was delivered via correspondence courses and, later, videotapes and broadcast radio and television (Berg, 2002). Within a 25-year period, this significant technological growth enhanced educational opportunities throughout the world (Zucker & Kozma, 2003). Before the WWW, distance learning was managed through mail, video-based courses, and extension services (Zucker & Kozma, 2003).

The growth of the Internet, while far different now from its original purpose, has significantly influenced our society (Roblyer & Doering, 2010). The Internet grew out of a project called ARPAnet, originally funded by the U. S. Department of Defense in the 1970's (Roblyer & Doering, 2010; Kahn, 1994). The original intent of the Internet was to create a medium for communication for DOD researchers working on projects in 30 locations (Roblyer & Doering, 2010). As a result of ARPAnet's standard communication protocol, established in 1971, major breakthroughs such as email file exchanges were made possible. In the 1980's, when desktop computers were growing in popularity, "The National Science Foundation funded a high-speed connection among university centers based on the ARPAnet structure. By connecting their individual networks, universities could communicate and exchange information in the same way the DOD's projects had" (Roblyer & Doering, p. 210). With the development of the graphical web browser (Mosaic) in 1993, the sharing of resources and network exchanges became more common, paving the way for virtual learning (Roblyer & Doering, 2010; Kahn, 1994).

While there are many differences among individual institutions' requirements and delivery systems, virtual schooling allows students to learn without attending class and with a much higher opportunity and expectation for student-to-teacher and teacher-to-student interaction than ever before (Furey & Murphey, 2005; Russell, 2004). Virtual learning uses the Internet and other forms of distance delivery (e.g., videoconferencing) to provide course instruction to students (Florida Tax Watch, 2007). Prior to 1996, online learning was limited primarily to postsecondary institutions. However, legislators realized the opportunities online learning could provide for K-12 students and in 1995 began to provide high quality education via the Internet through federally-funded programs such as *Preparing Tomorrow's Teachers to Use Technology, Learning Anytime Anywhere Partnerships, Technology Literacy Challenge Fund,* and *Technology Innovation Challenge Grants* (Zucker & Kozma, 2003, p. 7).

The main differences in virtual schooling and traditional schooling, according to Hassel and Terrell (2004), are location of the actual educational resources and accessibility as a result of the multimedia medium for content delivery in comparison to traditional pedagogy. The *Keeping Pace with K-12 Online Learning: An Annual Review of State Level Policy and Practice* report (2008) recognizes the geographic barriers that online learning overcomes but points out that online courses are teacher-led.

There are several terms commonly used interchangeably for online learning such as: e-learning, distance learning, networked learning, tele-learning, computer assisted learning, and web-based learning (Anderson, 2008). A virtual school is a K-12 organization offering partial or complete, government approved, web-based curriculum programs to students. Some states or local education agencies allow for full time

attendance, while others let students take one or two courses at a time (Anderson 2008; Hassel & Terrell, 2004; Russell, 2004). In virtual schooling, the instructor is separated from the student via distance, and the student uses a form of technology to interact with the instructor and construct personal meaning and learn from the experience (Anderson & Elloumi, 2008; Anderson, 2008; Cavanaugh, Gillan, Kromrey, Hess, & Blomeyer, 2004). According to Roblyer (2006), virtual learning is one of the fastest growing fields in K-12, and it is expected to grow significantly over the next ten years. Virtual learning programs can be characterized in four ways: by entities delivering courses, by organizational structure, by delivery technologies, and by curriculum delivery format.

Entities delivering courses. There are a variety of organizations delivering courses to students, including: state-based virtual schools serving in-state and out-of-state clients, post-secondary institutions, private vendors, local school districts, other school districts within the same or other states, their own district, a state-based, or via a local or district charter school (Watson, 2008; Zandberg & Lewis, 2008). State-led programs that provide full time teachers and student funding and course opportunities may be found in: Illinois, Michigan, Florida, Idaho, Kentucky, Missouri, and Georgia. State-led initiatives, which provide online resources such as a vendor clearinghouse for district use, may be found in Washington, Wyoming, Texas, and Oregon (Watson, 2008). Examples of vendors include but are not limited to: Class.com, Apex Learning, Aventa, Compass, Oddyseyware, Florida Virtual School, and SAS Institute, Inc. (Zucker & Kozma, 2003).

Organizational structure. There are five administrative structures for virtual schools: “statewide supplemental programs, district-level supplemental programs, single-district cyberschools, multidistrict cyberschools, and cybercharter schools” (NCES, 2006,

p. 2). Florida Virtual School and Georgia Virtual School are examples of statewide supplemental programs (Watson, 2008; Zucker & Kozma, 2003). A state-level agency authorizes students who are in a traditional brick-and-mortar or cyber school to enroll in these online courses, and opportunities for online learning through statewide supplemental programs are offered on a statewide basis (NFES, 2006).

District-level supplemental programs offer courses to students within a single school district, but the state does not necessarily monitor the course offerings (NFES, 2006). An example of a district-level supplemental program is Hamilton County Virtual School in Chattanooga, Tennessee (Ryan & Watson, 2008). Similarly, a single-district cyberschool also resides in and serves one district. These are schools that provide online curriculum for students with a variety of needs as a full time enrollment option (NFES, 2006).

Multidistrict cyberschools are housed in a single school district, but they serve students in a variety of schools (NFES, 2006). Cyber charter schools typically operate as multidistrict cyberschools but use commercial vendors as well (Berge & Clark, 2005; NFES, 2006; Watson, 2008; Zucker & Kozma). Charter school legislation and legislation specific to virtual schools applies to both multidistrict cyberschools and cyber charter schools (NFES, 2006).

Delivery technologies. There are a variety of technological methods that make virtual schooling possible. All users require a computer or handheld device with an Internet connection. The program the participant is using will specify software or any additional hardware needed to run activities (Delivery methods for distance education, 2007). For example, many online courses require that the user install Flash Reader. The

terms “anytime” and “any place” are often used to describe the flexibility in online learning, meaning that students can complete assignments at any time during the day or night, and there is no physical location from where they are required to take the course (Roblyer, 2006; Zucker & Kozma, 2003). Students can take courses from home, an office, or on vacation, for example. Some virtual courses, however, do have scheduled meeting times for students and teachers (Roblyer 2006; Zucker & Kozma, 2003).

Many virtual learning programs offer curriculum through a learning management system (LMS) such as Blackboard, Desire2Learn, Angel, or eCollege (Horton & Horton, 2003; Zucker & Kozma, 2003). An LMS allows administrators to manage the content, student data, knowledge sharing, and collaboration with other students and teachers, if chosen, and the platform or LMS is web accessible (Horton & Horton, 2003). Therefore, students, teachers and administrators can access courses via the Internet. Learning management systems allow for two-way communication. A course management system (CMS) (e.g., Moodle) allows for similar features of an LMS, however, it may not allow for two-way communication (Horton & Horton, 2003; Zucker & Kozma, 2003).

Students and teachers may communicate through the LMS or students may be required to submit work via fax or electronically as stated by Zucker and Kozma (2003) regarding the infrastructure of the VHS and Apex. The infrastructure of both of these programs allows the instructor to interact with students via the chatroom in the LMS. Content is delivered via slide shows and tutorials, for example, and teachers and students have a variety of ways to communicate by using the bulletin board, chat room, announcements and the discussion forums (Zucker & Kozma, 2003).

Curriculum delivery. Virtual schooling may be offered through a variety of instructional configurations. An increasingly popular one is blended learning, where a portion of the curriculum is online, but some curriculum is offered traditional with the instructor being present in person with the students. It may also be an asynchronous format, where students work completely independently of the teacher and have no scheduled "live" hours to be online. Or online curriculum may be offered in a synchronous setting, where the teacher is online "live" with the students, and students are required to login at a set time and day (Hassel & Terrell, 2004; Russell, 2004; National Center for Educational Statistics, 2005).

For example, a school district may want a synchronous virtual class of German in four schools at one time, but the district may only have one qualified German teacher. A solution, then, would be to offer a virtual course in a blended format. One teacher could broadcast lectures in a synchronous format throughout the district. If an online curriculum component were to be added to that, all students would be participating in a blended course. The State of Alabama offers a program called ACCESS that blends Internet and video-based coursework in a manner similar to that described above (Watson, Gemin, & Ryan, 2008).

History of the virtual schooling movement. Distance learning, which led to the emergence of virtual schooling, is an innovation that began in higher education. According to MacKenzie and Christensen (1968) (as cited in Berg, 2002) the first correspondence course offered in 1874 was in a language course, Chautauqua, offered by The University of Chicago's first president, William Rainey Harper. American workers and intellectuals, such as Thomas Edison took the course, which met during the summers

(Berg, 2002). A group of Baptists then began a course in Hebrew asking Harper to lead it (Berg, 2002). In 1886 they visited England and were very impressed with the University Extension movement (Berg, 2002). By 1890 there were 200 independent Chatauqua programs in the United States (Berg, 2002). Harper's efforts influenced the later-named Home-Study Department at the University of Chicago. By World War I, 12 universities had correspondence courses (Berg, 2002). The University of California at Berkley had the largest correspondence program in 1964 (Berg, 2002). In 1911, Thomas Edison released a film series for distance learning about the American Revolution (Berg, 2002). There were many training films required for WWI, which opened an opportunity for colleges and universities to offer more film based distance courses (Berg, 2002). World War II increased this demand and opportunity for video, audio and paper correspondence courses (Berg, 2002; Zucker & Kozma, 2003). Radio, by 1936, and television, beginning in the 1950's, of course, increased opportunities for one way correspondence courses (Daniel, 1995). The growth for which these two-way audio correspondence courses allowed transitioned into a new, more advanced distance learning, or the third phase of distance learning, virtual learning (Moran, 1993).

In 1980, according to Morabito (1999), the personal computer became more affordable, allowing for a new type of communication in email, forums, chats and two-way video conferencing. The University of Phoenix and On-line and Nova University began in the1970's. With the combination of the personal computer, the wide use of the Internet and the increase of correspondence courses, the push for computer-based distance learning began (Morabito, 1999).

The earliest instance of distance learning at the K-12 level was based on independent study concepts introduced by Michigan's superintendent of schools in the 1920's. According to Berge (2005), Superintendent Sydney Mitchell of Michigan integrated home-study vocational courses into traditional brick and mortar curriculum. Instructional television became popular beginning in the 1950's and remained the leading medium for distance learning until web-based courses (Taylor, 2001). In the 1990's, pioneers from the Concord Consortium in Massachusetts, Utah, Florida, and Michigan began initiatives that have shaped the current online learning initiative in the K-12 arena (Gemin, Ryan & Watson, 2008; Roblyer, 2006; Zucker & Kozma, 2003).

The Utah State Department of Education began a statewide virtual school in 1994 using state line-item funds, called the Utah Electronic High School (EHS) and a statewide online charter school, the Utah Virtual Academy. In 1995, the Concord Consortium was funded (VHS) by a federal Technology Innovation Challenge Grant. The start of the consortium included 50 charter members, staff development and co-development of content (Zucker & Kozman, 2003).

In 1996, the University of Nebraska received federal funding to write curriculum to create an online accredited high school program. This program is now private and known as Class.com (Gemin, Watson, & Ryan, 2008; Zucker & Kozman, 2003). The next virtual learning program in the U.S. and currently the largest in terms of enrollment, was the Florida Virtual School (FLVS), which began in 1997 with a $1.3 million allotment for "Break the Mold" status (Florida Tax Watch , 2007; Gemin, Watson & Ryan, 2008; Zucker & Kozman, 2003). In 2008, the program grew to more than 120,000 enrollments (Florida Tax Watch Center for Educational Performance and Accountability, 2007;

Gemin, Watson & Ryan, 2008). In 2002, FLVS was listed as a Full Time Equivalent (FTE) school of choice for parents in Florida (Gemin, Watson & Ryan, 2008). FLVS is funded based on successful completions and allows districts to franchise the model at the local level (Gemin, Watson & Ryan, 2008). Michigan's original program, founded in 1999, has led to the development of programs such as K-12, inc. in 2001 and North Carolina Virtual Public in 2007, to name a few (Berge & Clark, 2005; Gemin, Watson & Ryan, 2008).

Virtual or online learning programs have seen consistent growth since 1996 when the first programs began in the United States (Watson & Ryan, 2007; Zandberg & Lewis, 2008). As of September 2007, 42 states had "… significant supplemental online learning programs, significant full-time programs, or both," 38 of which are state led or led by state virtual school policies (Watson & Ryan, 2007, p. 18). The latest National Center for Education Statistics (2008) report on virtual schools found that in 2004–2005, there were an estimated 506,950 technology-based distance education course enrollments in public school districts (Zandberg & Lewis, 2008).

The Florida Virtual School, the first public, state-sponsored online school in the U.S. had a 50% growth in enrollment in 2007 (Ryan & Watson, 2007). The *Keeping Pace with K-12 Online Learning* report completed by the North American Council of Online Learning (NACOL) of 2007 states that Idaho Academy and Louisiana Virtual both grew approximately 18%, and the Virtual High School program grew by 24 % with students in 30 different states and 25 different countries by the conclusion of 2007.

The most common explanation of the success of the virtual learning movement in high schools is that of equity and high quality education for all students (Furey &

Murphey, 2005; Zucker & Kozma, 2003). With online courses a school district can broaden the course selection menu offered to students. Another reason school districts offer online courses is to meet requirements of the No Child Left Behind Act of 2001 (NCLB). Online learning helps school districts meet NCLB requirements for highly qualified teachers, for example.

According to the Reauthorization of the Elementary and Secondary Act in 2002 teachers must be highly qualified in the subject area they teach. If a school district does not employ a teacher who is highly qualified or if there is a teacher shortage, an online instructor with online curriculum may be hired (Berge & Clark, 2005; Paloff & Pratt, 2001). Additionally, traditional brick and mortar schools may use district, state, or private virtual schools to meet NCLB requirements that students in low performing Title I schools identified as not making AYP (adequate yearly progress) be offered virtual learning as a school of choice (Hassel & Terrell, 2004).

Proposed benefits of virtual schooling. Research and practices support the position that there are many advantages to online learning such as flexibility, increased access, engaging curriculum and greater teaching opportunities (Berge & Clark, 2005; Furey & Murphey, 2005; Hassel & Terrell, 2004; Russell, 2004). Participants can partake in online learning from any place with Internet access at any time of the day; this provides flexibility and a reduced need to travel, thereby reducing fees for training and increasing access (Russell, 2004). The flexibility of students being able to work at their own pace creates an opportunity for student-driven learning (Berge & Clark, 2005; Hassel & Terrell, 2004). Such flexibility and student choice allows for early graduation in some states, provides opportunities for advancement or solutions in hardship situations,

allows flexibility for reasons such as philosophical choices, and offers alternative means of education for students who are medical homebound or other situations where students may not have unconventional options prior to the online learning movement (Berge & Clark, 2005; Furey & Murphey, 2005).

Increased access creates an opportunity for equitable opportunities in school districts where students may not otherwise have various options to take traditional courses in upper level electives or foreign languages, for example (Berge & Clark, 2005; Furey & Murphey, 2005; Hassel & Terrell, 2004; Russell, 2004). Engaging online curriculum meets the needs of various learning styles through simulations and interactive content, where they can read and reread lessons, take tests, and students can get answers to questions individually without having to ask in front of a class (Furey & Murphey, 2005). A teacher in a brick-and-mortar situation may only appeal to auditory learners through lecture, for example. The combination of all of these creates a learning environment that can be customized to the individual learner's needs, an educational situation that is difficult to meet in a brick-and-mortar situation where a teacher may be responsible for managing 35 students at one time (Berge & Clark, 2005; Hassel & Terrell, 2004).

A great benefit to the field of education as a result of online learning is that students in various locations can have access to the best instructors. Additionally, former teachers may return to the field of education by way of teaching online as a result of part time opportunities because online courses decrease the stress or routine of traditional teaching, due to the benefit of flexible scheduling or because they can focus on individual

students without traditional classroom management concerns (Barker & Wendel, 2001; Hassel & Terrell, 2004).

On the other hand, there are also limitations to online learning. For example, students must have access to the appropriate hardware and software and the skills to operate the machinery; students must be aware of the expectations of a virtual class; and teachers must have appropriate pedagogy training and skills to teach virtual courses (Berge & Clark, 2005; Hassel & Terrell, 2004; Furey & Murphey, 2005; Russell, 2004).

A key to student success is that students must have access to the Internet and the appropriate technical skills necessary to run the software programs and the equipment such as the computer for an online course. The slowness of a dial-up connection, as opposed to broadband access, may distract, frustrate or discourage a student, just as an outdated computer may. Students in higher poverty areas are less likely to have access than students in more affluent areas (Russell, 2004)

A misconception that students tend to have is that online learning requires less time and work. Students must dedicate the same amount of time and effort to an online course as a traditional course. If a particular student in an online course has poor study habits in a brick-and-mortar setting, then he/she will most likely have such poor habits in the online setting, as well (Berge & Clark, 2005; Hassel & Terrell, 2004). In the online environment, daily routines, may not seem as obvious, so it is possible that students may become confused regarding the organization of assignments, deadlines and submission procedures. Online teachers need to be particularly aware of this tendency for confusion since the daily informal communication methods may be lost online (Russel, 2004). Additionally, students may feel isolated from the instructor, since there is no actual

traditional time and no in-person interaction (Hoffmaster, 2009; Russell, 2004). In other words, online learning is not suited for all learners, and teachers need professional development to assist in compensating for the different medium (Barker & Wendel, 2001). Some students, by nature, will always be more successful in traditional environments (Furey & Murphey, 2005).

Teachers must also learn new skills in order to be effective online. According to Furey and Murphey (2005), online teachers develop a high level of ownership for their online courses. Furey and Murphey (2005) explain that online teachers have an authority to design materials and the flexibility to address learner needs in a different medium without traditional constraints associated with a brick-and-mortar environment where time is limiting, and students may not be as willing to be open as a result of a lack of individual interaction (Furey & Murphey, 2005).

Teachers must be prepared for all the previously-mentioned variables. According to Russell (2004), teacher training for online delivery often seems to be in the form of virtual training modules. For example, a 15-week program is required of Canadian teachers in Fairfax County, and there are virtual training options for teachers with Virtual High School. Florida Virtual School has online training and a mentoring program for new teachers, as does the e^{4}TN program in Tennessee, and both states require highly-qualified and certified qualifications for online teachers (P.Lane, personal interview, March 29, 2009; Russell, 2004).

One common misconception about distance learning research – or, indeed, research with any technology in education – is that studies should focus on comparing the quality of distance and traditional instruction. Though early studies tended to compare the

two, Clark (1983) found that there is no benefit based on the medium for instruction. Rather quality instruction depends on instructional design and delivery, rather than delivery technologies. For example, in computer-based instruction (CBI), Clark (1985) found that when the same teacher designed the content in traditional and the CBI courses, there was no significant difference in effect sizes between the two. The effect the technology provided was a "washout." It was the instructor who had the most impact on learning (Clark, 1985, p. 250). Clark's study proved that the teacher is the variable that influences student achievement most, especially if he or she designs the curriculum and instructs the lessons. This applies in all settings regardless of the medium (Clark, 1985). Therefore, in terms of virtual schooling, quality is impacted by teacher quality. Thus, it seems even more necessary that an instrument is necessary to measure the effectiveness of teacher quality online.

Studies of virtual schooling quality. There are numerous studies of distance learning with adult learners. However, since the field of virtual schooling is still relatively new, there are fewer extensive studies (Clark, 2003; Murphey & Rodriguez-Manzanares, 2009; Russell, 2004).

Cavanaugh's (2001) meta-analysis indicated that, with equal attention and appropriate implementation, virtual schooling will have comparable success to traditional brick and mortar education. The Florida Tax Watch Center for Educational Performance and Accountability's (2007) evaluation of the Florida Virtual School found that students scored higher in online classes than in traditional courses. However, teachers of online courses are only compensated if students are successful, and FLVS has no requirement to

serve students with exceptional needs (Florida Tax Watch Center for Educational Performance and Accountability, 2007; Ryan & Watson, 2007).

In Smith's (2000) study, high school students from Alberta who were enrolled in virtual schools showed positive changes in social and emotional growth as a result of receiving more personal attention in virtual classes (Furey & Murphey, 2005). Mills (2002) studied the Virtual Greenbush program and found the teacher interaction and intervention it provided allowed for what Mills referred to as teachable moments, as defined by students. Without such teacher interaction, according to students, they could not have moved forward with content in an online course (Mills, 2002). Zucker and Kozma (2003) found that students were, in some cases, more satisfied with online courses rather than traditional courses as a result of online teacher interaction.

Smith (2000) also found that online teachers were concerned with workload as a result of online communication, course development, new pedagogy and technology, and preparation time. Teacher satisfaction was a result of an opportunity for innovation and creativity, flexibility, and camaraderie (Barker & Wendel, 2001; Hunter & Smith 2002; Murphey & Rodriguez-Manzanares, 2009; Smith, 2000).

Barker and Wendel (2001) completed a three-year study in 2001 that focused on virtual schools in Canada. They found ideal traits of online teachers that affected program success in "innovation and in technology; creativity and enthusiasm; a desire and ability to work collaboratively; a commitment to put students first; a willingness to work with parents; technology skills; and the ability to adapt quickly to change" (p. 122).

Overall, research findings on virtual schooling seem similar to those on distance education in general: potential for achievement and student satisfaction are equivalent,

given the right combination of design and facilitation. It also seems clear that virtual schooling offers unique opportunities for interaction and personal attention, provided the right design and facilitation take place.

Current controversies related to virtual schooling quality. Though it seems clear that virtual schooling can be of high quality, it is still currently a controversial topic in American education. Some of the controversies seem to be a by-product of the movement's rapid growth, but others reflect misunderstanding of the concept of virtual schooling (Tsai, et al., 2008). Still others are social/political in origin. Controversies related to virtual schooling include: drop rates, funding sources, and policy issues (Clark, 2001; Diaz 2002; Roblyer, 2006).

Drop rates in online classes exhibit a higher rate than in the traditional setting, and this must be considered when comparing the two formats (Carr, 2000; Diaz, 2002; Roblyer, 2006). This higher drop rate may not be indicative of instructional quality, however. There are several reasons to consider as to why students drop online classes at a higher rate than in the traditional setting. Demographics must be closely examined, for poverty and parental education level are indicators of student success (Jain, 2002). Some students state that they drop because the teacher is not experienced enough in the online program; therefore, the students prefer to drop the online course and take it in a traditional setting (Carr, 2000). Critics of online learning suggest that the environment is not personalized; however, students and instructors of gifted students disagree with this conclusion (Carr, 2000). Data collected from state virtual schools by the Southern Regional Education Board (SREB, 2007) suggest that some students are simply not well-prepared for courses in which they enroll. They lack content background and/or online

skills that would allow them to be successful in this new environment. Another reason SREB cites for higher student dropout is technical problems resulting from inadequate technical support.

Funding sources for virtual schooling are also a major topic of ongoing disagreement and discussion. Sixty percent of the 44 virtual schools surveyed by Clark (2001) cited start up funding for technology and staff as a major challenge for success. This includes funding qualified staff and appropriate infrastructure for development of and ongoing updates to online courses. In the past, government funding such as state grants and tuition have been the most common forms of funding (Clark, 2001). “Cyberschools are the 800-pound gorilla of the choice movement, although vouchers and charter schools get a lot more attention," said William Moloney, education commissioner in Colorado, where state financing for online schools has increased almost 20-fold in five years (Glass, 2009, p.9). It has grown to $20.2 million for 3,585 students today from $1.1 million for 166 full-time students in 2000” (Glass, 2009, p. 9). In the fall of 2006, the State of Colorado was paying for the schooling of 8,236 online students” (Glass, 2009). Public school districts often see themselves as in competition for funding with traditional schools (*Education Next*, 2009), which makes virtual schooling a provocative topic where many legislators are concerned.

Policies governing virtual schooling are controversial, particularly in relation to the use of private companies and to the size of the role virtual learning should play in a given education system. Private and state and local government agencies must work together in order to meet policies (Glass, 2009). Some legislators have successfully lobbied to institute virtual courses in traditional classrooms. In Florida, for example, a

law was passed in 2007 to make online learning available to all K-8 students by 2009/2010. Alabama requires high school students to take one online course before graduation (Glass, 2009). Policy also affects funding in areas such as full-time enrollment dates, impacts of assessment, level of teacher involvement, certification requirement of virtual teachers, and reciprocity of online teachers with various states. All of these issues impact both traditional and virtual school funding, which creates an ongoing debate on the role of virtual schooling (Glass, 2009).

One area of ongoing controversy is how to assure that instruction in virtual settings is of high quality (Carr, 2000). Since instruction depends on teachers and their performance, measurement of instructional quality focuses on teacher effectiveness. Background on this topic is presented here in two parts: methods of measuring effectiveness of traditional instruction, followed by methods of measuring teacher effectiveness in online settings.

Research on Measuring Effective Instruction: Traditional and Online

Overview of various strategies: Strengths and limitations. There are many ways in which teacher effectiveness is currently measured, including but not limited to: observation, peer-observations, self-assessment, portfolio, and standardized tests. Each of these has measurement limitations and issues. Qualitative measures such as observation, interviews, peer-observations, and portfolios leave opportunities for administrative bias, are costly and time-consuming, and may be unreliable (Braun, 2004; Hammond, 1996; Lengeling, 1996). Standardized tests that measure Adequate Yearly Progress (AYP), on the other hand, provide an opportunity for an objective, quantifiable method to measure the effect teachers have on student performance (Marthers, Olivia, & Laine, 2008).

Classroom observations are usually conducted by administrators and may be the most common form of teacher evaluation (Marthers, Olivia & Laine, 2008). Classroom observations do reveal teacher-student rapport, which may not be assessed via another model; however, this method of evaluation can be biased, unreliable, and invalid (Barret, 1986; Hammond, 1996; Marthers, Olivia, & Laine, 2008; Peterson, 1994). Hammond (1996) found that observations are often limited in time and provide little to-no-feedback for improvement. Observations are also seen as interruptions to class (Peterson, 1994). Goldstein (2008) found a contradiction to the perception of ineffectiveness in current teacher evaluation methods with a longitudinal peer review study in an urban district. He concluded that participants were able to engage and willing to improve through peer review as a method for evaluation. This was so even though it requires training of peers and learning how to observe others, and it may cause conflict between staff members or peers (Barret, 1986). Peer assessment is a time consuming and cost-prohibitive method for evaluation (Barret, 1986). Reflections from self-assessment may include pre- and post- observation conversations, conversations with peers, or even a portfolio development. Some school districts require teachers to record teaching so they can reflect and analyze their own instruction (Marthers, Olivia, & Laine, 2008). This method for evaluation is difficult to measure and time-consuming (Marthers, Olivia, & Laine, 2008). Portfolios are another method for teacher evaluation; however, they too are less common than teacher observation (Marthers, Olivia, & Laine, 2008). Portfolios may include self-analyses, recordings of self-taught lessons, lesson plans, and examples of student work and feedback. There are no conclusive findings to support enhancements in instruction based on teacher portfolios, and they are very time-consuming to construct and review.

However, they do allow administrators a greater opportunity to understand the non-instructional components of a teacher's classroom such as lesson planning, student relationships, and self-reflection and analysis (Marthers, Olivia, & Laine, 2008; Tucker, Stronge, Gareis, &Beers, 2003).

In general, as was stated earlier, qualitative assessments of teacher effectiveness measures leave opportunities for administrative bias and are costly and time-consuming, while standardized testing scores have margins for error. Standardized testing, however, provides a quantifiable and measurable result that can be tied directly to student gains (Alliance for Excellent Education, 2008; Braun 2005; Elmore 2002; Gore 2007; Marthers, Olivia, & Laine, 2008). Most educators agree that effective instruction can be measured by student improvement; however, determining how to measure a teacher's effectiveness, even when including student improvement, is an area upon which experts cannot agree (Alliance for Excellent Education, 2008; Darling-Hammond, 2007; Gordon, Kane, & Staiger, 2006).

It seems reasonable to measure teacher success based on the contributions to student learning. However, engagement and enthusiasm are difficult to measure, and standardized testing does not take into account factors such as this. Standardized tests are usually based on a variety of variables such as the interpretations of state standards by a group of experts as well as the actual instructional design of the test. If a state standard is to determine fact from opinion in language arts, then this is easily distinguishable on a multiple-choice standardized test; but if the standard is to write an essay, then this would be better measured by a written essay, not a multiple-choice standardized test item. Expense vs. budget may impact the way such a question is actually graded. For example,

training professionals to score an essay question versus using a computer-graded system for a multiple choice question have significant cost differences. Psychometric decisions, designing and the implementation of the test, the scale score, and the validity of the test itself are all influenced by a variety of factors (Braun, 2004). Also, with standardized tests, there are usually multiple versions that measure the same standards or skills, often referred to as a horizontal statistical procedure.

Using Value-Added Assessments to Measure Teacher Effectiveness

One standardized measure that has come into increasingly popular use is the Value-added Model (VAM). The VAM is a statistical method to determine the impact a teacher has on student's learning/achievement over a one-year period. This formula considers many factors, making it a complicated activity. It calculates a student's expected academic growth for a school year and compares it to how the student actually performs on standardized tests. Students who have higher-than-expected gains are considered to have teachers who added value (hence, the name). Students who score lower than expected are with teachers who did not add value. This test can only be performed with students and teachers who are taking courses that have end-of-course exams.

For a VAM, in order for data to be collected over multiple years, a scale score must be created, which is referred to as a vertical scaling statistical procedure. Second-grade math must be on the same scale as sixth-grade math, for example. Should there be the same expected gain for students regardless of the grade in which they are, even though content is more difficult in the eighth grade? Is a gain at the low end of the scale the same as a gain the upper end of the scale (Braun, 2004)? The tool for measurement

must provide consistency throughout the state in order to generate quantifiable data that is valuable to teachers, educators, and administrators.

Regardless of complaints or compliments regarding No Child Left Behind (NCLB), impact of legislative policy regarding the reauthorization of the Elementary and Secondary act is to make our nation's educators more aware of test scores. Annual yearly progress (AYP) is required by NCLB, and this is based on test scores. Cohorts in schools must meet AYP based on a target goal. Value-added assessments are a quantitative way to measure whether schools meet NCLB goals.

Under NCLB requirements schools must provide highly-qualified teachers, which means that they meet requirements for academic training and the state's licensure criteria. This has resulted in a focus on teacher quality. Some experts are connecting the idea that quality instruction equates to student learning. The connection of the two in a quantitative model has been studied over the last decade in what is referred to as valued-added models (Braun, 2004).

VAM is a way to measure schools, districts, and teachers on how much students progress over a year or their academic growth in a year, not on their level of achievement (Ballou, 2002; Braun, 2005). This is a new, nontraditional method for evaluating teachers and students, and it offers a way to evaluate them based on student progress rather than mean scores. Rather than ranking students on individual standardized scores, the statistical measures of VAM allow students, parents and the public to measure student and teacher success based on student gains from year to year. These gains may be measured at the school and district level so that administrators can focus on data at the most appropriate level to identify success strategies and problems.

AYP may also be measured using VAM, which helps schools to meet AYP since it is at the individual student level. The value-added formula controls for poverty, ethnicity, and other circumstances that would affect initial achievement (Ballou, 2002). Simple averages don't tell educators about a given school or teacher. For example, high averages in one school may suggest that students are from more highly educated or more affluent families. On the other hand, lower scores in another school may only imply that poverty affects test scores, regardless of how strong the teacher student effect is (Ballou, 2002). Schools and teachers are not given credit for students who enter at a higher level or behind grade level, according to Ballou (2002), when the VAM is used.

At the end of the value-added analysis, a number in the form of scale-scored points is associated with each teacher, and it may be in the form of a positive or a negative value. This number is reflective of the teacher's performance in comparison with that of other teachers based on student growth. Although randomized assignment is the best scenario for setting up an experiment, and it seems that VAM would allow for that, the reality is that, in schools, randomized class assignments are not always possible. In some situations, for example, parents may be able to request which teacher their child has. Given that the design of the value-added model is for random assignment, in situations where parents have influence as to the teacher that their child is assigned, a statistical distraction is in place that impacts the representative scores of the general population (Braun, 2005; Braun, 2004). Additionally, all teachers do not have the same resources to teach their students and all teachers are not using the same reform strategies in the classroom (Braun, 2004).

The Tennessee Value-added Assessment System. The State of Tennessee has found a correlation between teacher effectiveness and student achievement. "Researchers in Tennessee have found that students given the most effective teacher for three years in a row made over twice the gains of comparable students assigned the least effective teachers" (Alliance for Excellent Education, 2008, p. 2; Sanders & Rivers, 1996).

The best known and most frequently-used VAM model is the Tennessee Educational Value-added Assessment System (TVAAS). It is widely implemented in the State of Tennessee (Braun, 2004). The TVAAS model for evaluating teachers in the State of Tennessee is a direct result of the Tennessee Educational Improvement Act of 1992. Tennessee is geographically diverse with many rural areas. As a result of inequities in funding, many small, rural districts brought a lawsuit against the state. The business community then put pressure on the state to reform the system. As a result, the legislation put a strong accountability system into place.

The TVAAS model or Tennessee's VAM was founded and designed by David Sanders and his colleagues at the University of Tennessee at Knoxville in 1993. Originally the TVAAS formula was designed to determine the effectiveness of food in livestock with the Department of Agriculture at UTK (K. Kelly, personal communication, September 4, 2008). However, as a result of pilot studies that Sanders and his partners completed in the 1980's, Tennessee legislators embraced the TVAAS model as the accountability system for education. As a result, schools and school systems must demonstrate adequate progress in five subjects, all of which have standardized tests. The value-added scores began being publicized and assigned in 1993. The public report does

not include individual teacher scores, which are provided to respective teachers and administrators. TVAAS data collection began in 1996 (Kupermintz, 2003).

This particular VAM is based on specific components that begin with the student's cohort. The cohort is defined by the year the student enters second grade. All Tennessee students have seven tests that are labeled A-F upon completion of the eighth grade (Bock, Wolfe, & Fisher, 1996).

TVAAS is designed to measure the performance of schools, school systems and teachers through learning outcomes by means of data analysis. In theory, a teacher or school can be held accountable by the amount of learning a student gains over one-year. Expected and average gains then would be set by the State Department of Education (Bock, Wolfe, & Fisher, 1996). For example, if district X, based on the TVAAS prediction formula, expects a district mean prediction score, using the general linear formula that is copyrighted by SAS, of 528, and the actual mean TVAAS score is 522, then the district effect or teacher effect score is -7. In this example then, the teachers did not have a positive effect on student gain in academic learning (K. Kelly, personal communication, July 23, 2009).

Various value-added models differ in the number of years of data they employ, the kind of adjustments they make, and how they handle missing data (Braun, 2004). Tennessee uses a standardized testing model to assess students in second-through-eighth grades in math, language arts, reading, science, and social studies. These standardized tests, given during the last two weeks of the school year are referred to as TCAP tests at the middle school level. Reading, language arts and math contain both norm referenced tests and criterion referenced metric items that are on an IRT scale, while the science and

social studies tests have only norm referenced metric items that are on an IRT scale. Only the norm referenced items on the test are used for TCAP or the Tennessee assessment program. The tests are provided by CBT McGraw Hill/Terra Nova for the State of Tennessee (Bock, Wolfe, & Fisher, 1996). Tennessee also measures value added and teacher effect scores in high school core content areas using longitudinal data from middle school TCAP testing. Tennessee rescales the prediction scores for the previous three year's data each year as part of the model (K.Kelly, personal communication, July 23, 2009).

There are some considerations that have to be taken with gain scores. To calculate gain scores for Tennessee's TVAAS, for example, one must find the difference in two test scores that are approximately 12 months apart. Each individual test, required in Grades 3-8 may have had distractions during the administration, such as a disruption of the testing environment, during the test that may cause statistical error. This error is compounded when determining gains or value added because the statistician is finding the difference in two scores to determine the gain. Each individual score may have been influenced. Another complexity of the TVAAS formula is that statisticians have an adjustment to the raw scores to minimize any "noise" that is a result of testing distractions. The more participants in the testing, the more accurate the testing is (Ballou, 2002; Braun, 2005).

The teacher effect is determined by three factors once the student value added has been found: (a) district average for that specific school year and specific grade (b) class or teacher effect for the specific grade or specific year and (c) system or unsystematic

variations for that specific class and specific year (Braun, 2004; Bock, Wolfe, & Fisher, 1996).

At the second year of student assessments or the end of third grade, another component is added into the VAM, the teacher effect from the previous year. Classes of students are judged based on the class effect, not the test mean (Braun, 2004; Bock, Wolfe, & Fisher, 1996). The average class effect is attributed to the teacher. Therefore, it is the teacher effect. There are four components to this score: (a) district average for that specific school year and specific grade (b) class or teacher effect for the specific grade or specific year (c) system or unsystematic variations for that specific class and specific year, and (d) teacher effect for the previous year (Braun, 2004). According to Braun (2004), by subtracting the first year score from the second year score, statisticians yield the students' effect from year to year.

Each subject and grade is assigned its own equation, and "statisticians can add equations for the data from subsequent years. Sanders refers to this as a 'layered model' to capture the notion that the data from each succeeding year are added to those from the previous years" (Braun, 2004, p. 14).

The TVAAS model makes use of all data available for students for a 5-year-period, with the exception of student characteristics. Student characteristics are not taken into consideration with this model, according to Braun (2004). However, according to Dr. Kirk Kelly in a personal communication (July 23, 2009), in Tennessee, income status and race factors are required on student answer sheets for testing, so these variables are available to be included in the general linear formula if SAS chooses. Sanders, the founder of the model, is certain that there is no need for such considerations, though

(Ballou, 2002). While there are substantial correlations, the correlations represent zero gains in reference to student characteristics, according to Sanders; therefore, Sanders does not include additional calculations for student characteristics such as race and socioeconomic status. However, this is not universally valid or accepted (Braun, 2004). Estimation algorithms are in place for any missing data. This is an advantage of the approach as it applies to district-level data but not to teacher or student effect data, as systematic errors can be introduced.

As a result of TVAAS testing, Tennessee has substantial data collection on each student. The Tennessee State Department of Assessment has data from year to year, school to school, and system to system throughout every student's education. Tennessee, unlike other states, has not relied on home information such as income to predict accountability information, as it is not always reliable, and the relationship of the variables is not strong enough to predict gain (Ballou, 2002; Bock, Wolfe, & Fisher, 1996).

Teacher effect is determined by many factors including the progress in learning teachers' current and previous students show on standardized testing each year after the test scores are adjusted for gains from previous teachers and the growth, students show on previous standardized tests. The student gain scores are also adjusted for contributions from other teachers and subjects over a number of years (Bock, Wolfe, & Fisher, 1996).

A value-added analysis is applied to the gains the students make each year during grades 3-8, and that score is compared it to the gains made by the normative sample for that same content area and same grade level. The student must be present in the teacher's class 150 days of the school year in order to "count" towards the TVAAS for that

particular teacher (K. Kelley, personal communication, September 4, 2008). If the normative sample score gain is 12 points for fourth grade math, and a student gains 12 points in fourth grade math, then the teacher scores 100%. A teacher whose student gained only 10 would score 83% (10/12), for example. Random assignments of teachers and students to classes create much more credible statistics (Braun, 2005).

To compute the teacher-effect score, Tennessee's Department of Assessment first determines the TVAAS, score which is based on the Tennessee Comprehensive Assessment Program (TCAP) for elementary and middle school students and the Gateway tests for high school students. These tests are vertically-aligned, standardized testing series so that all skills are measured on grade level in grades 3-8. Student TVAAS scores are measured through a statistical mixed-model theory and methodology that requires a layered model. The first calculation is based on student TCAAP scores and gains or losses in those scores at the end of each school year in grades 3-8. The actual formula is copyrighted by SAS (2009), and they do not release the actual variables. However, the symbols in the general linear equation represent variables for each year a child participates in testing in Tennessee (K. Kelly, personal communication, July 23, 2009). For example, in the following formula, $Y_{96}{}^{5} = B_{96}{}^{5} + U_{94}{}^{3} + U_{95}{}^{4} + U_{96}{}^{5} + E_{96}{}^{5}$, the $Y_{96}{}^{5}$ part of the formula represents the student TVAAS score for a student who was in 5^{th} grade in 1996. B represents the district mean test score in 1996 on TCAP testing. U represents the teacher's contribution to the students' learning in previous years (1994 and 1995 in this example) and E is the "student level scholastic component," or TCAP test score, in the year represented. This example represents the student's scholastic component in 1996 (Ballou, Sanders, & Wright, 2004, p. 40).

In the state's model, once the TVAAS score is determined, then the formula is engaged to solve for U, which is the teacher effect, $U_{95}{}^{4} = (Y_{95}{}^{4} – Y_{94}{}^{3})- (B_{95}{}^{4} – B_{94}{}^{3}) - (E_{95}{}^{4} – E_{94}{}^{3})$. The teacher effect is then based upon "year-to-year gain after removing the district mean gain and the contribution of factors idiosyncratic to the student" (Ballou, Sanders & Wright, 2004, p. 41). When looking at a standard bell curve, the teachers that administrators are actually looking for fall to the extreme right (very high effect) or left (very low effect). Teachers who are labeled "Non-detectable difference" fall into the -1 to +1 area. The outliers in the extreme +2 and -2 are the teachers who have a detectable difference. Teachers who score a -2 for a teacher effect need attention for professional development. Teachers who score +2 can serve as lead or mentor teachers, for example (K. Kelly, personal communication, July 23, 2009).

When Y= Xb + Zu + e (Ballou, Sanders, & Wright, 2004) the district mean score "is a vector of fixed effects," the teacher effect is a "vector of random effects," and X and Z are incidence matrices (indicating which students have been assigned to which teachers in which subjects in which years, and e is a vector of random error terms" (Ballou, Sanders & Wright, 2004). In other words, set variables are responsible for the district mean score. The teacher effect has set variables but they range depending on how the student scores and the outcome of the district mean, for example. At least ten students within the same cohort must be included in the formula to ascertain a teacher-effect score (Ballou, Sanders, & Wright, 2004).

Methods of Measuring Effective Online Instruction

Effective teaching strategies for the traditional classroom may be different than those of the online classroom. Because online learning is based on the Internet there is a

learning curve for teachers. Therefore, training in the online environment and the online pedagogy is key to successful online instruction (McKenzie, Mims, Bennett, & Waugh, 2000; Pallof & Pratt, 2001).

Effective online instructors have some common characteristics with teachers in the traditional classroom, but there are also distinctive skills that are necessary. Savery (2005) captures some of the traits of an effective online instructor with the acronym "VOCAL." In order to be effective online, Savery (2005) suggests instructors be Visible, Organized, Compassionate, Analyze, and Lead by example. The *Southern Regional Educational Board* (2003) identifies traits of effective online instructors. For example, quality online teachers must be excellent time managers, strong communicators, well-organized, and content experts.

Time management is an essential skill for online instructors because students can work "anytime and anyplace" in an online course. The flexibility of "anytime and anyplace" learning offered to students is one of the most popular reasons for citing virtual schooling as a beneficial program to students (Berge & Clark, 2005; Furey & Murphey, 2005; Hassel & Terrell, 2004; Russell, 2004). Since SREB (2003) finds that high-quality online teachers communicate and interact with students in a timely manner, where students have 24/7 access quality online teachers excellent time-management skills is an essential characteristic of an effective online instructor.

Online teachers must be highly communicative and timely in responding to students (Treacy, 2007). This also requires good organization. Students are more successful in an environment that is more interactive, calling for assimilation and knowledge construction through social interaction, and because students may never

actually see their teacher excellent written and verbal communication skills are necessary (SREB, 2003).

Summary of Findings from Literature

As a result of the benefits virtual learning offers to students, teachers, and school districts (e.g., flexibility, increased access, engaging teaching and greater teaching opportunities), virtual schools have seen consistent growth since 1996 (Berge & Clark, 2005; Furey & Murphey, 2005; Hassel & Terrell, 2004; Russell, 2004; Watson & Ryan, 2007, Zandberg & Lewis, 2008). Another reason for such growth is because virtual schooling provides students with access to the best teachers regardless of the teachers' or students' geographic location (Barker & Wendel, 2001; Hassel & Terrell, 2004). Evidence of this can be seen in the growth of Florida Virtual School, for example, with more than 120,000 enrollments in 2008 (Florida Tax Watch Center for Educational Performance and Accountability, 2007; Gemin, Watson & Ryan, 2008; Zucker & Kozman, 2003). This growth requires attention to delivery models and student success.

Student success rates are particularly visible in online courses given the controversial concern over students in virtual schools having a higher attrition rate than in traditional courses (Carr, 2000; Diaz, 2002; Roblyer, 2006). According to SREB (2003) and Carr (2000) student success in virtual schools depends on the quality of the instructor and the methods used to deliver instruction. SREB (2003) states that it is essential for states and schools to "select, hire, train, and evaluate teachers to ensure that they can teach effectively online" (p. 2). Because online courses may be delivered in a variety of methods such as blended learning, synchronous or asynchronous models, they require a specific skill set and pedagogy training in order for teachers and students to be successful

in the online medium (Berge & Clark, 2005; Hassel & Terrell, 2004; Furey & Murphey, 2005; National Center for Educational Statistics, 2005; Russel, 2004).

Thus, particular attention needs to be given to identifying effective online instructors. SREB (2003) has provided a checklist for states and schools to use when evaluating online teachers with three categories for evaluation to indicate if the online instructor does exceed, doesn't exceed or exceeds expectations in each category. However, the measurement is subjective, leaving room for administrative bias.

Florida Virtual School, one of the pioneers in virtual learning has defined teaching standards for online instructors; however, Liz Azukas, an Instructional Leader and Sue Steiner and Program Director for FLVS (2009) state that "The old paradigm of teachers being supervised by administrators who share a physical space no longer applies to cyber-education, thus causing practitioners to rethink how they enact evaluation" (p. 1). Azukas and Steiner (2009) describe their method for evaluating teachers at FLVS to be one in which supervisory instructors virtually "observe" teacher practices in the online classroom by monitoring pillars of FLVS: "Communication, Collaboration, Flexibility, Learner-Centered and Organization" (p. 1). This model allows for subjective concerns associated with qualitative data such as bias (Barret, 1986; Hammond, 1996; Marthers, Olivia & Laine, 2008; Peterson, 1994). In 2006, SREB provided an updated evaluation tool based on supervisor or administrator observation for states and schools to evaluate online instructors and online programs. Again, qualitative concerns with such data exist.

Implementing a value-added model to measure online teacher effectiveness, such as what Tennessee currently applies in the traditional classroom (TVAAS), would measure student academic growth in a year or a course in a quantitative manner in order

to generate a teacher-effect score or a gain score (Ballou, 2002; Bock, Wolfe, & Fisher; Braun, 2005). This would provide an opportunity for the public, students, teachers and administrators to evaluate a quantitative measure of the effectiveness of a teacher.

Evidence strongly supports virtual schooling as a growing movement, and it seems clear that there are effective teachers online; however, there is no objective way currently published to measure the effectiveness of online teachers. As a result of the growing movement of virtual schooling, a method needs to be in place to confirm the quality of online teachers as readily as do current methods for traditional teachers. Though a different skill set is involved in online teaching, it may be that the most important skills are those that teachers bring from the traditional environment to the virtual medium that allow them to be effective, quality online instructors. Administrators and virtual school program coordinators need to know how to measure online teacher quality based on quality indicators and teacher effectiveness data as a method to improve student success and reduce retention in virtual courses. This research will provide evidence of whether or not we can use the same measures to gauge online teacher effectiveness as we do in traditional environments, thus addressing the need for effective methods of judging quality in online teaching. In light of the still-controversial topic of virtual school quality and popular concerns about its higher-than-usual dropout rates, this study will also help determine educator perceptions of the most important quality indicators of online teachers in comparison to traditional teachers. The latter will identify practical and conceptual perspectives that influence Tennessee's ability to use various measures of teacher quality.

Chapter 3: Methodology

Study Design

The study used a mixed-methods design to examine the relationship between teacher quality indicators in the traditional and online environments. It would be valuable for education system stakeholders to explore the feasibility of using a parallel measurement to determine if teachers are equally effective in the online classroom environment and the traditional classroom environment. In order to do this, there must be a standard by which to measure the individual teacher in each environment so that a correlation between the two skill sets may be made. End of course (EOC) scores from a proportional sample of traditionally-taught students in a Tennessee school district were compared with those from a convenience sample of Tennessee online students. The analysis also determined if program effects between online vs. traditional formats differed by subject area in Algebra I, Biology, and English I courses.

Another primary purpose of the study was to examine how teachers and administrators felt online teachers should be assessed and if there are conceptual and practical issues related to the way traditional teacher assessment is currently handled versus the way educators feel it should be handled. An anonymous survey that contained both Likert-scale and open-ended questions was used to gather evidence about the nature of these perceptions.

Setting and Population

The setting for the study included Tennessee students and teachers. The traditional student population consisted of de-identified students who attended Hamilton County Department of Education schools in Tennessee. Hamilton County Schools were selected

as a convenience sample; however, the district includes urban, suburban, magnet, and rural schools. The online student sample included students who attended public schools in Tennessee. The online population of student data was from a pool of statewide students, not one specific school district. For the Likert survey, teachers from Tennessee who are affiliated with the Tennessee online program, e^{4}TN, were asked to participate. Participants from all groups are associated with Tennessee public schools.

There were three target groups for the study; two were used to compare program effects, and one was used to gather survey data. The first two groups were from the Tennessee school district (Hamilton County) and the state-wide online program. Data for students in both groups were de-identified. Their 6th 7th, and 8th grade Normal Curve Equivalent (NCE) scores were collected, along with their EOC test scores, and demographic data that included gender, race, exceptional education status, gifted status, GPA, ELL status, and socioeconomic status. Each sample (online and traditional) group consisted of 102 Algebra I students, 26 Biology students and 34 English I students for a total of 204 Algebra I students, 52 Biology students and 68 English I students in the sample.

The online student samples were selected based on all participants from spring and fall semesters 2007-2008 and 2008-2009 who had all NCE scores, completed the respective online course, took the respective EOC test, and had all demographic qualifiers. A sample that was proportional and systematic was drawn from the traditional district population to match the sample for the online population. Samples were matched as to numbers of students in each socioeconomic group (as indicated by participation in free-and-reduced lunch program), ethnicity, gender, gifted status, ELL status, and

exceptional education status. Once this proportional sample was drawn, a systematic sample was selected. For example, for the Algebra I sample in the district-level program, there were a potential of 827 students in the district's traditionally-taught courses who qualified for selection (i.e., they were Algebra I students and had all the required data). Every 7th student was selected for a purposeful sample of 102 students to match the number of students in the online sample. The same procedure was followed to select the samples of students in traditional courses for Biology and English I courses.

Approximately 100 Tennessee teachers and administrators who are particiants in the e4TN program were invited to participate in an anonymous Likert survey as part of the study's focus on educator perceptions of teacher effect scores and TVAAS. All Likert survey questions were on non-controversial topics. Teachers and administrators were invited to participate in the online anonymous survey via email if they participated in the e4TN program during the spring 2009 semester.

Materials and Instruments

In the SAS teacher-effect model, which the Tennessee Department of Education uses, teacher-effect scores are dependent on TVAAS scores, which allow administrators to measure student achievement through gain. Longitudinal analysis of student TCAP data in grades 3-8 is needed to calculate the student gains in high school content areas where Gateway and EOC exams are offered. NCE gain is generated based on 3 years of NCE scores (Glossary of Terms and Abbreviations, 2009). The data needed to calculate student scores is based on: (a) the Tennessee Comprehensive Assessment Program (TCAP), a group of tests in five subject areas (math, science, social studies, reading, and language arts) administered annually to all Tennessee elementary and middle school

students in grades 3-8; and (b) end-of-course tests in Tennessee high school subjects through spring 2009 in five core subjects (English I, Physical Science, U.S. History, Algebra I, or Biology.

The State of Tennessee's online learning program began in January 2006. Since there are limitations with the number of online teachers that tie to students who have been in public schools since third grade, and since many online students may be in the exceptional education category because online learning meets the needs of a diverse population of students, the sample size that meets state criteria is currently too small for SAS to run a true teacher-effect score (J. Rivers, personal communication, November 17, 2009).

SAS is an outside vendor who created the formula used to calculate teacher effect and also has the state contract to complete the calculations each year for teacher effect. For this study, the researcher did not have access to the SAS model to calculate teacher effects, since the state's model is the copyrighted property of SAS. Instead, the approach used in this study compares program effects by considering many of the same variables as the SAS model. To address research questions 1 and 2, de-identifed data on EOC scores and other variables were obtained from a Tennessee school district and a Tennessee online program.

Data to address research question 3 and 4 were collected from a survey that included Likert-scale items on a scale where 1 was equivalent to "Strongly Agree," 2 was equivalent to "Agree," 3 was equivalent to "Neutral," 4 was equivalent to "Disagree," and 5 was equivalent to "Strongly Disagree." Open-ended questions were also included (see Appendix A). Data were collected on each teacher and administrator perceptions

regarding assessment of teacher quality in both traditional and online environments. The survey requested descriptive data on respondents, e.g., teacher experience, teacher area of highly qualified status, and semesters teaching online. Other sections of the survey asked participants to rank the types of appropriate teacher assessments (e.g., portfolio, observation, peer-review, self-evaluation) for the medium in which they are applied, either traditional or online. Participants were asked to share attitudes regarding TVAAS and teacher-effect scores as tools for measurement in both the traditional and online environments. To allow comparisons of teacher perceptions about appropriate measures of quality in both types of course environments, all ranked questions and descriptive data questions were the same in the sections on online and traditional environments. The survey was delivered via an electronic survey tool, *Survey Methods*, and responses were anonymous.

Data Collection and Analysis Methods

Program effect data collection. To provide data to answer research questions 1 and 2, the researcher received two sets of de-identified student information. One was from a traditional school district in which students had been taught in traditional classrooms, and the other was from the state's online program. Both were in the format of an *Excel* tab-delimited file that included a course unique identifier, 6th grade NCE scores, 7th grade NCE scores, 8th grade NCE scores, EOC exam scores, economic status, grade level, ethnicity, gender, GPA, gifted status, exceptional education status, and ELL status.

In order for data to be used, students in both the traditional and online groups had to have NCE scores for all grades 6 through 8 and have taken the EOC exam. All students in both the traditional and online environments were present in an Algebra I, Biology, or

English I course for a minimum of 150 days, as Tennessee's traditional teacher effect formula requires (K. Kelley, personal communication, September 4, 2008). This eliminated any students who had been dropped or withdrawn from the online environment based on e^{4}TN's 10-day drop and 30-day withdrawal policies. All online students earned "complete" status in the online course. A key difference between the SAS model and the approach the researcher used is that exceptional education students are not included in the SAS model. In that model, they are dropped from the calculation for teacher effect. However, they were included in the samples for this study, and exceptional education status was included as a variable in the regression analysis.

Survey data collection. The Likert survey was communicated to potential respondents as an emailed Internet link (Patten, 2005). Potential, anonymous participants were comprised of approximately 100 teachers and administrators who worked with the e^{4}TN program in Spring 2009 semester. The survey itself was generated and housed in an online survey service provided by Survey Software. The researcher set the survey tool to allow participants to take the survey once. Participants could choose to save their submissions in mid-survey and later return to that point where he or she left off later without losing any data (Survey Methods, 2008).

Survey participants had 30 days to respond to survey questions. On day 0, an email was sent that included survey information and the explanation of the study. On day 15, a thank-you email was sent to thank participants for their participation and to remind remaining potential participants of the opportunity to complete the survey. On day 31, a thank-you email was sent to all potential participants for their willingness to participate and for completing the survey.

Data analysis methods. In order to address research question 1, “Is there a significant difference in program effects of traditional classrooms (as measured by end-of-course scores for a sample of traditionally-taught students in a Tennessee school district) and online classrooms (as measured by end-of-course scores for a sample of Tennessee's online students)?,” the researcher first split the data file according to subject area (Algebra, Biology, and English 1). This by-subject split was necessary since the EOC tests for the content areas had different top scores and, therefore, could not be compared across areas. Then a *t* test was used to compare the EOC scores between the two programs by subject area. A *t* test is a procedure that is often used to test the null hypothesis by observing the difference in two means in outcomes from two groups (e.g., traditional and online). A null hypothesis, according to McMillan & Schumacher (2006) is a “formal statistical statement of no relationship between two or more variables” (p. 475). If the *t* test finds that means are different, at, or below the pre-test probability value of .05, then they are said to be significantly different (Patten, 2005).

A regression analysis was then done to determine any variables that could affect students' performance on EOC exams. Once variables were determined, an Analysis of Covariance (ANCOVA) was able to be done to control for those factors in order to get a true comparison of program effects.

For research question 2, “Do program effects between traditional and online environments vary significantly by subject area (i.e., Algebra, Biology, and English 1)?,” the researcher also used a data file that was split according to subject area (Algebra, Biology, and English 1) and examined results from the ANCOVAs.

Research questions 3 and 4 were addressed by analyzing a Likert-scale and qualitative data from a survey. Research question 3 asks, “Do Tennessee educators perceive that Tennessee's model for teacher-effect scores can be used equally well in both traditional and online environments?” Research question 4 asks, “What factors do administrators and teachers perceive should be considered in determining teaching quality in the online environment?” By asking this question, the researcher used Likert-scale and ranked items (quantitative data) and open-ended questions (qualitative data). Since the literature does not provide definitive evidence to indicate factors that administrators and teachers perceive should be considered in determining teaching quality in the online environment, the researcher included an opportunity for respondents to add their own factors in addition to the Likert-scale measures. Likert-scale data was analyzed with Mann-Whitney *U* and Wilcoxon tests. For the qualitative analysis of open-ended items, the researcher employed a content analysis to examine the survey feedback for themes and patterns in perceptions by using a constant-comparison method (Patten, 2005).

Likert-scale data was reported by mean scores, standard deviation, SE (sampling error), the *F*-score and *p* value (<.05) (Patten, 2005). A Mann-Whitney *U* test was done to compare the same participants' Likert-scale responses between online and traditional environments to measure attitudes toward using various measures of evaluation in traditional versus online environments. Wilcoxon analyses compared teacher and administrator responses on two of the items.

Summary of Methodology

Using a mixed-methods design, this study addressed the question of whether teacher quality indicators used in a traditional teacher evaluation can be used to measure teacher quality in online courses. For the quantitative data to address research questions 1 and 2, comparisons were made between traditional and online EOC scores on 162 de-identified Tennessee high school traditional students and 162 Tennessee high school online participants who have NCE scores for 6^{th}, 7^{th}, and 8^{th} grades and who took the respective Gateway or EOC exam in Tennessee in 2007-2008 or 2008-2009. Additionally, through an anonymous, electronic Likert-scale survey, the study examined how teachers and administrators who worked with e^4TN during spring 2009 felt both traditional and online teachers should be assessed. The Likert study also examined perceptions of conceptual and practical issues associated with assessment of both traditional and online teachers. Descriptive data of survey participants (e.g., experience teaching traditional, experience teaching online, and highly qualified endorsement areas) was collected in order to determine patterns in endorsement area and teacher effect, and attitudes towards assessments of both online and traditional teachers.

The researcher used a data file that was split according to the subject areas of Algebra I, Biology, and English I and used *t* tests and ANCOVAs to compare EOC scores of online and traditional students. Since a regression analysis identified variables that contributed to students' EOC scores, ANCOVAs were used to control for these variables in order to get a true comparison of program effects.

The anonymous Likert survey was completed electronically in Survey Methods (an online survey resource) with a 30-day window for completion. The survey was analyzed

with Mann-Whitney *U*, and Wilcoxon analyses. Finally, for open-ended comments, a content analysis for themes and patterns using a constant-comparison method was employed.

Chapter 4: Results

Overview of the Chapter

The purpose of this chapter is to present the results of the study in terms of findings related to the research questions. This chapter is organized by the four research questions posed in Chapter 1. To address research questions 1 and 2, the chapter first reports on findings of quantitative data collected to compare program effects in the traditional and online environments. Finally, this chapter addresses research questions 3 and 4 by reporting findings of survey results from educators in Tennessee that indicate factors and strategies that they perceive should be considered in determining teaching quality in the traditional and online environments.

Overview of Study Purpose and Design

This study collected quantitative data to examine the effects of online and traditional programs in terms of students' End of Course (EOC) scores, and collected qualitative data to determine educators' perceptions of factors that should be considered when gauging teacher quality in the online and traditional environments. The purpose of this study is to discover whether or not educational systems can use traditional indicators of teacher quality such as value-added scores to select effective online instructors.

Using a mixed-methods design, this study addressed the feasibility of whether teacher-effect scores originally designed for use in traditional teaching environments can and should be used to measure quality of online teachers. For the quantitative data to address research questions 1 and 2, comparisons were done between traditional and online EOC scores in order to gauge program effects in subject areas that require EOC or Gateway testing in Tennessee. Because a regression analysis identified several variables

having to do with students' prior ability to contribute to EOC scores, these were controlled statistically with an ANCOVA so that program effects could be measured and compared.

Additionally, through an anonymous, electronic survey that included Likert-scale, ranking, and open-ended items, the study examined how teachers and administrators who worked with e^{4}TN during spring 2009 felt that both traditional and online teachers should be assessed. The Likert-scale items examined perceptions of conceptual and practical issues associated with assessment of both traditional and online teachers. Descriptive data of survey participants (e.g., experience teaching traditional, experience teaching online, and highly qualified endorsement areas) were collected in order to determine patterns in attitudes toward evaluations of both online and traditional teachers. Data from Likert-scale items were examined using Mann-Whitney *U* and Wilcoxon analyses. Data from open-ended items were examined using a constant-comparison procedure to look for common themes and patterns in responses.

Research question 1 focus and methods. Research question 1 asks "Is there a significant difference in program effects of traditional classrooms (as measured by end-of-course scores for a sample of traditionally-taught students in a Tennessee school district) and online classrooms (as measured by end-of-course scores for a sample of Tennessee's online students)?" This represented the primary purpose of this study: whether or not online teachers can be evaluated for quality instruction using the same value-added model as the one used for teachers in traditional classrooms, or if they need an alternative model. If program effects differ, then variables outside those usually used

to measure teacher effect are having an effect on students. Therefore, teacher effects will be different in the two environments.

The data to analyze this question were two sets of de-identified student information. One was from a school district in which students had been taught in traditional classrooms, and the other was from the state's online program. Both were obtained in the format of an *Excel* tab-delimited file that included a course unique identifier, 6^{th} grade NCE scores, 7^{th} grade NCE scores, 8^{th} grade NCE scores, EOC exam scores, economic status, grade level, ethnicity, gender, GPA, gifted status, exceptional education status, and ELL status. Files were uploaded to an SPSS program for analysis. Since EOC top scores possible varied by subject area, the data were first split by subject area.

The researcher acquired data from a Tennessee school district for students who had also participated in Algebra I, Biology, and English I. The sample of traditional students was matched with the online student population in terms of proportions of students in various ethnicity, gender, socioeconomic status, gifted status, exceptional education status, and ELL groups. Both groups of students had variables such as grade levels, GPA, NCE scores, and EOC scores reported for their respective content areas. Since subject area EOC tests differed by top possible scores, the data were split by content area before analyses were done. Then EOC tests were compared for traditional and online groups in order to address research questions 1 and 2. Analysis of data for research question 1 was done by using a *t* test to compare the EOC scores in the two programs (Patten, 2005). This allowed the researcher to compare program effects in the traditional and online environments. A regression analysis was used to determine variables contributing

significantly to students' performance on EOC tests, and ANCOVAs were calculated by subject area to control for these contributions in order to compare program effects.

Research question 2 focus and methods. Research question 2 examines if "program effects between traditional and online environments vary significantly by subject area (i.e., Algebra I, Biology, and English 1)?" For research question 2, student data were used from three subject areas: Algebra I, Biology, and English I, in two environments, online and traditional. Again, ANCOVAs allowed a comparison of effects in these areas.

Research question 3 focus and methods. Likert-scale survey data for research question 3 were obtained from a survey of 304 educators that asked, "Do Tennessee educators perceive that Tennessee's model for teacher-effect scores can be used equally well in both traditional and online environments?" The data from the respondents were analyzed for mean score and standard deviations, and Mann-Whitney *U* tests were performed to compare Likert-scale responses between online and traditional environments and to determine if there were any differences in participant attitudes toward using various measures in traditional versus online environments.

Research question 4 focus and methods. Likert-scale survey data were used to address research question 4, "What factors and strategies do teachers perceive should be considered in determining teaching quality in the traditional and online environments?" As with research question 3, data were analyzed for mean score and standard deviation, and Mann-Whitney *U* tests were performed to compare Likert-scale responses between online and traditional environments and to determine if there were any differences in participant attitudes toward using various measures in traditional versus online

environments. Wilcoxon tests also were performed to determine if there was a significant difference in administrator and teacher responses (Patten, 2005).

Results for Research Question 1

Overview of Research Question 1 focus and methods. Research question 1 (Is there a significant difference in program effects of traditional classrooms and online classrooms?) represents the primary purpose of this study: to explore the relationship between online and traditional instruction in terms of student outcomes. The data to analyze this question were generated by comparing two samples, one in the online environment and one in the traditional environment for students who completed Algebra I, Biology, and English I, and had NCE scores for grades 6, 7 and 8.

Report and analysis of descriptive data collected on research question 1. The traditional sample was selected in a way that matched the online sample for important factors such as socioeconomic status and ethnicity. The online student population was selected based on those who were available with complete data sets. Then a sample of students who also met the data requirements (grades 6, 7, and 8 NCE scores and EOC scores) in the traditional environment were selected to include equivalent proportions of participants based on socioeconomic status, ethnicity, gender, gifted status, exceptional education status, and ELL status as shown in Tables 4.1 and 4.2.

Student participants in each group ranged by grade level from 8 to 12, with the majority of students being in 9^{th} grade and the fewest students being in the 8^{th} grade (see Table 4.1).

Table 4.1

Frequencies and Percentages of Student Participants in Demographic Categories, by Program

	Free Lunch Y/N		Exceptional Ed. Y/N		Gifted Y/N	
	Y	N	Y	N	Y	N
Online	79 (48.8%)	83 (51.2%)	28 (17.3%)	127 (78.4%)	1 (.003%)	161 (49.7%)
Traditional	82 (50.6%)	80 (49.4%)	35 (21.6%)	134 (82.7%)	3 (.009%)	159 (49.1%)
Total	161 (49.7%)	163 (50.3%)	63 (19.4%)	261 (80.6%)	4 (1.2%)	320 (98.8%)

Table 4.2

Frequencies and Percentages of Student Participants in Ethnicity Categories, by Program

	Asian	African American	Hispanic	Native American	White	Total
Online	2 (1.2%)	60 (37.0%)	4 (2.5%)	1 (0.6%)	95 (58.6%)	162 (100%)
Traditional	1 (0.6%)	57 (35.2%)	10 (6.2%)	0 (0.0%)	94 (58.0%)	162 (100%)
Total	3 (0.9%)	117 (36.1%)	14 (4.3%)	1 (0.3%)	189 (58.3%)	324 (100%)

Table 4.3

Frequencies and Percentages of Student Participants at Each Grade Level, by Program

	8^{th}	9^{th}	10^{th}	11th	12th	Total
Online	8 (4.9%)	69 (42.6%)	31 (19.1%)	25 (15.5%)	29 (17.9%)	162 (100%)
Traditional	1 (0.6%)	111 (68.5%)	41 (25.3%)	7 (4.3%)	2 (01.2%)	162 (100%)
Total	9 (2.8%)	180 (55.6%)	72 (22.2%)	32 (9.9%)	31 (09.6%)	324 (100%)

Results of inferential data analysis for research question 1. A *t* test was done between EOC scores of students in online and traditional programs, by subject area. There was no significance difference in two of the three courses. English I was the only

course in which when found: $t(66)=2.049$, $p=.045$. Algebra I and Biology reflected no significant differences ($t(202) = -1.551$, $p=.122$ and $t(50)=-.354$, $p=.725$, respectively).

However, since there were observed differences between the two groups regarding GPA, a *t* test ($p=.05$) was done to determine if there was a significant difference between online and traditional groups. As reported in Table 4.4, there was no overall difference between GPAs of participants ($p=.45$). However, Algebra I students in the online program did have a significantly higher GPA ($p=.002$) than students in the traditional program, and the English I students from the traditional population had a significantly higher GPA than the online students ($p=.00$).

Table 4.4

Results of t-test Comparison of GPA for Student Participants, by Program and Subject Area

Content Area	Traditional	Online	Difference	*p* values
Algebra I	1.71	2.05	0.34	.00
Biology	2.46	2.74	0.28	.26
English I	2.48	1.60	-0.88	.00
Total	1.99	2.07	0.08	.45

Since variables other than programs used could be accounting for EOC performance, the researcher used a stepwise regression analysis to identify variables in addition to GPA that could contribute significantly to EOC scores (Green & Salkind, 2005). It was determined that five variables were significant contributors: 6^{th}, 7^{th}, and 8^{th} grade NCE scores, GPA, and grade level. These are shown in Table 4.5.

Table 4.5

Stepwise Regression Results for Factors Contributing Significantly to End-of-Course Test Scores

Step	Variable	*B*	*Std Error B*	*B*	*t*	*p*
0	(Constant)	370.95	20.36		18.22	.00
1	7th	0.73	00.20	.32	3.62	.00
2	GPA	6.63	02.30	.13	2.88	.00
3	8th	0.53	00.19	.22	2.82	.01
4	Grade	6.120	01.90	.14	3.27	.00
5	6th	0.37	00.18	.16	2.04	.04

After these variables were identified, an average NCE was calculated and separate ANCOVAs were done using GPA, grade level, and average NCE as covariates to compare programs while controlling for these prior-ability characteristics.

The independent variable in all three ANCOVAs was program (online and traditional), and the dependent variable in all cases was EOC score. A preliminary analysis evaluating the homogeneity-of-slopes assumption indicated that the relationship between the covariate and the dependent variable did not differ significantly as a function of the independent variable. This was done to assure that the populations were similar so that ANCOVAs could be done to control for average NCE, GPA, and grade levels (Green & Salkind, 2005). The results of the test confirmed that populations were similar, so ANCOVAs were done for each content area and each covariate.

Results of the ANCOVAs, shown in Table 4.6, showed that in Algebra I and Biology, the average NCE and GPA were major contributors to the variance between the EOC scores in programs, with the covariates accounting for between 5% and 69% of the

variance between programs. The Grade Level covariate was a significant contributor only in English, though Algebra was close to being significant.

Table 4.6

Results of ANCOVA Contributions for NCE, GPA, and Grade Level as Covariates, by Subject Area

Covariates	Algebra I	Biology	English I
Average NCE	$F(1,200) = 72.62$ $p = .00$* (27% of variance)	$F(1,48) = 106.40$ $p = .00$* (69% of variance)	$F(1,64) = 60.93$ $p = .00$* (49% of variance)
GPA	$F(1,200) = 11.50$ $p = .00$* (5% of variance)	$F(1,48) = 42.60$ $p = .00$* (47% of variance)	$F(1,64) = 4.94$ $p = .03$* (7% of variance)
Grade Level	$F(1,200) = 3.66$ $p = .057$ (2% of variance)	$F(1,48) = 0.21$ $p = .65$ (0% of variance)	$F(1,64) = 6.20$ $p = .02$* (1% of variance)

As Table 4.7 shows, when results of EOC score comparisons were controlled for significant contributors (average NCE, GPA, and grade level), there were no significant differences between programs in Biology and English I. Although Algebra I did reflect significant differences between programs when the covariates GPA and grade level were considered, it did not when average NCE was considered as a covariate, so results were inconsistent.

Results for Research Question 2

The population was the same for research question 2 as it was for research question 1, and the analyses were also the same. As Table 4.7 indicates, there were no significant differences in programs in Biology and English I, but differences were found

in Algebra I for two of three analyses. Results generally indicate no differences by content area.

Table 4.7

Results of ANCOVA Comparisons of Program Effects, by Subject Area, Controlled for Covariates

Subject Area	Covariate	Program Status	N	Est. Mean	Std. Error	*p*
Algebra I	Average NCE	Online	102	506.90	3.62	.12
		Traditional	102	490.87	3.63	
	GPA	Online	102	503.30	4.10	.00*
		Traditional	102	498.40	4.10	
	Grade level	Online	102	502.96	4.33	.03*
		Traditional	102	488.37	4.83	
Biology	Average NCE	Online	26	551.14	5.33	.53
		Traditional	26	556.14	5.35	
	GPA	Online	26	548.85	7.07	.07
		Traditional	26	554.71	7.03	
	Grade level	Online	26	555.09	9.47	.44
		Traditional	26	550.76	9.48	
English I	Average NCE	Online	34	526.35	4.43	.51
		Traditional	34	514.53	4.41	
	GPA	Online	34	513.24	5.67	.72
		Traditional	34	520.26	5.37	
	Grade level	Online	34	512.09	5.09	.08
		Traditional	34	518.51	5.88	

Results for Research Question 3

Report and analysis of descriptive data collected on research question 3. Sixty-eight people completed the survey for a return rate of approximately 22%. As Table 4.8

indicates, over three-quarters of respondents were teachers. Five of the responses were neither, which means that they were serving in a facilitator role for the online learning program in Tennessee. Those who answered both (five) are administrators in the traditional setting, but they teach online, as well. Table 4.9 indicates that the online experience of teachers who participated in the study is evenly distributed among teachers with various levels of traditional experience. The majority of administrators who participated in the study had less than one year of experience, as shown in Table 4.10. Table 4.11 indicates that the majority of online teachers have 0-3 semesters experience. This group of online teachers is composed of administrators in the traditional environment, as well as teachers in the traditional environment, based on the reporting in Table 4.8. More teacher participants have endorsements in English and Social Studies; however, other endorsement areas seem evenly distributed, as shown in Table 4.12.

Table 4.8

Survey Respondents by Role (Frequencies and Percentages)

		Respondents	
		Frequency	Percentage
Role	Teacher	51	71.8%
	Administrator	10	14.1%
	Neither	5	7.0%
	Both	5	7.0%
	Total	71	99.9%

Note. Five participants fell into the category of neither teacher nor administrator; therefore, the total for this survey question is 71, rather than 66.

Table 4.9

Survey Respondents by Years of Experience Teaching in the Traditional Environment in Tennessee

Years Teaching	Frequency	Percentage
1-5	10	19.60%
6-10	11	21.60%
11 and 20	16	31.40%
21-30	14	27.50%
Total	51	100.0%

Table 4.10

Survey Respondents by Years of Experience as Administrators in Tennessee

Years Teaching	Frequency	Percentage
<1	11	45.8%
1-5	4	16.7%
6-10	3	12.5%
11 and 20	4	16.7%
21-30	2	8.3%
Total	23	100.0%

Table 4.11

Survey Respondents by Semesters Experience Teaching in the Online Environment in Tennessee

Semesters Teaching	Frequency	Percentage
0-3	31	63.30%
4-7	12	24.50%
8-10	06	12.20%
Total	49	100.0%

Note. Sixteen of the total respondents were not teachers in Tennessee or had not taught in Tennessee.

Table 4.12

Teacher Survey Respondents by Teaching Content Area

Content Area	Traditional	Online	Total Combined
English	15 (28.8%)	13 (29.5%)	28 (29.17%)
Fine Arts	1 (1.9%)	–	1 (1.04%)
Foreign Language	3 (5.8%)	5 (11.5%)	8 (8.33%)
Health PE	1 (1.9%)	1 (2.3%)	2 (2.08%)
Math	10 (19.2%)	5 (11.4%)	15 (15.63%)
Science	7 (13.5%)	5 (11.4%)	12 (12.5%)
Social Studies	3 (5.8%)	8 (18.2%)	11 (11.46%)
Career and Technology	3 (5.8%)	1 (2.3%)	4 (4.17%)
Other	6 (11.5%)	4 (9.1%)	10 (10.42%)
Not Teaching in Traditional Environment	3 (5.8%)	2 (4.5%)	5 (5.20%)
Total	52 (100.0%)	44 (100.0%)	96 (100.0%)

Note. At the time of the survey, no online fine arts courses had been offered by e^{4}TN.

Results of descriptive and inferential data analysis for research question 3.

Table 4.13 shows percentages of respondents agreeing that teacher effect scores were a good method for evaluating traditional and online teachers. Results showed that educators felt teacher-effect scores were a more appropriate way to assess teachers in the traditional environment than online environment; however, open-ended survey responses indicate that educators prefer a combination of assessment strategies such as teacher-effect combined with evaluation, portfolio, student feedback, and peer-feedback. Foreign language and social studies are the only subject areas where more educators said online teachers should be evaluated using teacher-effect scores as opposed to other subject areas where educators felt the evaluation method of teacher-effect scores was more appropriate in the traditional environment.

Table 4.13

Survey Respondent Perceptions on Whether Teacher Effect Scores Are a Good Method for Assessing Traditional and/or Online Teachers

	TVAAS: Good way to determine teacher quality	TVAAS: Only way to determine teacher quality
FTF environments	33.3% Strongly Agree/agree	0% Strongly Agree/agree
Online environments	18% Strongly Agree/agree	0% Strongly Agree/agree

A Wilcoxon test was used to compare responses on two survey items related to whether teacher-effect scores are a good way to measure teacher quality. One item asked whether teacher-effect scores were a good way of measuring teaching quality in the online environment and the other asked the same question about the traditional environment. The results showed a significant difference (z = -2.926, p =.003), indicating that respondents felt that teacher effect is a better way to measure teacher quality in the traditional environment than it is in the online environment. Based on the Likert-scale results (1=Strongly Agree and 5 =Strongly Disagree), the mean of the Wilcoxon ranks in favor of teacher effect as a way to measure teacher quality online was 10.65, while the mean of the ranks in favor of teacher effect as a way to measure teacher quality traditional was 9.67. Since a lower score indicates more agreement, there was greater overall agreement on use of teacher effect for traditional environments.

A Wilcoxon test was also used to compare responses on two other survey items related to use of teacher-effect scores. One item focused on whether teacher-effect scores should be the *only* way to measure teacher quality in the online environment and the other asked the same question about the traditional environments. The results showed no

significant difference (z = -1.091, p = .275), indicating that the respondents held the same opinion: that teacher-effect scores should not be the only way teacher quality ought to be measured both in the traditional and online environments. Based on the Likert-scale results (1=Strongly Agree and 5 =Strongly Disagree), as indicated in Table 4.13, no respondents supported using teacher effect as the *only* means of gauging teacher quality.

A Mann-Whitney *U* test was used to determine if there was a significant difference in teachers' and administrators' perceptions of using teacher-effect scores as a good method for evaluating teacher quality in the traditional classroom. Results were significant (z = -2.039, p = .041), indicating that administrators and teachers have different attitudes regarding the use of teacher effect scores in the traditional classroom. Based on the Likert-scale results (1=Strongly Agree and 5 =Strongly Disagree), the lower mean rank score indicates a more favorable response towards teacher effect scores as a good evaluation method for teachers in the traditional environment. The administrator mean rank was 26.75 and teacher mean rank was 34.82, indicating that respondents who were administrators viewed teacher-effect scores as a good measure for determining quality teachers in the traditional environment. Respondents who were non-administrators, however, tended not to rank teacher-effect scores as highly as did administrators for determining teacher quality in the traditional environment.

A Mann-Whitney *U* test also was used to determine if there was a significant difference in teachers' and administrators' perceptions of using teacher effect scores as the *only* method for evaluating teacher quality in the traditional classroom. Results were not significant (z = .000, p = 1.0). Neither group seems to feel that teacher-effect scores should be the only way teacher quality is measured in a traditional environment.

However, of those 21 who responded to the open-ended question, "Describe how you feel about the TVAAS approach and its appropriateness for evaluating teacher quality in the traditional, traditional teaching environment," one third gave answers that indicated clearly that they did not understand how teacher-effect scores are calculated. For example, some responded that a baseline should be established and then pre- and post-tests given for students in order to determine if a teacher is effective, which is the way scores are, indeed, calculated.

The evidence from these open-ended responses contradicts the self-reported Likert-scale responses, as shown in Table 4.14, when respondents were asked if they understood how teacher-effect scores were calculated. There is an observable difference between how educators responded, indicating that more respondents understood teacher effect scores and TVAAS after they read the survey explanation than before they read it. After reading the explanation of how Tennessee determines teacher-effect scores, no participants indicated in Likert-scale items that they didn't understand how the scores were determined. However, even after reading the explanation, their open-ended comments indicated that teachers and administrators did not, in fact, understand how teacher effect and TVAAS are calculated.

Table 4.14

Comparison of Respondent Understanding of Teacher Effect Before and After Reading the Survey Explanation

	Yes	Somewhat	No	Total
Prior	32 (50.8%)	16 (25.4%)	15 (23.8%)	63
Post	43 (68.3%)	20 (31.7%)	0	63

Note. Eleven of the74 total participants did not respond. Percents represent those who responded to this question.

Results for Research Question 4

Results of inferential data analysis for research question 4. The sample for research question 4 was the same as research question 3. Data in the form of comments from open-ended questions indicated support for evaluating teachers' effectiveness in both the online and traditional environments based on student feedback. The most common comments from participants stated that students are the clients, and they should be asked for feedback; teachers should be evaluated on how well they communicate with the student. Feedback regarding the same question in the traditional environment produced similar results including comments that focused on student improvement; student feedback; and that students, particularly adults, know when learning is taking place.

When participants were asked if they thought teacher-effect scores, which are based on TVAAS, are a good method of measuring teacher quality, a third of respondents gave positive feedback to their use in the traditional environment, while about half that number reported agreement with using these measures to evaluate teacher quality in the online environment. Thus, respondents seemed to feel teacher-effect scores are more appropriate for use in the traditional environment than in the online one. However, they also feel that teacher effect should never be the only method used to gauge teacher quality. All participants disagreed or were neutral to the idea that teacher-effect scores were the *only* way to assess teacher quality in both the online and traditional environments.

Survey participants were asked to rank methods to evaluate teacher effectiveness in the online and traditional environments from the following list: portfolio, observation, peer review, self-evaluation, teacher effect, or other. The majority of respondents seemed

to favor using a combination of items to evaluate teachers in both mediums. The top-ranked individual method was observation for both mediums, online and traditional. A summary of responses is given in Table 4.15.

Table 4.15

Respondent Ranking of Best Methods for Assessing Teachers in Traditional and Online Environments

Items rated	Number respondents Rating as 1		Number respondents Rating as 2		Number respondents Rating as 3	
	FTF	Online	FTF	Online	FTF	Online
Portfolio	1 (1.5%)	4 (6.2%)	2 (16.2%)	17 (26.2%)	3 (29.4)	17 (26.2)
Observation	25 (36.8)	23 (35.4%)	23 (33.8%)	15 (23.1%)	12 (17.6%)	14 (21.5%)
Peer Review	4 (5.9%)	7 (10.8)	15 (22.1%)	14 (21.5%)	13 (19.1%)	12 (18.5%)
Self-Evaluation	2 (2.0%)	1 (1.5%)	8 (11.8%)	10 (15.4%)	11 (16.2%)	10 (15.4%)
TVAAS	5 (7.4)	5 (7.7%)	9 (13.2%)	6 (9.2)	10 (14.7%)	10 (15.4%)
Combination	31 (45.6%)	15 (38.5%)	2 (2.9%)	3 (4.6%)	2 (2.9%)	2 (3.1%)
Total Respondents	68	65	68	65	68	65

Finally, participants were also asked to describe their perceptions about TVAAS and teacher effect as a method to evaluate teacher quality. To analyze the open-ended comments, a content analysis was done using a constant-comparison technique (Patten, 2005). The researcher found common themes from respondents regarding teacher quality indicators in the online environment. Educators state they are intimidated by teacher-effect scores as a method for evaluating online teachers because they feel that they have

little control in an online classroom; they do not create the curriculum, nor can they control some variables such as technology. They also consistently (24%) indicated that online teaching is very different than traditional instruction. They feel these factors should be considered when measuring a teacher's effectiveness, since they are afraid they will be penalized because of these factors if teacher-effect scores are applied to online courses. However, respondents also observed that, since students voluntarily take online courses, they may be more motivated to do well in courses, thereby enhancing teacher-effect scores. Another common theme that appeared from the open-ended comments was that a test score by itself is not indicative of teacher quality and a combination of evaluation methods would be more appropriate.

Summary of Results

Using a mixed-methods design, this study examined the feasibility of using teacher assessment scores from traditional teaching environments to measure quality of online teachers. Research questions 1 and 2 were addressed by comparing EOC scores from a systematically-selected sample of traditionally-taught students in a Tennessee school district with those from a convenience sample of Tennessee online students. The analysis also determined if program effects between online and traditional formats differed by subject area in Algebra I, English I, and Biology. Findings indicate that there are generally no significant differences in effects of online and traditional programs when results are controlled for prior differences between groups, though results are inconsistent in the algebra area (Patten, 2005).

To address research questions 3 and 4, a survey with open-ended questions and Likert-scale items was completed by 68 Tennessee educators. Likert-scale data was

analyzed with Mann-Whitney *U* and Wilcoxon tests (Patten, 2005). Open-ended items were examined using constant-comparison methods to analyze content for themes and patterns (Patten, 2005). Survey data indicated a lack of understanding for Tennessee's teacher-effect model, and a general perception that traditional teacher quality indicators cannot be used to assess teachers in the online environment. Results provide inconclusive evidence as to whether different models must be used to gauge teacher quality in traditional and online environments.

Chapter 5: Discussion

Overview of Discussion

Online learning is a growing field, and as more students become involved in online learning, measurements to indicate the quality of online instruction become of greater concern (SREB, 2006). This study examined teacher quality indicators in the online and traditional environments in order to determine if identifying a teacher in the traditional environment as effective is indicative or predictive of an effective teacher in the online environment. One way to examine this was by measuring and comparing students' end-of-course (EOC) scores in both the traditional and online environments to determine if there was any significant difference in the effects of the two programs. Results indicating no significant difference would suggest teachers in the two different environments could be evaluated using the same model since the program effects are the same. A significant difference, on the other hand, would indicate that similar methods could not be used to evaluate teachers in each program because the programs differ significantly. Wood (2008) says good traditional teachers are not necessarily good online teachers, even though some research says that online and traditional teachers share many of the same quality indicators such as knowing their respective content, being organized, having positive attitudes, having high verbal and written communication skills, and knowing student learning styles (Deubel, 2008; Education Week, 2004; Hammond & Prince, 2007; Pierce, 2008; Roblyer & McKenzie, 2000; Where we stand on teacher quality, 2004; Yang, 2005).

Tennessee evaluates traditional teachers using a value-added model (Hammond & Prince, 2007; Olson, 2005), a practice that is a growing trend as states look for a

quantifiable method to measure teacher effectiveness (Hammond & Prince, 2007). With increasing growth in enrollments in online education and quality indicators or standards defined by iNACOl and SREB (SREB, 2006; NACOL, 2008; Trotter, 2008), more quantifiable methods for evaluating online teachers are needed (Watson, Gemin, Ryan, & Wicks, 2009).

Purpose of the study. The purpose of the study was to shed light on whether or not educational systems could use traditional indicators of teacher quality such as value-added scores to select effective online instructors. This information could be useful to both Tennessee and other states that are searching for such measures to guide selection of online teachers.

If the EOC scores between traditional and online programs did not show a significant difference in programs, then this study would support using the same value-added scores to assess teacher quality in online courses as are used in traditional courses to assess teacher quality because the programs would show similar results. Similar results would indicate that comparable teaching processes for both programs are taking place; therefore, a common measurement for teacher assessment could be used. States could consider this as a method for teacher recruitment or incentive pay (Florida TaxWatch Center for Educational Performance and Accountability, 2007; Ryan & Watson, 2007). However, if the program showed differing EOC scores between traditional and online courses, then results would indicate that another method for assessing teacher quality in online courses must be found because something different is transpiring in the two programs.

If EOC scores differed significantly between traditional and online courses, the study would indicate that differing conditions between the two environments call for different ways of measuring teaching quality in online courses and traditional courses. If results of the two programs were similar, the study would provide data to support a more standardized, easy-to-calculate way of measuring teacher quality, a practice that has the potential to increase the quality of online programs and student instruction and a practice that is much needed (Marthers, Olivia, & Laine, 2008). Quantifiable teacher quality data for online teachers would allow administrators to make data-driven decisions to drive professional development plans and hiring decisions (Ballou, 2002; Braun, 2005). For example, better methods of evaluating teacher quality would assist administrators in deciding who would be the most effective teacher mentors and curriculum or pedagogy specialists.

Research questions. As online learning continues to grow, it is necessary to find a method for evaluating instruction to comply with NCLB requirements for highly-qualified, effective teachers (Crane, 2002; SREB, 2006). Tennessee has found a correlation between student achievement and teacher effectiveness, asserting that " … students given the most effective teacher for three years in a row made over twice the gains of comparable students assigned the least effective teachers" (Alliance for Excellent Education, 2008, p. 2; Sanders & Rivers, 1996). This study will provide evidence to determine whether such a relationship can be established in the same way with Tennessee's online instructors. The following research questions focus on the feasibility and practicality of using the same teacher quality indicators obtained in traditional environments as evidence of online teaching quality:

1) Is there a significant difference in program effects of traditional classrooms (as measured by end-of-course scores for a sample of traditionally-taught students in a Tennessee school district) and online classrooms (as measured by end-of-course scores for a sample of Tennessee's online students)? This question seeks to address whether or not the same value-added model can be used as the one used for teachers in traditional classrooms or if they need an alternative one. Significant differences in EOCs would indicate the presence of differing learning environment conditions (e.g., a different variety of learners in online classrooms vs. traditional classrooms, differing pedagogical requirements for effective online teaching than in a traditional classroom).

2) Do program effects between traditional and online environments vary significantly by subject area (i.e., Algebra I, Biology, and English 1)? According to data from past evaluations, students seem to do better online in some subject areas than others (Florida TaxWatch Report, 2008). This study will provide evidence on whether or not this trend is reflected in data from a Tennessee online program and, consequently, if it can be expected to be reflected in future data.

3) Do Tennessee educators perceive that Tennessee's model for teacher-effect scores can be used equally well in both traditional and online environments? If a value-added model is planned for use in online courses, it would be helpful to have the support of Tennessee educators for its use. Negative perceptions could affect recruitment of online teachers. For example, they may feel that they should not be assessed in the same way online as they are traditional and, therefore, would not be as likely to apply for such teaching opportunities.

4) What factors and strategies do teachers perceive should be considered in determining teaching quality in the traditional and online environments? Legislation related to online learning is in the infancy stages in Tennessee. The first legislation was passed on August 22, 2008. As the State of Tennessee becomes more aware of both the challenges and opportunities offered by online learning, this study will help identify educator perceptions of problems and potential concerns prior to the model's implementation, should the state choose to implement teacher-effect scores with online instruction.

Summary of results. Results from the study indicate that, when prior differences between groups are controlled statistically, there are no significant differences in programs in a Tennessee online program and a Tennessee traditional program, except in Algebra I, where differences were inconsistent depending on which covariate was considered. The results indicate that students show similar end-of-course performance in the online environment as in the traditional environment. The study also indicates that there is no significant difference in EOC scores by subject areas in English I and Biology. Results of an ANCOVA indicated that there were no significant differences in programs when contributions of two of the three covariates were controlled in Algebra. Therefore, results are inconclusive.

Results of the survey of the sample of Tennessee educators indicated that they felt strongly that teacher-effect scores should not be the only method in which teachers are evaluated for quality in either the online or traditional environments. They did think that it is better to evaluate traditional teachers than online teachers using teacher-effect scores. As a matter of fact, no respondents felt teacher-effect scores should be used to evaluate

teacher quality of online instructors. Administrators more frequently indicated approval of Tennessee's model of teacher-effect measures as a method for indicating teacher quality than did teachers. However, responses from the majority of survey participants indicated that they did not understand the factors or formula used to calculate teacher-effect scores. Tennessee educators feel that a variety of methods should be used to assess teachers such as student feedback, student scores, and teacher communication skills.

Limitations of the Study

Only students enrolled in Tennessee's e^{4}TN online courses with EOC exams in during the semesters of spring 2008 and spring 2009 in Tennessee were included in the study for the online sample. In addition, the students had to have NCE scores from grades 3-8. As a result of these data collection requirement in order to do the study, the sample of online students was small. Since the face-to-face population was selected to match the online sample, it was also small. The sample, therefore, is a small data set, specific to Tennessee, and results cannot be generalized to other programs. For the Likert survey and open-ended questions, only Tennessee educators were surveyed. This sample is not reflective of a national interpretation of value-added scores. Responses are specific to Tennessee educators regarding Tennessee models.

Interpretation of Findings

Differences in EOC scores between traditional and online programs in Tennessee were not found to be significant. Although there were differences found in programs in the subject area of algebra, these results were not consistent across analyses. Since the effect of prior performance, as indicated by NCE, was probably more important than the contributions of GPA and grade level. The fact that this comparison was not significant

should probably carry the most weight. This is true since the ANCOVA indicated that average NCE accounted for about a quarter of the variance between the programs in Algebra I, 69% in Biology, and 49% in English I (per Table 4.6).

Online students tend to take courses for different reasons than do traditional students. Many take online classes because they need to retake a traditional course due to a previous failure or because the school may not offer the course at a time that is convenient to their needs. Some take courses so that they can work at individualized paces (Watson, Gemin, Ryan & Wicks, 2009). Others take the online courses outside of the regular school day. It is reasonable to think a program that requires more independence by student-learners would attract a different type of a student than a traditional program. Since the Tennessee online student sample was a convenience sample, using this narrow sample may have biased the findings; however, this is somewhat unlikely since the researcher controlled for socio-economic status, race, gender, and grade in the purposeful sample of students from the traditional Tennessee school district. Although online students have been taking courses as recovery credit (i.e., they are repeating the course), the online participants had a mean GPA 0.08 points higher than the traditional students.

Online student GPA scores were self-reported by guidance counselors, whereas the sample of traditional students' GPA scores was pulled from the school district's student management system. Participants were not matched originally by GPA for the study; however, because it could impact the study, GPA was controlled for by the ANCOVAs. As indicated by Roblyer and Davis (2008), GPA was the most significant indicator in student success in online courses. Students with a higher GPA would be expected to

perform better than students with a lower GPA. Ideally, in a larger sample size, participants would also be matched by GPA. Finding no evidence of significant differences between online and traditional programs may also be a result of the small sample size from the convenience sample, because the online program is still in its early stages of development (McMillan & Schumacher, 2006). This limited convenience sample, in comparison to a much larger traditional sample, may also contribute to observed program effects. As the online program grows, a purposeful sample including students with matched GPAs may yield more reliable results. Students who take online courses self-select into the online environment. Therefore, they may be more inclined to be independent learners or prefer the option of online learning environments (Watson, Gemin, Ryan, & Wicks, 2009).

Finding no significant difference in program outcomes may indicate no core difference in the instructional environments. Even though there are some unique quality indicators for online instructors, online instruction requires similar pedagogical indicators as traditional instruction: being masters of time, being knowledgeable of content, being strong communicators and being able to analyze student responses and data, and being flexible (Berge & Clark, 2005; Furey & Murphey, 2005; Hassel & Terrell, 2004; Russell, 2004).

Standards for online teaching quality have been prepared by various organizations (SREB, 2006; NACOL, 2008; Trotter, 2008). The similarity in programs found by this study contradicts the general assumption that traditional teaching quality does not necessarily predict online teaching quality (Wood, 2008). Participants in the study were all associated with Tennessee's online learning program; therefore, they were all

recommended to teach by an administrator. Additionally, e^{4}TN trains all staff on best practices in online learning. Having all survey participants previously trained in best practices of online learning and being recommended by an administrator would suggest that they are strong in pedagogy and curriculum in both the traditional and online environments. Using these kinds of teachers, therefore, may also be a limitation of the study. Curriculum and program design also impact pedagogy, so these results may not generalize to another online environment.

Finally, it is evident from survey data that Tennessee educators have strong views against using value-added data to measure online teacher quality. Those objections, combined with their deficiency in understanding the value-added model, create a difficult political climate for using a value-added model to assess teachers in the online environment.

Implications for Practice

Although results were not consistent across all indicators, results overall indicated no significant differences between programs in at least two content areas. If these results hold true in future research, then the same data for teacher evaluation in the traditional environments can be used in online environments. At the same time, it seems clear that additional pedagogical skills are probably needed for teachers in the online environment than in the traditional environment. Teachers in the online sample had this additional training, and results of the survey data indicated that such training is important. For example, educators' responses in the open-ended comments section of the survey indicated that online teachers should be evaluated on communication with online students. A sample of this practice would be evaluating teachers' online exchanges with

students for effective communication. Educators also shared a concern that curriculum in online courses may not be as easy to modify as in the traditional setting, leaving online teachers at a disadvantage not faced by traditional teachers. The online value-added formula should probably account for student completion of online curriculum materials and time active in the online course, as well as the varying contributions of hybrid versus asynchronous courses. Per the ANCOVA results, the model would also need to control for prior ability by including variables such as student GPA and NCE.

In the traditional teacher-effect model in Tennessee, exceptional education students are not included in teacher-effect scores. Since the current online program allows for students to self-select and allows educators to enroll students with varying needs and abilities, it seems necessary that students with exceptional needs may need to be considered as a population to be included in the formula. Approximately 20% of the samples for this study were made up of exceptional education students for the semesters of spring and fall for two years in Algebra I, Biology and English 1. Excluding 20% of the population reduces the opportunities to run a teacher-effect model in the online environment. After four years of Tennessee's online program, the sample size is still currently too small for a true teacher-effect model, unless exceptional students are allowed to be included. Using Tennessee’s traditional teacher-effect model in the online environment will be delayed until the sample size is larger. Excluding exceptional education students from the calculation for the online program will delay this opportunity, perhaps for several years.

Some exceptional education students may see success in the online environment because of the technology and differentiation in instruction (Bransford, Brown, &

Cocking, 2000). The technology associated with online learning allows for students with diverse learning needs to participate in curriculum that meets their individual learning styles. They also have flexibility and can learn at their own pace (Watson, Gemin, Ryan, & Wicks, 2009). Therefore, there seem to be several reasons to consider including exceptional education students in a modified version or model of teacher-effect in the online environment.

Results of this study also have implications for teacher training that addresses the additional skills needed for online teaching. Pedagogical training is key for online instructors (McKenzie, Mims, Bennet, & Waugh, 2000; Pallof & Pratt, 2001), but training that addresses the value-added model is also needed, as evidenced by open-ended responses to the survey items. Educator comments indicated a lack of understanding of how teacher-effect score is calculated in Tennessee.

To address training in online pedagogy, teacher preparation programs should focus on social aspects of online learning, course facilitation, technology skills, communication, as well as strong content knowledge. Since communication in online courses can be nonverbal, teacher education programs need to train future educators to interact with students to convey emotions and human interaction. Additional trainings for online instructors are crucial, since teachers cannot see body language to determine student engagement or comprehension. Teachers must understand how to evaluate discussion boards and implement various strategies that make best use of the online environment. Meeting the needs of each student through the medium of technology requires different approaches than meeting needs in a face-to-face environment. Pedagogical training for the online environment also needs to be a focus of teacher

preparation programs, since the pedagogical characteristics of the program affect the success of students and the overall effectiveness of the program (Cassidy, 2004; Deubel, 2008; Education Week, 2004; Hammond & Prince, 2007; Pierce, 2008; Roblyer & McKenzie, 2000; Where we stand on teacher quality, 2004; Yang, 2005).

As online learning becomes more readily accessible to students, and they self-select into courses in traditional and online programs, some students may be more successful in the online environment than in the traditional environment, at least in some subject areas, as indicated by the significant difference found in programs in Algebra I and other current studies (Florida Tax Watch Report, 2008; Watson, Gemin, Ryan, & Wicks, 2009). This would imply that access to online courses should be readily available to students with diverse socioeconomic backgrounds, academic backgrounds, gender, and grade levels and in a variety of subject areas. It also implies that some students may be more successful in online courses than others, given their prior abilities and learning styles.

Qualitative data indicated that Tennessee educators may not understand the current TVAAS formula or teacher-effect model used to evaluate traditional teachers in Tennessee. Administrators should consider this when developing professional development for all educators in Tennessee. TVAAS data contains valuable information for student prediction indicators and advisement regarding student ability. Teachers can use prior NCE scores and prediction scores to assess student knowledge and abilities. Additionally, teachers need to understand the evaluation model under which they are being measured. Teacher preparation programs in Tennessee should consider including training on this model.

Implications for Future Research

Further research needs to be done with larger samples and in other online programs. Based on results from this study, it seems clear that these future studies should control for prior ability as part of the study design. Additionally, future research should control for the exceptional education variable in online and traditional environments. A larger sample would allow a researcher to determine if students with exceptional needs in specific areas see a positive gain in EOC test scores more frequently in online, rather than in traditional settings. Since ANCOVA results found a significant difference in Algebra I with GPA as the covariate, and The Florida Tax Watch (2008) and The Sloan Consortium (2009) found that students may be successful in different subject areas online, future studies need to be done to identify what consistent factors are characteristics of successful online students.

The original intent of the researcher was to compare teacher-effect scores of Tennessee teachers who taught the same courses (Algebra I, Biology, and English I) in both the online and traditional environments, However, the population of teachers in Tennessee who have taught online and traditional courses in the same subjects was too limited, based on the criteria for inclusion established by SAS. The researcher also would have liked to evaluate differences in teacher-effect scores across courses in the online and traditional environments to determine if teachers are equally effective in the online and traditional environments across content areas (Patten, 2005). As online learning grows and the online teaching population increases in Tennessee, future studies would allow for comparing teachers who teach the same EOC subjects both online and in the traditional setting to see if they show similar teacher-effect scores. This would allow administrators

to determine more definitively if traditional instructors are equally effective in the online environment, with comparable students.

Finally, as hybrid courses become more popular due to growing technology integration in the classroom, budget constraints, and teacher shortages (Watson, 2008), a future study could measure effective instruction in content areas in hybrid, online, and traditional settings in order to determine differences in program effect. For example, two school systems may partner to share a teacher by providing the primary curriculum online and providing synchronous access to an instructor via web cam technology, as opposed to offering only asynchronous courses. This increased communication and higher level of interaction with students may create a more effective learning environment for students. It will also provide equity for school districts and allow for teachers to have more control of the curriculum, which they indicated as a concern in the open-ended survey.

Summary of Discussion

The purpose of this study was to explore the feasibility of using the same method Tennessee currently uses to gauge quality of teaching in traditionally-delivered courses to the quality of teaching in the online environment. Research questions were:

1) Is there a significant difference in program effects of traditional classrooms (as measured by end-of-course scores for a sample of traditionally-taught students in a Tennessee school district) and online classrooms (as measured by end-of-course scores for a sample of Tennessee's online students)?

2) Do program effects between traditional and online environments vary significantly by subject area (i.e., Algebra I, Biology, and English 1)?

3) Do Tennessee educators perceive that Tennessee's model for teacher-effect scores can be used equally well in both traditional and online environments?

4) What factors and strategies do teachers perceive should be considered in determining teaching quality in the traditional and online environments?

Findings indicate that there is no significant difference in EOC scores in online and traditional programs in Biology and English when prior differences between groups were controlled statistically. However, there is a significant difference in EOC scores in Algebra I, depending on which covariate was considered. The significant difference in Algebra I is most likely attributed to the small sample size; therefore, the results need to be compared with those from future studies.

The majority of the data contradicts the responses from Tennessee educators regarding the use of the same evaluation method for online teachers as traditional teachers. Finding no significant difference in Tennessee online and traditional programs in two of three subject areas suggests that there should be no difference in how teacher quality is measured. Since similar results were found in the programs, then it would be logical that similar evaluations would be used for the programs. However, survey participants, who were all Tennessee educators, all agreed that the Tennessee teacher-effect model used to evaluate traditional instruction should not be the only method used to evaluate online instructors. Since pedagogy in the online environment may be different than in the traditional environment, additional types of online-teacher evaluation(s) may need to be implemented to measure teaching quality indicators in the online environment (SREB, 2003; Treacy, 2007; Berge & Clark, 2005; Hassel & Terrell, 2004; Furey & Murphey, 2005; National Center for Educational Statistics, 2005; Russel, 2004). Findings

may also suggest that some students are more successful in the online environment, and opportunities should be provided for them to select the online environment rather than the traditional environment (Carr, 2000; Diaz, 2002; Roblyer, 2006).

Since the results were inconsistent, future studies should be done to determine if the findings hold true with larger sample sizes. A more detailed study should be performed using teacher-effect scores, with the assistance of TDOE and SAS, as more online teachers who also teach the same end-of-course subjects in Tennessee in the traditional environment enter the teaching population. Teacher-effect scores in each environment could be compared to determine if equivalent teacher-effect scores are achieved by the same teacher in both environments. Future studies could control for exceptional education status, if large enough samples exist or if SAS changes its criteria for the online teacher-effect model. Further studies should be done that address hybrid, online, and traditional models to measure program effect in these environments. If no significant differences continue to be found between traditional and online programs when prior ability for GPA is controlled, then there will be further evidence that online learning and traditional learning are equally effective.

References

Alliance for Excellent Education (2009). *Measuring and improving the effectiveness of high school teachers*. Washington, DC.

Azukas, L., & Steiner, S. (2009). Assessing online instructors: The virtual observation. *Instructional Leader*. Retrieved from http://www.uwex.edu/disted/conference/Resource_library/handouts/05_1947H.pdf

Ballou, D. (2002, Summer). Sizing up test scores. *Education Next, 2*(2). Retrieved from http://www.hoover.org/publications/ednext/3365706.html

Ballou, D., Sanders, W. & Wright, P. (2004, Spring). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics, 29*(1), 37-65.

Barker, K., & Wendel, T. (2001). E-learning: *Studsxvying Canada's virtual secondary schools.* Kelowna, BC: Society for the Advancement of Excellence in Education.

Barrett, J. (1986). *The evaluation of teachers* (Report No. SP 028 569). Washington, DC: Office of Educational Research and Improvement. (ERIC Document Reproduction Service No. ED 278 657)

Berge, G.A. (2002). *Why distance learning: higher education administrative practices*. Washington, DC: American Council on Education.

Berge, Z. L,. & Clark, T. (2005). *Virtual schools planning for success*. New York: Teachers College Press.

Berman, P. (Ed.). (2007*).* E-Learning. *In P. Berman (Ed.), E-learning concepts and techniques (pp. 26-41) Bloomsburg: Institute for Interactive Technologies.* doi:10.1002/piq.20029

Bielick, S., Chandler, K. & Broughman, S. (2001, July). *Home schooling in the United States 1999-2001* (NCES No. 2001-033). Washington, DC: U.S. Department of Education, National Center for Educational Statistics. Retrieved from http://nces.ed.gov/pubs2001/2001033.pdf

Bransford, J., Brown, A., & Cocking, R. (Eds.). (2000). *How people learn: Brain, mind, experience and school.* Washington, DC: National Academy Press.

Braun, J. (1986). The evaluation of teachers. *Eric Digest*. 12. Retrieved from http://www.ericdigests.org/pre-925/evaluation.htm

Braun, H. I. (2004). *Using student progress to evaluate teachers: A primer on value-added models*. Princeton, NJ: Educational Testing Service.

Braun, H. I. (2005). *Value-Added modeling: what does due diligence require*? Princeton, NJ: Educational Testing Service.

Bock, R. D., Wolfe, R., & Fisher, T. (1996). Review & analysis of the Tennessee value added assessment system (part one). *Educational Researcher*, *37*(2), 65-75.

Carr, S. (2000, February 11). As distance education comes of age, the challenge is keeping the students. *The Chronicle of Higher Education*, *23*(A1). Retrieved from http://chronicle.com/free/v46/i23/23a00101.htm

Cassidy, S. (2004). Learning styles: An overview of theories, models, and measures. *Educational Psychology*, *24*(4), 419 - 444.

Cavanaugh, C.S. (2001). The effectiveness of interactive distance education technologies in K-12 learning: A meta-analysis. *International Journal of Educational Telecommunications,7* (1), 73-88.

Cavanaugh, C.S., Gillan, K.G., Kromrey, J., Hess, M., & Blomeyer, R. (2004). *The effects of distance education on K-12 student outcomes: A meta-analysis*. Naperville, I11: Learning Point Associates. Retrieved from http://www.ncrel.org/tech/distance/k12distance.pdf

Clark, R.E. (1983). Reconsidering research on learning from media. *Review of Educational Research,53* (4), 445-449. doi:10.3102/00346543053004445

Clark, R. E. (1985). Evidence for confounding in computer-based instruction studies: analyzing the meta-analyses. *Educational Communications and Technology Journal,33* (4), 249-262.

Clark, T. (2000). *Virtual high schools: State of states*. Macomb, IL; Western Illinois University, Center for the Application of Information Technologies. Retrieved from http://www.cait.org/shared_resources.docs/vhs_files/vhs_study.pdf

Clark, T. (2003). Virtual and distance education in American schools. In M.G. Moore, & W.G. Anderson (Eds.). *Handbook of distance education* (pp. 673-699). Mahwah, NJ: Lawrence Erlbaum.

Clark, T. (2001). *Virtual schools: Trends and issues. A study of virtual schools in the United States.* Phoenix, AZ: Distance Learning Resource Network at WestEd. Retrieved from http://www.wested.org/online_pubs/virtualschools.pdf

Crane, J. (November, 2002). *The promise of value-added testing.* Progressive Policy Institute. Retrieved from www.ppionline.org

Daniel, J. S. (1995) *The mega- universities and the knowledge media: implications of new technologies for large distance teaching universities*. MA Thesis, Concordia University, Montreal, Canada.

Darling-Hammond, L. (2007). Recognizing and enhancing teacher effectiveness: A policymaker's guide. In L. Darling-Hammond & C. D. Prince (Eds.), *Strengthening Teacher quality in high-need schools-policy and practice* (pp. 1 and 26*).* Washington, DC: The Council of Chief State School Officers.

Delivery methods for distance education. (2007, May 11). Retrieved from Distance Education Clearinghouse: http://www.uwex.edu/disted/delivery.cfm

Deubel, P. (2007, April). What's online education all about? *THE Journal*. Retrieved from http://www.thejournal.com/articles/20516

Diaz, D.P, (2002, May-June). Online drop rates revisited. *The Technology Source*. Retrieved from http://ts.mivu.org/default.asp?show=article&id=1034

U. S. Distance Learning Association (USDLA). *Distance learning.* (n.d.). Retrieved from USDLA at http://www.usdla.org/Glossary_Distance.pdf

Dwyer, N., Joseph, S., Medina, R., Suthers, D., & Vatrapu, R. (2008). Beyond threaded discussion: representational guidance in asynchronous collaborative learning environments. *Computers & Education, 50*(4), 1103-1127

Education Week (2004, September 21). *Teacher quality*. Retrieved from http://www.edweek.org/rc/issues/teacher-quality/?print=1

Eckes, M., & Law, B. (2000). *The more-than-just surviving handbook ESL for every classroom teacher* (2nd ed.). Winnipeg, Canada: Portage and Main Press.

Eisner, E. W. (2002). *The educational imagination: On the design and evaluation of school programs* (3rd ed.). Upper Saddle River, NJ: Merrill.

Elmore, R. (2002, September-October). Testing trap. *Harvard Magazine, 105,* 35-39.

Feigenoff, C. (2006). Metrics for high quality teaching. *Metrics for High Quality Teaching*

Explorations. Retrieved from http://oscar.virginia.edu/explorations/x8006.xml

Florida TaxWatch Center for Educational Performance and Accountability. (2007). Final report: A comprehensive assessment of Florida Virtual School. Tallahassee, Florida. Available at http://www.floridataxwatch.org/resources/pdf/110507FinalReportFLVS.pd

Furey, D., & Murphey, E. (2005). An introduction to K-12 virtual schools and schooling: Factors enabling growth and impacts on traditional systems, schools and roles. *The Morning Watch: Educational and Social Analysis, 32,* 3 and 4.

Gemin, B., Ryan, J. & Watson, J. (2008). *Keeping pace with K–12 online learning: A review of state-level policy and practice.* Evergreen, CO: Evergreen Consulting.

Glass, G. V. (2009). *The realities of K-12 virtual education*. Boulder and Tempe: Education and the Public Interest Center & Education Policy Research Unit. Retrieved from http://epicpolicy.org/publication/realities-k-12-virtual-education

Glossary of terms and abbreviations used in Tennessee's value-added assessment reports (n.d.). Retrieved March 9, 2010, from Education Consumers Foundation website: http://www.education-consumers.org/Tennesseeproject/vaglossary.htm

Goldstein, J. (2007, May). Easy to dance to: Solving the problems of teacher evaluation with peer assistance and review. *American Journal of Education, 113*(3), 479-508. doi:10.1086/512741

Gordon, R., Kane, T., & Staiger, D., (2006). *Identifying effective teachers using performance on the job*. Washington, DC: The Brookings Institution.

Gore, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.

Green, S.B. & Salkind, N.J. (2005) (4th ed.). *Using SPSS for windows and macintosh: Analying and understanding data.* Upper Saddle River, NJ: Pearson Education.

Hammond, D.L. (1986, March). Policy initiatives for developing a teaching profession. *The Elementary School Journal, 86*(4), 530-551.

Hammond, D.L. & Prince, C.D. (2007, July). *Strengthening teacher policy quality in high-need schools-policy and practice.* A report prepared under a grant from The Joyce Foundation to the Council of Chief State School Officers. Retrieved from http://ti_sp.alsde.edu/qt/Shared%20Documents/Strengthening%20Teacher%20Quality%20in%20High%20Need%20Schools-Policy%20and%20Practice.pdf

Hassel, B. & Terrell, G. (2004). Managing virtual schools: The Canadian experience. In C. Cavanaugh (Ed.), *Development and management of virtual schools: Issues and trends* (pp. 50-67). Hershey, PA: Information Science Publishing, Idea Group Inc.

Horton, K. & Horton, W. (2003). *E-learning tools and technologies*. Wiley Publishing Inc., Indianapolis.

Hunter, W., & Smith, R. (2002) Virtual schooling: integrating schooling into technology. In B. Barrell (Ed.), *Technology, teaching and learning: Issues in the integration of technology* (pp.197-219). Calgary, AB: Detselig Enterprises.

Jain, K.K. (2002). Motivating e-learners – A case study of e-learners seeking an advanced degree from Unitar. *Student Affairs, 3*(4). Retrieved from http://www.studentaffairs.com/ejournal/Fall_2002/motivatingelearners.html

Kahn, R. (1994). The role of government in the evolution of the internet. *Communications of the ACM, 37*(9), 15-19. doi:10.1145/179606.179729

Kupermintz, H. (2003). Teacher effects and teacher effectiveness: a validity investigation of the Tennessee value added assessment system, *Education Evaluation Policy, 25*(3), 287-298. doi:10.3102/01623737025003287

Lengeling, M. (1996, Spring). The complexities of evaluating teachers. *Mextesol Journal, 19*(4), 15-22.

Marthers, C. Olivia, M. & Laine, S.W. (2008, February). Effective teacher evaluation: Options for states and districts. *National Comprehensive Center for Teacher Quality*. Retrieved from http://www.tqsource.org/publications/February2008Brief.pdf

McKenzie, B. K., Mims, N., Bennett, E., & Waugh, M. (2000, September 25). Needs, concerns and practices of online instructors. *Online Journal of Distance Learning Administration, 1*(3). Retrieved from http://www.westga.edu/~distance/ojdla/fall33/mckenzie33.html

McMillan, J., & Schumacher, S. (2006). *Research in education. Evidence-based inquiry.* Boston: Pearson.

Mills, S.C. (2002). School isn't a place anymore: *An evaluation of virtual Greenbush online courses for high school students*. KS: Southeast Kansas Education Service Center.

Mitchell, S. C. (1923, June). For the 90 percent. *School Review, 31*, 444-493.

Morabito, M. G. (1999). *Online distance education: Historical perspective and practical application*. Universal Publishers/uPUBLISH.com.

Moran, L. (1993). Genesis of the open learning institute of British Columbia. *The Journal of Distance Education, (8)* 1, 43-70. doi:10.1007/978-1-4020-8299-3_3

Murphey, E., & Rodríguez-Manzanares, M. (2009). Research on the e-teacher in the K-12 Distance Education Classroom. In P. Rogers, G.Berg, J,Boetcher, C. Howard, L. Justice, & K. Schenk (Eds.), *The Encyclopedia of distance and online learning* (2nd ed.) *Information Science Reference.*

National Center for Educational Statistics (NCES). (2005, March). Distance education at post-secondary institutions:2002-2003. Washington, DC: US Department of Education.

National Forum on Educational Statistics. (2006). Forum guide to elementary/secondary *virtual education* (NFES 2006-803). Washington, DC: U.S. Department of Education.

NACOL (October, 2008). NACOL fast facts about online learning. Retrieved from http://www.inacol.org/

National Forum on Educational Statistics. (2006). Forum guide to elementary/secondary virtual education (NFES 2006–803).Washington, DC: U. S. Department of Education.

Paloff, R. M., & Pratt, K. (2001*). Lessons from the cyberspace classroom the realities of online teaching*. San Francisco: Jossey-Bass.

Patten, M.L. (2005). *Understanding research methods: An overview of the essentials.* Glendale: Pyrczak.

Peterson, K.(2004, June). Research on teacher evaluation. *NASSP Bulletin, 88*(639), 60-79. doi:10.1177/019263650408863906

Pierce, D.P. (2008, November). These traits make online teachers successful: Clarity, creativity and compassion are key characteristics for online instructors says UCF researcher. *eSchool News*, Retrieved from http://www.eschoolnews.com/news/top-news/?i=55930

Roblyer, M.D., & Doering, A. (2010) (5th ed.). *Integrating educational technology into teaching. (5th Ed).* Boston, MA: Allyn & Bacon

Roblyer, M. D., Davis, L., Mills, S., Marshall, J., & Pape, L. (2008). Toward practical procedures for predicting and promoting success in virtual school students. *The American Journal of Distance Education, 22*(2), 90-109. doi:10.1080/08923640802039040

Roblyer, M. D., Marshall, J., & Mills, S. (2006). *When are students "virtually successful? Testing a model to identify at-risk virtual high school students*. Paper presentation proposal for the 2007 Annual Conference of the American Educational Research Association.

Roblyer, M. D. (2006). Virtually successful: Defeating the dropout problem in online school programs. *Phi Delta Kappan, 88*(1), 31-36.

Roblyer, M. D. & McKenzie, B (2000). Distant but not out-of-touch: What makes an effective distance learning instructor? *Learning and Leading With Technology, 27*(6), 50-53.

Russell, G. (2004). The advantages and disadvantages of virtual schools and of *ICT in Education*, *7*(1), 6-8.

Russell, G. (2005). Virtual Schools: A critical view. In C. Cavanaugh (Ed.), *Development and management of virtual schools: Issues and trends* (pp. 1 and 25). Hershey, PA: Information Science Publishing, Idea Group Inc.

Ryan, J., & Watson, J. (2006, October). *Keeping pace with K-12 online learning: A review of state level policy and practice*. Retrieved from http://www.nacol.org/

Ryan, J., & Watson, J. (2007). *Keeping pace with K-12 online learning: A review of state level policy and practice*. Retrieved from http://www.kpk12.com/downloads/KeepingPace_2007.pdf

Sanders, W. L., Saxton, A. M., & Horn, S. P. (1997). The Tennessee value-added system: A quantitative outcomes-based approach to educational assessment. In J. Millman (Ed.) *Grading teachers, grading schools: Is student achievement a valid measure?* (pp. 247-256). Thousand Oaks, CA: Corwin Press, Inc.

Sanders, W., & Rivers, J. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville, Tennessee: University of Tennessee Value-added Research and Assessment Center.

Savery, J. R. (2005). Be VOCAL: Characteristics of successful online instructors, *Journal of Interactive Online Learning*, *4*(2), 141-152.

Smith, R. (2000). *Virtual schools in the K-12 context.* Unpublished doctoral dissertation, The University of Calgary, Calgary, Canada. Retrieved from http://66.51.101.171/rosina/dissertation.doc

Southern Regional Education Board (SREB). (2003). *Essential principles of high-quality online teaching: Guidelines for evaluating K-12 online teachers.* Atlanta: SREB.

Southern Regional Education Board (SREB). (2006). *Essential principles of high-quality online teaching: Guidelines for evaluating K-12 online teachers.* Atlanta: SREB.

Southern Regional Education Board (SREB). (2006). *Online teaching evaluation for state virtual schools.* Atlanta: SREB.

Taylor, J. (2001). Fifth generation distance education. Higher education series, 40, 1-7.

Trotter, A. (2008). Voluntary online-teaching standards come amid concerns over quality. *Education Week*, *27*(26), 1-15.

Treacy, B. (2007, October). What's different about teaching online? How are virtual schools changing teaching? *Kentucky Virtual High School Newsletter*, *1*(2).

Tsai, I-C., Kim, B., Liu, P.-J, Goggins, S.P., Kumalasari, C., & Laffey, J.M. (2008). Building a model explaining the social nature of online learning. *Educational Technology & Society*, *11*(3), 198-215.

Tucker, P., Stronge, T., Gareis, C. & Beers, C. (2003). The efficacy of portfolios for teacher evaluation and professional development: Do they make a difference? *Education Administration Quarterly*. *(39)* 5, 572-602. doi:10.1177/0013161X03257304

Tucker, B. (2007, June). Laboratories of reform: virtual high schools and innovation in public education. *Education SectorReports*. Retrieved from http://www.educationsector.org

Tucker, B. (2009, Summer). Florida's online option. *Education Next. 9*(3) Retrieved from http://www.hoover.org/publications/ednext/floridas_online_option.html

U.S. Census Bureau. (2004, March 18). *Census Bureau projects tripling of Hispanic and Asian populations in 50 years.* Retrieved from http://www.census.gov/Press-Release/www.releases/archives/population/001720.html

Watson, J. (2008). *Blended learning: The convergence of online and traditional education.* Washington, DC: North American Council for Online Learning. Retrieved from http://www.inacol.org/research/promisingpractices/NACOL_PP-BlendedLearning-lr.pdf

Watson, J., & Ryan, J. (2007). *Keeping pace with K-12 online learning: A review of state-level policy and practices*. Retrieved from http://www.nacol.org/docs/KeepingPace07-color.pdf

Watson, J., Gemin, B, Ryan, J. &Wicks, M. (2009). *Keeping pace with K-12 online learning: A annual review of state-level policy and practices.* Retrieved from http://www.kpk12.com/

Where we stand on teacher quality. (2004). Princeton, NJ: Educational Testing Services. Retrieved from http://www.ets.org/Media/Education_Topics/pdf/teacherquality.pdf

Wood, C. (2005). Highschool.com. *Edutopia, 1*(4), 32–37.

Yang, Y. (2005, Spring). Preparing instructors for quality online instruction. *Online Journal of Distance Learning Administration, 8* (1). Retrieved from http://www.westga.edu/~distance/ojdla/spring81/yang81.htm

Zandberg, I., & Lewis, L. B. (2008). *Technology-based distance education courses for public elementary and secondary school students: 2002-3 and 2004-5 statistical analysis report (NCES 2008-008).* Washington, DC: National Center for Educational Statistics.

Zucker, A., & Kozma, R (2003). *The virtual high school: Teaching generation V.* New York: Teachers College Press.

Appendices

Appendix A

Informed Consent and Likert Survey

Likert Survey – Oliver – will be collected electronically through a survey generator

Date

Dear _____,

I am a student under the direction of Dr. M.D. Roblyer, Professor in the Learning and Leadership doctoral program at the University of Tennessee at Chattanooga. I am conducting a research study titled: *Measuring and Comparing Teaching Quality in Online and Traditional High School Classroom Environments*.

Please click on the link to respond to a 22-item anonymous online survey if you are a teacher and a 19-item anonymous online survey if you are an administrator, which will take approximately 10 minutes. You may choose not to participate or discontinue the survey at any time. There are no foreseeable risks to your participation in the study. Your completion of the anonymous survey will constitute your informed consent to participate.

If you have any questions concerning the research study, please call me at 423-209-8810 or email me at oliver_wendy@hcde.org.

This research has been approved by the UTC Institutional Review Board (IRB). If you have any questions concerning the UTC IRB policies or procedures or your rights as a human subject, please contact Dr. M.D. Roblyer, IRB committee Chair, at (423) 425-5567 or email instrb@utc.edu.

Completion of the survey will be considered your consent to participate. Please refer to the following explanations for the Tennessee Value Added Model as you complete the survey. Login information is provided below. Thank you.

Sincerely,

Wendy L. Oliver
150 Templeton Lane
Hixson, Tennessee 37343

Login Address:

As a Tennessee teacher or administrator you are probably familiar with TVAAS, which is a way Tennessee assesses student gains in knowledge and teacher effectiveness. This study is evaluating TVAAS and teacher effect scores as a way to assess online

instruction. I would appreciate your candid answers to determine if and how TVAAS scores should used to evaluate online teachers.

Here is how TVAAS scores are currently calculated and used:
TVAAS scores for Tennessee students are calculated by using their end-of-course test results. A predicted "gain score" is calculated for each student based on his/her achievement over several previous years. The teacher receives a teacher effect score by comparing her/his students' scores in the year s/he taught them to their predicted ones.

The teacher receives points or a teacher effect score by adding “gain” to a student’s learning or may lose points or remain neutral by not helping the student make a predicted increase. The sum of these scores over the group of students the teacher taught is considered the "teacher effect score."

For background on how TVAASS is calculated and used, please refer to the explanation in the email you received with log-in information for the survey.

1. Please indicate your gender.
a. Male
b. Female

If you are currently an administrator in Tennessee, please skip to question ***7***.
If you are a currently a teacher in Tennessee, please begin with question ***2***.

Questions 2-9 apply to the ***traditional/traditional*** or brick and mortar setting.

2. Are you currently teaching a secondary course (grades 6-12) in Tennessee in the traditional/traditional environment?
a. Yes
b. No

3. If you answered Yes to number 1, in what content area are you teaching?
a. English
b. Fine Arts
c. Foreign Language
d. Health and PE
e. Math
f. Science
g. Social Studies
h. Career and Technology
i. Other
j. Not Applicable/I'm not teaching in a traditional environment at this time.

4. If you marked an area in 2, is your license in this area:
a. Apprentice
b. Professional
c. Alternative
d. Interim
e. I do not have a current license.
f. Not Applicable/I'm not teaching in a traditional environment at this time.

5. How many years have you taught secondary education in *Tennessee* in the *traditional/ traditional environment?*
a. 0-1
b. 1-5
c. 6-10
d. 11 and 20

e. 21-30

6. How many years have you taught secondary education in the *traditional/ traditional environment anywhere*?

f. 0-1
g. 1-5
h. 6-10
i. 11 and 20
j. 21-30

(Teachers will be branched to Question 10.)

7. What type of administrative license do you hold?

a. Beginning
b. Professional

8. How many years have you served as an administrator in secondary education in *Tennessee?*

k. 0-1
l. 1-5
m. 6-10
n. 11 and 20
o. 21-30

9. How many years have you served in administration in secondary education in the *anywhere*?

p. 0-1
q. 1-5
r. 6-10
s. 11 and 20
t. 21-30
u.

10. Prior to this survey did you understand how teacher effect scores were calculated in Tennessee?

a. yes
b. somewhat
c. no

11. After this survey do you understand how teacher effect scores are calculated in Tennessee?

a. yes
b. somewhat
c. no

Instructions for questions 12 and 13. **The following questions relate to how to evaluate traditional/traditional instruction ONLY.** Mark the answer that indicates how much you agree with the statement.

12. TVAAS scores are a good way to determine teacher quality in the traditional/traditional setting.

a. Strongly Agree
b. Agree
c. Undecided
d. Disagree
e. Strongly Disagree

13. TVAAS scores should be the ONLY way that *traditional, traditional teachers* are evaluated.

a. Strongly Agree
b. Agree
c. Undecided
d. Disagree
e. Strongly Disagree

14. Place a 1, 2, and 3 beside the top 3 ways you feel teacher effectiveness should be measured in the traditional, traditional classrooms:

a. Portfolio
b. Observation
c. Peer Review
d. Self-Evaluation
e. TVAAS
f. Combination of all or any of the above. Please specify.
g. Other. Please specify.

15. Use numbers 1 to 9 to rank of the following from MOST VALUABLE (1) to LEAST VALUABLE (9) in determining the effectiveness of teachers in traditional, traditional classrooms.

a. Student outcomes
b. Professional development/training
c. Scores on evaluation observations for teachers
d. Professional years of experience
e. Level of education
f. Grade point average in school
g. Technology skills
h. Communication skills
d. Other. Please specify.

16. Describe how you feel about the TVAAS approach and its appropriateness for evaluating teacher quality in the traditional, traditional teaching environment.

The following questions apply strictly to the online or e-learning environment.
If you are currently in an administrative role with e-learning in Tennessee, please skip to question ***22.***

If you are a currently a teacher in Tennessee, please begin with question ***17***.

17. Are you currently teaching secondary courses in Tennessee in the *online* or e-learning environment?
a. Yes
b. No

18. If you answered Yes to number 11, in what content area do you teach?
a. English
b. Fine Arts
c. Foreign Language
d. Health and PE
e. Math
f. Science
g. Social Studies
h. Career and Technology
i. Other
j. Not Applicable/I'm not teaching in a traditional environment at this time.

19. If you marked an area in 18, is your license in this area:
a. Apprentice
b. Professional
c. Alternative
d. Interim
c. I do not have a current license.
d. Not Applicable/I'm not teaching in a traditional environment at this time.

20. How many *semesters* have you taught *online in the State of Tennessee*?
a. 0-3
b. 4-7
c. 8-10
d. 11-13
e. 14+

21. How many *semesters* have you taught *online anywhere*?
a. 0-3
b. 4-7
c. 8-10
d. 11-13
e. 14+

(Teachers will be branched to Question 25.)

22. What type of administrative license do you hold?

a. Beginning
b. Professional
c. Not applicable

23. How many *semesters* have you served in an administrative role with e4TN, the State of Tennessee's online learning initiative?

f. 0-3
g. 4-7
h. 8-10
i. 11-13
j. 14+
k. Not applicable

24. How many *semesters* have you served as an *online* administrator *anywhere*?

f. 0-3
g. 4-7
h. 8-10
i. 11-13
j. 14+
k. Not applicable

Instructions for questions 25 & 26. **The following questions relate to how to evaluate online or virtual learning ONLY.** Mark the answer that indicates how much you agree with the statement.

25. TVAAS scores are a good way to determine teacher quality in the online or virtual environment.

a. Strongly Agree
b. Agree
c. Undecided
d. Disagree
e. Strongly Disagree

26. TVAAS scores should be the only way *online* teachers are evaluated.

a. Strongly Agree
b. Agree
c. Undecided
d. Disagree
e. Strongly Disagree

27. Place a 1, 2, and 3 beside the top 3 ways you feel teacher effectiveness should be measured in online, virtual classes.

a. Portfolio
b. Observation
c. Peer Review
d. Self-Evaluation
e. TVAAS
f. Combination of all or any of the above. Please specify.
g. Other. Please specify.

28. Use numbers 1 to 9 to rank the following from MOST VALUABLE (1) TO LEAST VALUABLE (9) in determining the effectiveness of teachers in online, virtual classes.

a. Student outcomes
b. Professional development/training
c. Scores on evaluation observations for teachers
d. Professional years of experience
e. Level of education
f. Grade point average in school
g. Technology skills
h. Communication skills
i. Other. Please specify.

29. Describe how you feel about the TVAAS approach and its appropriateness for evaluating teacher quality in the online teaching environment.

Vita

Wendy Oliver was born in Chattanooga, Tennessee to the parents of Thomas Oliver and Henny O'Bryant. She has two sisters, April and Lindsey Oliver. She attended Hamilton County Schools in Chattanooga, Tennessee graduating with honors in 1994. A dedicated student-athlete, as an undergraduate she attended The University of Tennessee at Chattanooga for two years and then completed a BA at Lincoln Memorial University in December of 1997. After a successful academic and athletic undergraduate career Wendy returned to Chattanooga to teach and coach. Opportunities in administration spurred continuing education for Wendy, and she graduated with her M.Ed. from Lincoln Memorial University in 2000. As an inspiring leader and advocate for technology in education Wendy found interest in and completed an Educational Specialist degree in Leadership and Educational Technology. In late 2005 Wendy was a founding member of the State of Tennessee's online learning program, e^4TN, and is continuing to serve in leadership roles such as the Tennessee Distance Learning Association's Board of Directors. She is continuing her education with a terminal degree in Leadership in Learning from the University of Tennessee at Chattanooga.