Summer 2017

Evaluation and Student Learning Objectives in Connecticut Schools

Vittoria Fielosh
Sacred Heart University, vfielosh@gmail.com

Evaluation and Student Learning Objectives in Connecticut Schools

Vittoria Fielosh

Sacred Heart University

Advisor: Dr. Charles Britton

Abstract

Over the past several years, states across the country have been restructuring their teacher evaluation models due to changes in federal education policies. The focus of many of these models has become student achievement, and there are many ways states are measuring a teacher's impact on student achievement. Connecticut, along with more than 20 other states, uses student learning objectives to measure a teacher's impact on student learning. Unlike other states, Connecticut has not specified what assessment be used to measure student growth. This case study, of one school district in Connecticut, examines the type of assessments used as a part of student learning objectives. This study also examines teacher and administrator's perceptions of the process, and finally, student learning objectives will be assessed to determine if student learning objectives meet the state criteria outlined by the Connecticut State Department of Education. Electronic surveys, including both open-ended and close-ended questions, were developed and distributed to teachers and administrators in the participating district. Quantitative data was analyzed using descriptive statistics while qualitative was coded inductively. Findings suggest teachers have mixed feelings regarding the SLO process as it relates to teacher evaluation. Several types of assessments are being used to measure student growth, and many student learning objectives developed by teachers and administrators do not meet the SMART goal criteria set forth by the state.

Table of Contents

Chapter 1: Introduction

Background

In recent years, teacher evaluation has been at the center of education reform. In an effort to improve student learning, the Department of Education has encouraged states to take a closer look at teacher evaluation as effective teachers have been found to have the greatest impact on student success (Warring, 2015). Federal policies, like No Child Left Behind and Race to the Top, linked grant funding to education reform. As a result, states made changes to their teacher evaluation plans to meet the demands of the newly instated policies. Teacher evaluation models were greatly influenced by the Measure of Effective Teaching Project (MET). Based on the MET Project, funded by the Gates foundation, states developed teacher evaluation models that evaluated teachers using multiple measures that include student achievement, classroom observations, and student feedback (Gates Foundation, 2009).

Connecticut began to restructure its teacher evaluation plan in 2012. At that time, 14 school districts in Connecticut volunteered to pilot the new teacher evaluation plan known as System for Educator Evaluation and Development (SEED) (Donaldson, et al, 2014). Connecticut's SEED includes multiple measures to assess teachers as suggested by the MET Project. According to SEED, teacher performance and practice would make up 40% of a teacher's evaluation, parent feedback would account for 10%, student growth and development would account for 45%, and whole school student learning/student feedback would make up the remaining 5% (CSDE, 2015). Over the past five years, districts throughout the state have implemented new teacher evaluation plans that included student learning objectives (SLOs). However, unlike other states, Connecticut has not mandated standardized measures be used for developing student learning objectives. SEED does however require that SLOs meet the criteria for SMART goals in that SLOs or Indicators of Academic Growth and Development (IAGD) be

specific, measurable, attainable, result-oriented, and time-bound (CSDE, 2014). The question then becomes, what measures are Connecticut teachers using to develop their SLOs to measure student growth and development and do they meet the SMART criteria established by the State Department of Education?

Statement of Problem

Student growth and development makes up 45% of a teacher's overall, annual evaluation in the state of Connecticut. Connecticut does not mandate standardized assessments be used for this 45%, however, Connecticut suggests that 22.5%, or one SLO, focus on student growth on a standardized measure if there is a standardized measure available for the content being taught. The remaining 22.5% can measure student growth using standardized or non-standardized measures (CSDE, 2014). Having said that, there has also been much debate at the state level regarding the use of the Smarter Balanced Assessment (SBA) scores to evaluate teachers. Finally, in April 2017, Connecticut's Department of Education came to the conclusion that SBA scores will not be used to evaluate teachers (CSDE, 2017). Since the assessments used to measure teacher effectiveness are not the focus of the plan, and Connecticut educators do not have one universal assessment to measure student growth, therefore the SLO process becomes the focus of the 45% for the evaluation model. Connecticut SEED explains the SLO process as follows:

> SLOs are carefully planned, long-term academic objectives. SLOs should reflect high expectations for learning or improvement and aim for mastery of content or skill development. SLOs are measured by Indicators of Academic Growth and Development (IAGDs) which include specific assessments/measures of progress and targets for student mastery or progress. Research has found that educators who set high-quality SLOs often

> realize greater improvement in student performance. The SEED model asks teachers to set more specific and measurable targets. (CSDE, 2014, p. 25)

Though SLOs are to be broad statements, IAGDs should be written using SMART goal language (CSDE, 2014).

Teachers throughout the State of Connecticut are developing SLOs with guidance and approval from their administrators. A variety of instruments are being used to measure student growth because Connecticut's State Department of Education has not mandated that statewide assessments be tied to teacher evaluation. Across districts, grades, and content areas, there is great opportunity for variation in this process.

Purpose of the Study

Though many studies have been conducted looking at teacher evaluation throughout the United States, there is little research that focuses specifically on the SLO process. This case study will examine the SLO process within two Connecticut school districts to gain a better understanding of how this process is impacting teaching and learning. Though there are over 160 school districts in the State of Connecticut, the district participating in this study will serve as a representative sample for the following reasons: the district services students in grades K-12, all teachers will have the opportunity to participate in the case study across grades and content areas, the district is following the SEED model set forth by the state, and the participating district is a DRG D district. The state of Connecticut classifies school districts using District Reference Groups (DRGs). This classification system groups schools together based on similar socioeconomic status (CSDE, 2006). DRGs range from A-I. Furthermore, the district participating in this study volunteered to do so hoping that the findings would benefit their district.

The findings of this study will provide a sampling of SLOs used by Connecticut teachers. It will also provide insight relating to the SLO process from the teacher's and administrator's point-of-view. Lastly, this case study will look at sample SLOs to determine if SLOs and IAGDs are meeting the SMART goal criteria set forth by CT SEED.

The overall purpose of this study is to gain a better understanding of the SLO development process and if it has had a positive impact on teaching and learning. The findings of this case study may encourage other districts to take a closer look at the SLO process in their districts. As stated earlier, many researchers have looked at teacher evaluation, however, when looking specifically at SLOs, research is lacking. This case study is being conducted to gather information in this area and provide research where it may be lacking.

Significance of the Study

According to Lacireno-Paquet, Morgan, and Mello (2014), 30 states in the United States have included SLOs as one component of their teacher evaluation system. Teacher evaluation has undergone may changes within the last five years. As educators, it is crucial that we examine this process closely as it impacts teachers and students in more than half of our states. This study will will examine the SLO process and hopefully provide data and evidence to help inform the teacher evaluation process, at minimum, in the participating district.

Research Design

This research will take the form of a case study, more specifically, a collective case study examining the SLO process in two school districts in the state of Connecticut. Surveys developed mirrored those used in a similar study conducted in Arizona and Utah in 2015 by Makkonen, Tejwani, and Rodriguez. Surveys will be made available to all teachers and administrators in the participating districts electronically using *Google Forms*.

Surveys were made up of both open-ended and close-ended questions. Both qualitative and quantitative data will be collected and analyzed. Teacher surveys contain both types of questions, while the administrators' survey contains only close-ended questions. Participants will be asked to respond to close-ended questions using a five-point Likert scale. Open-ended questions, ask teachers to provide their SLOs. SLOs will be coded and evaluated to see if they meet the SMART goal criteria.

Reliability and validity will be established by using methods triangulation, triangulation of sources, and peer debriefing. Qualitative data will be compared to quantitative data. Data from each school will be analyzed separately and then the data will be compared to determine if there are themes, trends, similarities, or differences.

Research Questions

1. When setting their SLOs, what assessments did Connecticut teachers use to measure student growth during the 2016-2017 school year?

2. What are teachers' perceptions of the impact of SLOs on student achievement and teacher evaluation?

3. What are administrators' perceptions of the impact of SLOs on student achievement and teacher evaluation?

4. Do SLOs developed by teachers and administrators meet the SMART goal criteria set forth by CT SEED?

Assumptions and Limitations

Assumptions and limitations for this case study relate to the survey developed to collect data and the level of participation in the study. An electronic survey, using *Google Forms*, will allow the researcher to collect the most responses in a time effective manner. This method also

allows for quantitative and qualitative data to be collected. The survey consists of fewer than 15 questions, most of which are close-ended, to encourage participation. As with most research, there are also limitations within this case study. First, teachers may not provide their SLOs when asked or they may provide incomplete SLOs. Second, teachers may be skeptical about answering questions honestly because findings will be reported to their district administrators. Likewise, this may deter teachers from participating in the study. Finally, since data will be collected from two school districts, the data collected may not be representative of the entire state.

Definitions of Terms

Student Learning Objective (SLOs)- "Student learning objectives are broad statements about the knowledge and skills a teacher wants students to demonstrate as a result of instruction, address the central purposive a teacher's assignment, take into account baseline data on student performance, pertain to a large proportion of a teacher's students, reflect content mastery or skill development, and reflect ambitious but attainable goals for student learning." (CT SEED, 2013) (Lacireno-Paquet, et. al, 2014)

Indicators of Academic Growth and Development (IAGDs) include specific assessments/measures of progress and targets for student mastery or progress to measure SLOs. (CT SEED, 2014, p. 25)

SMART Goals- Connecticut SEED states that IAGDs should be written using SMART goal language, S = Specific and Strategic, M = Measurable, A = Aligned and Attainable, R = Results-Oriented, T = Time-Bound (CT SEED, 2014, p. 29)

Expected Findings

Based on my experience as a teacher, a review of teacher evaluation plans and a comprehensive review of literature, I expect that this study will result in mixed findings. Prior to

developing research questions for this study, I reviewed teacher evaluation plans for approximately ten districts in Connecticut. A takeaway from that process was that there is variation in plans throughout the state. Though all districts follow the basic outline set forth by Connecticut's SEED model, there are differences in how district's implement the model. Further, it is expected that a variety of assessments used to measure student achievement as outlined in SLOs and IAGDs. The state guidelines specify that one SLO must be based on a standardized measure, but only if there are standardized measures available in the content area in which you teach (CSDE, 2015). Based on a previous study looking at SLOs, Makkonen et. al (2015), teachers used a variety of assessments to measure student achievement, for example, district-wide assessments, teacher-made assessments, and student portfolios to name a few. This case study will be surveying teachers in a K-12 district who teach various grades and content areas, therefore, findings may suggest that teachers in Connecticut are also using a variety of measures to indicate student growth. Though the majority of teachers in Connecticut have been successful based on the new teacher evaluation plan, survey results may indicate that not all teachers and administrators find the SLO process to be of benefit to teaching and learning. Lastly, it is expected that the majority of SLOs will not meet each of the SMART goal criteria outlined by CT SEED.

Chapter 2: Literature Review

Introduction: Teacher Evaluation: How Did We Get Here?

Over the past 10 years, there have been significant changes to educational policies in our country. From No Child Left Behind to Race to the Top, the adoption of Common Core Standards and high stakes testing, many students, teachers, and schools have been resilient in their attempt to adapt to these changes. In 2013, the US Department of Education extended No Child Left Behind waivers and 34 states were allowed to move ahead with their plans to "prepare all students for college and career, focus aid on the neediest students, and support effective teaching and school leadership" (USDE, 2013). In addition, Race to the Top also encourage reform relating to teacher evaluation. States across the country restructured their teacher evaluation models to align with policy changes to ensure that their schools would receive federal grant funding (Donaldson & Papay, 2015). Based on the Measure of Effective Teaching Project (MET), funded by the Gates Foundation, states developed teacher evaluation models that assess teachers using multiple measures that include student achievement, classroom observations, and student feedback (Gates Foundation, 2009). Various instruments are being used to measure student achievement and in turn to evaluate teachers. Value-added measures (VAM), state tests, standardized assessments, and student learning objectives (SLOs) are examples of ways states are evaluating teachers' effectiveness. When looking specifically at SLOs, there are also variations among states, districts, schools, and teachers in how SLOs are developed and the types of measures used to evaluate student progress (Lacireno-Paquet, Morgan, & Mello, 2014). Measures for student learning objectives may include, but are not limited to, standardized assessments, state tests, vendor generated assessments, value added measures, district developed assessments, classroom assessments, portfolios, and task completion.

In 2012, 14 school districts in Connecticut piloted a new teacher evaluation plan known as System for Educator Evaluation and Development (SEED) (Donaldson et al., 2014). Connecticut's SEED includes multiple measures to assess teachers and the configuration is as follows: teacher performance and practice 40%, parent feedback 10%, student growth and development 45%, and whole school student learning/student feedback 5% (CSDE, 2015). Since then, districts throughout the state have implemented new teacher evaluation plans that included student learning objectives. However, unlike other states, Connecticut has not mandated standardized assessments be used for developing student learning objectives. The state's assessment, Smarter Balanced Assessment (SBA), would only account for about 25% of teachers in grades K-12 because the SBA is only given in grades 3-8 in English Language Arts and math. The question then becomes, what measures are Connecticut teachers using to develop their SLOs to measure student growth and development?

The focus of this literature review was teacher evaluation and student learning objectives. The following key terms were used to search the *Education Resource Information Center* (ERIC) database: teacher evaluation, assessments, student learning, student learning objectives, Connecticut, value-added assessments, and Gates Foundation. Searches were limited to articles that were peer reviewed and published between 2010 and 2016. The search was limited to this time frame to encompass the implementation of new teacher evaluation models in Connecticut.

Teacher Evaluation in Connecticut

What measures of student achievement are Connecticut teachers using to develop student learning objectives as part of Connecticut's System for Educator Evaluation and Development? When looking at Connecticut public schools, there are great variations in what teachers teach

across grades and across content. The use of SLOs allows Connecticut's SEED to address the diversity in teachers' roles and responsibilities and address each teacher's effectiveness.

However, *The Hartford Courant*, in February 2016, reported that 98% of the state's teachers were rated as proficient or exemplary, the top two ratings out of four, for the 2013-2014 school year. This begs the question, is SEED an effective measure of teacher evaluation? Since SLOs make up the majority of a teacher's overall rating, people question whether the SLOs that are being developed by teachers and their evaluators are rigorous enough to accurately measure teachers' effectiveness. ELA is an area that is assessed by statewide assessments, but are ELA teachers using these standard measures? Further, a study was conducted by Neag researchers from the University of Connecticut's Center for Education Policy Analysis in 2013 to collect data in districts that piloted SEED the previous school year (Donaldson, et al., 2014). Relating specifically to SLOs, this study found teachers expressed mixed views on whether SLOs changed their practice, only 55% of administrators surveyed indicated that setting SLOs led teachers to make changes in their teaching practice, and the selection of IAGD targets is an area that deserves close attention as the SEED model evolves (Donaldson, et al., 2014). Future research will ask: what assessments did Connecticut teachers use as the Indicators of Academic Growth and Development (IAGD), from 2015-2016, when setting their Student Learning Objectives, what are teachers' perceptions of the impact of SLOs on student achievement and teacher evaluation, and what are administrators' perception of the impact of SLOs on student achievement and teacher evaluation?

The New Phase of Teacher Evaluation

In the Fall of 2009, the Bill & Melinda Gates Foundation started the Measures of Effective Teaching (MET) project to examine new approaches to evaluating effective teaching.

There was concern relating to the high percentage of teachers that were rated satisfactory or above using the evaluation models that were in place at that time. The purpose of the MET project was to take a closer look at how teachers were evaluated and to look specifically at how teachers' effectiveness impacted student achievement. This was a four-part project that began with this initial study. The goal of the project was to improve the quality of information about teaching effectiveness, to help build fair and reliable systems for teacher observation and feedback (Gates Foundation, 2009). For years, researchers have found that a teacher's effectiveness has the greatest impact on student success when compared to other factors (Warring, 2015). Taking this into consideration, the MET project looked at multiple measures to evaluate a teacher effectiveness. The project gathered data from 3000 teachers in six school districts: Charlotte- Mecklenburg Schools, Dallas Independent School District, Denver Public Schools, Hillsborough County Public Schools (including Tampa, Florida), Memphis City Schools, and the New York City Department of Education (Gates Foundation, 2009). The multiple measures included student achievement, classroom observations, student feedback, and lastly, these measures should promote teacher growth and development. For this study, student achievement was measured using state assessments and three supplemental assessments, that were also standardized measures, in grades four through eight in, English language arts and math (Gates Foundation, 2009). Classroom observations were conducted through the use of videos to enable multiple observers to use multiple scoring measures. Student feedback was acquired using the Tripod survey instrument, developed by Harvard researcher Ron Ferguson, which assessed whether students experience their classroom teacher to be engaging, demanding, and supportive of their learning (Gates Foundation, 2009). As noted earlier, this being the first of four parts of the MET project, the research only looked at student achievement in grades and subjects that are

assessed at the state level. The authors noted that there is much more to study, however, the research suggested multiple measures should be used to assess the effectiveness of teachers.

Critics of the MET Project

Kane (2012) reviewed the first two phases of the MET project in an article titled, "Capturing the Dimensions of Effective Teaching." He noted that the best predictor of a teacher's success from year to year, was a teacher's success in prior years. He also noted that teachers with larger student achievement gains, also seem to have students with greater long term success (Kane, 2012). When looking at the multiple measures used to evaluate teaching, student achievement was a more reliable predictor of the achievement of future students than classroom observations or student surveys. Several studies have indicated that teachers' effectiveness has the greatest impact on student success, but is student achievement on state assessments the best way to measure student gains or teacher effectiveness? Thomas Kane noted that there are limitations to using standardized assessments to measure student achievement and to evaluate teachers. For example, in many districts less than 25% of teachers teach in grades or subjects where student achievement can be measured by a state assessment. In addition, student achievement linked to state assessments does very little to help teachers develop their skills and grow professionally (Kane, 2012). Tests given once a year do little to inform instruction because the data on the students is reviewed after the students have matriculated through the school year and have a new teacher. If these measures are being used to evaluate teachers, will this data provide teachers enough information to inform their instruction? Kane (2012) noted the absurdity of teachers received feedback on student performance after-the-fact when he stated, "A performance-evaluation system should support growth and development, not just facilitate

accountability. Teachers need to be able to see their own strengths and weaknesses clearly and recognize where they need to hone their skills." (Kane, 2012 p. 36)

Kim Marshall, author of *Rethinking Teacher Supervision and Evaluation,* also examined the findings of the MET project. Marshall states that the effectiveness of a multiple measures approach will depend on how classroom observations, achievement data, and student feedback are used (Marshall, 2012). Though Marshall agrees that one classroom observation is not sufficient to evaluate a teacher's performance, he does not feel four observations will give evaluators enough insight to what children experience daily sitting in class, and the paperwork involved with each formal evaluation "creates an impossible workload for administrators." (Marshall, 2012). Marshall suggests, more frequent, informal observations, followed up with face-to-face coaching by the same evaluator, allows administrators to observe a variety of lessons and provides a clearer picture of a teacher's performance.

Marshall also states that using standardized assessments to evaluate teachers is "highly problematic." The following concerns were raised by Darling-Hammond, Amrein-Beardsley, Haertel, and Rothstein (2012), David (2010), and Goldhaber and Hansen (2008) as noted by Marshall (2012 p. 51):

- Standardized tests were not designed to evaluate teachers.
- For reliable data, districts will need three years of value-added scores and three years is an unacceptable time frame to support struggling teachers.
- High stakes testing tied to teacher evaluation may lead to teaching to the test verses teaching skills students need to grow and develop as students.
- Evaluating teachers using test scores may divide staff.

- How do we evaluate the 75% of our teachers that do not teach in grades or subjects where standardized tests are available?

- Is one teacher truly responsible for student success on standardized tests? How do we account for specialists, interventionists, support staff, teachers from previous years, and other educators our students work with?

Marshall describes a better way to incorporate student achievement by establishing learning outcomes for students, using formative assessments to guide instruction, and he even suggests that teams of teachers work together to focus on improving student learning. Teachers and administrators would discuss what measures they would use to assess student achievement. The team approach would also serve as a professional development opportunity as teachers work together to improve teaching and learning.

Lastly, Marshall suggests that student feedback should not be used to evaluate teachers. Student feedback is relevant, but there should be a conversation between teachers and their evaluators about student responses. Teachers should then be evaluated based on how they respond to the feedback (Marshall. 2012).

A Closer Look at Student Achievement in Teacher Evaluation

Given teachers' effect on student learning and achievement, policy makers and researchers have called for an increase in rigor and quality of teacher evaluation models (Donaldson, 2010). Because of Race to the Top funding, many states redesigned their teacher evaluation systems and passed legislation linking student performance to teacher evaluation. Further, student achievement influences teacher tenure and termination. Legally, this is cause for concern because research has not been able to show reliability and validity in the assessments used to measure student achievement. More than 20 states have required teacher evaluation

plans, which include "objective measures of student achievement growth" and make up 40% to 50% of a teacher's overall rating (Baker, Oluwole, & Green, 2013). Are the assessments used both valid and reliable?

Value-added measures (VAMs) and student growth percentiles (SGPs) are two types of measures used across the country to measure student achievement in new teacher evaluation models (Baker, et al., 2013). VAMs use pre and post assessment data, from the beginning of the year to the end of the year, to measure the value a teacher adds to a student's achievement as opposed to other factors (Baker, et al., 2013). On the other hand, SGPs do not take other factors into consideration, but they do compare a student's growth to that of other students who performed similarly on the same assessment.

Mark Ehlert, Cory Koedel. Eric Parsons, and Michael Podgursky (2014) looked at how to best measure student achievement for the purpose of school and teacher evaluation by looking at three methods: student growth percentile, one-step value added model, and two-step value added model. They defined SGP by stating, "SGPs calculate how a student's performance on a standardized test compares to the performance of all students who received the same score in the previous year." To calculate SGPs for schools or specific classes, the median SGP was used for that group. A concern with SGP is that it does not control for student characteristics like race and socioeconomic status. The one-step value added measure also looks at student growth, but does control for the following characteristics: race, gender, free or reduced lunch eligibility, English-language-learner, special education, and grade level (Ehlert, et al., 2014). The two-step value added measure controls for differences among students first and then controls for differences between schools. When looking at math scores from the Missouri Assessment program from 2007-2011, SGP and one-step value added both showed a negative correlation between student

performance and socioeconomic status. The two-step value added model eliminated any relationship between school poverty rates and student growth (Ehlert, et al., 2014). The authors cautioned that using SGP may cause teachers who work with disadvantaged students, to be unfairly rated as underperforming and suggest that high-poverty school performance should not be compared to low-poverty school performance.

There are additional limitations to both SGP and VAM models. First, these measures are only available in grades and subject areas where there is mandated annual testing (Warring, 2015). As mentioned earlier, statewide standardized assessments address only 25% of what our K-12 teachers teach. Next, this data does little to inform teacher improvement or professional development. While identifying effective or ineffective teachers, these measures do not identify what factors make a teacher effective (Warring, 2015). Lastly, the focus on test scores can unintentionally lead teachers to teach to the test. In states and districts were VAMs are used to rank teachers based on their effectiveness, test scores may also lead to teachers competing against each other instead of working together for what is best for students (Warring, 2015).

How Do We Evaluate Teachers Who Teach in Non-Tested Areas?

One of the most significant limitations of using standardized measures to assess student achievement, for the purpose of evaluating teachers, is the fact that more than 50% of public school teachers do not teach in subjects or grades where standardized assessments are available. As quoted in Geo and Holdheide (2011, p.1), Prince, et al. (2009) stated,

> Identifying highly effective teachers of subjects, grades, and students who are not tested with standardized achievement tests- such as teachers of art, music, physical education, foreign language, K-2, high school, English language learners, and students with disabilities- necessitates a different approach. It is important that states and districts

provide viable options for measuring the progress of these groups of students and the productivity of their teachers, both of which contribute to school performance.

Though federal guidelines do not spell out what these measures should be; however, they do specify that measures used to determine a teacher's contribution to student learning must be rigorous, must show growth between two points in time, and must be comparable across classrooms (Geo & Holdheide, 2011). Several states and districts have attempted to address this problem and have used various approaches though there are issues with each when looking at cost, reliability, and validity (Geo & Holdheide, 2011). The following is a list of options states and districts are using to assess student growth in non-tested areas: use existing tests designed for other purposes like-end-of course tests, create new tests in areas where assessments are lacking, use the four Ps (portfolios, products, performances, or projects), give teachers of non-tested areas a prorated score for collaborating with teachers in a tested subject, do not use student achievement for these teachers, or use student learning objectives (Geo & Holdheide, 2011). Research to determine which method is most valid or reliable is limited. Geo and Holdeide (2011) notes that research cannot keep up with the changes in teacher evaluation because changes are happening at such a fast pace.

Student Learning Objectives

Student learning objectives (SLOs) are becoming a popular alternative measure because they can be used to evaluate teachers in any grade or subject (Gill, Bruch, & Booker, 2013). The SLO process, as defined by Race To the Top Technical Assistance (2010 p. 1), is "a participatory method of setting measurable goals, or objectives, based on the specific assignment or class, the needs of students being taught, the subject matter taught, the baseline performance of the students, and the measurable gain in student performance during the course of instruction." A

case study by Lacireno-Paquet, Morgan, and Mello (2014) reviewed state websites and took a close look at the use of SLOs in teacher evaluation plans. Thirty states reported using SLOs in their teacher evaluation models. Twenty-five states defined SLOs, the definitions varied, but most did include that student growth and achievement had to be measurable. It is interesting to see that there are many similarities across states, yet there are also many differences in how SLOs are used to evaluate teachers. Of particular interest are the various assessments that are used or required to develop SLOs and measure student achievement. In 21 states SLOs apply to all teachers and require an evaluator to approve SLOs. In three states SLOs only apply to teachers in subjects or grades in which student progress cannot be assessed using state tests. The three most common types of SLOs were SLOs for individual teachers, SLOs for teams or grade levels, and school-wide SLOs. The types of assessments used to measure student growth included state standardized assessments, vendor-developed assessments, district-developed assessments, school-wide and classroom measures (Lacireno-Paquet, et al., 2014).

Taking a closer look at the implementation of student learning objectives, Makkonen, Tejwani, and Rodriquez (2015) looked at how Arizona and Utah used SLOs in a pilot program. SLOs set clear learning expectations for students, use formative assessments to track student progress, and allow for differentiated instruction based on student progress (Makkonen, et al., 2015). This study was conducted during the 2013-14 school year in both states as they implemented a pilot test of the SLO- based teacher evaluation process. Three hundred sixty-three teachers from four districts in Arizona participated in the process. Student learning objectives were collected and coded to determine what assessments teachers were using to determine student growth. Types of assessments included vendor-developed, teacher developed, national or

state standardized, district or school-wide. Types of goals set included demonstrates a skill, meets growth target, meets test benchmark, or meets proficiency (Makkonen, et al., 2015).

In Arizona, most SLOs focused on students attaining a certain score or mastery on a vendor-developed test versus what students would learn or what strategies teachers would use in their classrooms (Makkonen, et al., 2015). According to this study, when evaluating teachers using SLOs, performance spanned all four scoring levels, with four being the most common score (1= low, 4=high). Different types of SLOs also yielded different scores. Performance based SLOs, like playing a piece of music or running a mile is a specific amount of time, received fours more frequently than other SLOs. Elementary school teachers also scored higher than secondary school teachers. In addition, there was a correlation among the three components of the teacher evaluation model which were SLOs, classroom observations using the Danielson Framework, and student feedback surveys (Makkonen, et al., 2015).

Utah's pilot, though similar to Arizona's, was different in many ways. Utah's pilot began in January, so it only spanned half a school year and only 82 teachers participated with more than half being special educators. While teachers in Arizona used vendor-developed assessments, 73% of Utah's teachers focused their SLOs on project completion, 23% focused on students demonstrating a physical skill, and they used teacher-developed assessments. Eighty-nine percent of Utah's teachers met or exceeded expectations for their SLOs (Makkonen, et al., 2015).

This study also surveyed teachers in Utah as to their perceptions of the SLO process. Fifty-two percent of teachers who responded to the survey felt that this was a worthwhile process and 66% did not want to be held accountable for school-wide test scores in place of SLOs. Though some results were positive, teachers did not feel that the SLO process helped them develop professionally or benefitted students (Makkonen, et al., 2015).

Synthesis of Research Findings

As we look at the MET project and its recommendations for changes in the teacher evaluation system, we perhaps uncover more questions than answers. One piece of research that is consistent is that teachers are the one factor that have the greatest impact on student success. How we measure that impact is a question that remains to be answered. Student achievement, in many states, accounts for the majority of a teacher's overall rating. States across the country are using different measures to assess teachers' impact on student achievement. Value-added measures, student growth percentiles, and student learning objectives are all being used to measure student growth and, in turn, evaluate teachers. Research has revealed that all three models have benefits, but they also have limitations. One of the major issues with value-added measures and student growth percentiles is that they do not allow us to assess more than half of our country's public school teachers because the assessments are limited to English language arts, math and, in some states, social studies and science.

Looking specifically at student learning objectives, more than half of the states in the country have incorporated SLOs in their teacher evaluation systems. However, there are inconsistencies in the way that they are implemented. Teachers are using a variety of assessments to measure student achievement. States are allowing teachers to use a variety of assessment tools from skill completion to portfolios to standardized vendor-developed assessments. The fact that there is variability in the measures teachers are using can be viewed as both a benefit and a limitation. Student learning objectives allow teachers to customize student learning outcomes to fit the parameters of the content and the students they are teaching. Teachers and evaluators are developing their SLOs and, in many cases, formative assessments are used to monitor student progress as part of the SLO.

Research looking specifically at SLOs is limited at this time. Lacireno-Paquet, Morgan, and Mello (2014) reviewed state websites to gain a better understanding of which states are using SLOs and how they are using them. The research conducted by Makkonen, Tejwani, and Rodriquez (2015) provides insight to assessments used and teachers' perceptions of the process. In both Utah and Arizona, the SLOs depended on different assessments. Other states, like New York, require teachers to use state assessments or standardized measures for SLOs (Forman & Markson, 2015).

> In the state of Connecticut, forty-five percent of a teacher's evaluation is based on attainment of goals and/or objectives for student growth, using multiple indicators of academic growth and development to measure those goals/objectives. The process for assessing student growth using multiple indicators of academic growth and development for teacher evaluation will be developed through mutual agreement by each teacher and their evaluator at the beginning of the year. (CSDE, 2015)

With 30 or more states using SLOs to evaluate teachers, do we have enough data to say this is a valid and reliable approach? Perhaps we should take a closer look at the measures being used to gage student achievement and if those measures are valid and reliable. More research needs to be done to learn more about the implementation of student learning objectives, their validity in the teacher evaluation process, and the impact SLOs have on student learning.

Summary

Teacher evaluation has been a major focus of the many changes that our educational system has faced over recent years. States have been revamping their evaluation systems to meet the new federal requirements. Though student achievement must be a part of a state's teacher evaluation program the federal guidelines do not spell out how student achievement must be

measured. Therefore, student achievement has taken many different forms. States are using value-added measures, student growth percentiles, state assessments, student learning objectives, or a combination of the aforementioned. The state of Connecticut has implemented its teacher evaluation model using student learning objectives. Currently, the state has not mandated a specific measure be used to assess gains in student achievement. Even though the state is using the Smarter Balanced Assessment in English language arts and math in grades 3-8, teachers in those areas do not need to use this data to develop their SLOs. This study will look at what measures teachers are using to assess student learning for the purpose of teacher evaluation. It will also look at whether teachers and administrators perceive the process to be effective for improving both teaching and learning.

Chapter 3: Methodology

Purpose

Though there is much research relating to teacher evaluation, there is limited research looking specifically at using SLOs to evaluate a teacher's effectiveness. The purpose of this study was to determine what measures are being used by Connecticut teachers in the participating school district to measure student achievement which will then, in turn, measure a teacher's effectiveness. A study conducted in Arizona and Utah did look specifically at using SLOs as part of the teacher evaluation model (Makkonen, et al, 2015). Similarly, this study looked at SLOs and IAGDs developed by teachers in one Connecticut public school districts. Focus questions for this research were:

1. When setting their SLOs, what assessments did Connecticut teachers use to measure student growth during the 2016-2017 school year?
2. What are teachers' perceptions of the impact of SLOs on student achievement and teacher evaluation?
3. What are administrators' perceptions of the impact of SLOs on student achievement and teacher evaluation?
4. Do SLOs developed by teachers and administrators meet the SMART goal criteria set forth by CT SEED?

Philosophical Framework

As a veteran teacher of 20 years, I have experienced a variety of changes in teacher evaluation over this span of time. Recent federal changes in teacher evaluation have prompted states to change their teacher evaluation models. Connecticut implemented SEED as a pilot in 14 districts in September 2012 and was later adopted statewide. While many other states measure

student growth using standardized assessments, Connecticut does not. As a part of this new teacher evaluation model, teachers across the state are asked to develop student learning objectives (SLOs) to measure the effectiveness of their teaching. This study will look at what measures teachers are using to measure student growth when developing their SLOs, teachers' perception of the SEED process, as well as school administrators' perception of the process. Since this study will be looking specifically at the teacher evaluation system in Connecticut, the philosophical assumption taken is the epistemological assumption. According to Creswell (2015), conducting research under the epistemological assumption researchers "conduct studies in the 'field', where the participants live and work." As a teacher and future school leader, this is an issue that has an impact on me and my colleagues on a daily basis as we are very much "in the field."

The interpretive framework for this research is pragmatism. Creswell states (2015) that pragmatism is focused on the outcome of the research and how the outcome can bring about solutions. Creswell also states that pragmatism will require "multiple methods of data collection to best answer the research question, will employ multiple sources of data collection, will focus on the practical implications of the research, and will emphasize the importance of conducting research that best addresses the research problem." This study will collect both qualitative and quantitative data through questionnaires and interviews.

Research Design

This research was a case study of one school district in the state of Connecticut. Creswell (2013, p. 469) defines a case study as an "in-depth exploration of a bounded system based on extensive data collection." More specifically, it will be a collective case study. Stake, 1995, as quoted by Creswell (2015, p. 469) states that a collective case study looks at multiple cases and

are described and compared to provide insight into an issue. Teacher evaluation is a systematic procedure at both the district and state level. The main goal of the revised teacher evaluation model, in the state of Connecticut is to improve teaching, which in turn, will improve learning. The SLO process asks teachers to establish expectations for their students' learning, therefore impacting teaching practices. Finally, student growth and learning are the measures used to evaluate a teacher's effectiveness. The results of this study provided insight to Connecticut's use of SLOs in the teacher evaluation model.

This study also used a mixed methods design in that both qualitative and quantitative data was collected to address the research questions mentioned earlier. The use of both quantitative and qualitative methods provide a better understanding of the research problem and questions than either method alone (Creswell, 2015).

The Case

Prior to conducting this research, all superintendents in the state of Connecticut were contacted via email. The researcher was hoping to find a district that would benefit from the findings of the research. The state of Connecticut classifies school districts using District Reference Groups (DRGs). This classification system groups schools together based on similar socioeconomic status (CSDE, 2006). District Reference Groups range from A-I. Two school districts volunteered to participate in this study; however only one district participated in the study. The participating district was a DRG D, K-12 district, and is located in New London county. According to the District Profile and Performance Report, published by the Connecticut State Department of Education, for the 2015-2016 school year, the district serves approximately 2,500 students in eight schools. Seventy-nine percent of the student population is white, 5% is Asian, 3.4% is African American, and 9.1% is Hispanic or Latino. About 20% of their students

are eligible for free or reduced lunch and students with disabilities make up 14.5% of their student population. The district employs about 200 certified staff members, 11 school level administrators, and four district level administrators. Both teachers and administrators in both districts will take part in this study.

Data Collection Methods

Quantitative and qualitative methods were used to collect data to address the three research questions set forth. The research questions and methods used to collect data are illustrated in Table 1 below.

Table 1.

Research Questions and Methods

Research Questions	**Method**
1-What assessments did Connecticut teachers use as the IAGDs from 2016-2017 when setting their SLOs?	-Closed-ended Survey Questions -Open-ended Survey Questions
2-What are teachers' perceptions of the impact of SLOs on student achievement and teacher evaluation?	- Closed-ended Survey Questions
3-What are administrators' perceptions of the impact of SLOs on student achievement and teacher evaluation?	- Closed-ended Survey Questions
4. Do SLOs developed by teachers and administrators meet the SMART goal criteria set forth by CT SEED?	-Open-ended Survey Questions

The following section will go into further detail explaining how the data was collected from the participation school district using a mixed methods design.

Surveys

Creswell (2015, p. 379) defines survey research design as procedures in quantitative research in which investigators administer a survey to a sample or to the entire population of people to describe the attitudes, opinions, behaviors, or characteristics of a population. This survey looks at teachers' and administrators' attitudes, opinions, behaviors, and characteristics as they relate to teacher evaluation and more specifically, the SLO process. Further, the survey given will be of cross-sectional design; it will be given at one point in time and will compare the attitudes, opinions, behaviors, or characteristics of two groups, teachers and administrators (Creswell, 2015).

Surveys will include both open-end and close-ended questions. Open-ended questions require the participant to supply the answer and allows the participant to create responses based on their own experiences instead of the researcher's experience (Creswell, 2015). For the purpose of this research, the survey will ask teachers to include their SLO(s). Most of the survey is compiled of close-ended questions. These vary from multiple choice questions to statements that require a response on a five-point Likert scale. Close-ended questions allow the researcher to assign a numerical value to the responses and statistically analyze the data (Creswell, 2015).

Data, both quantitative and qualitative, will be collected using a web-based survey. A web-based survey is a survey instrument for collecting data that is available on the computer (Creswell, 2015 p. 386). *Google Forms* will be used to generate and distribute the surveys. The survey will be made available to all teachers and administrators in the two participating K-12,

Connecticut school districts electronically. Survey questions for teachers were previously used in a study conducted in Arizona and Utah looking at the SLO pilot programs in these states (Makkonen, et al, 2015). The questions for the survey administered were derived from a survey developed by the Utah State Office of Education (Makkonen, et al, 2015). Eleven of the 14 questions from this study were incorporated in the current study (see appendix A). These questions were then slightly modified for the administrator's questionnaire (see appendix B). Both surveys included close-ended questions, while the teacher survey also included open-ended questions. The majority of questions required responses a five-point Likert scale. The open-end question provided teachers an opportunity to give examples of their SLO(s). This data was later coded.

Data Analysis Methods

Quantitative and qualitative data was analyzed to look at both teachers' and administrators' views on the SLO process, the instruments teachers used to measure student growth as it relates to their SLOs, and if the SLOs meet the criteria for being SMART goals. Descriptive and inferential statistics will both be used to look at the data. Measures of central tendency were used to describe the quantitative data. Close-ended questions produced numerical data relating to teachers' and administrators' views relating to the SLO process, while open-ended questions, specifically asking for SLOs, were coded to determine what assessments teachers actually used to determine student growth. Combined, the data provided insight to whether the SLO process is a worthwhile component of the CT SEED.

Quantitative Data

Most of the data collected was quantitative data. Some of the data collected was categorical data (ie: subject you teach, years teaching) to gather information describing the group

of participants. Most of the quantitative data that was collected used a five-point Likert scale. Number values were assigned to responses. The data was analyzed using both descriptive and inferential statistics. Descriptive statistics analyzed the data using mean, median, and mode providing insight to responses for each question. Quantitative data was analyzed using *Statistical Package for Social Sciences (SPSS)*.

Qualitative Data

Through the survey, teachers were asked to share their SLO(s). This information was coded inductively. SLOs were analyzed and sorted based on the type of assessments used to measure student growth. Categories were created based on responses. NVivo was used to establish categories and sort the qualitative data collected. SLOs were also assessed using the following rubric to assess whether the each SLO met the criteria for being SMART goals as outlined by CT SEED. The Connecticut State Department of Education, 2014, outlined the criteria as follows:

> Student learning indicators are written in a S.M.A.R.T. goal format, i.e., student learning indicators are SPECIFIC and STRATEGIC (about what is to be learned and by whom), MEASURABLE (identifies the specific measure/assessment and target), ATTAINABLE (target is rigorous but appropriate to improving student learning), RESULTS-ORIENTED (states what results can reasonably be achieved, given available resources), and TIME- BOUND (specifies when the results can be achieved).

See table 2 below.

Table 2.

Rubric for Assessing SMART Goals

Criteria	Yes- Criteria Met	No- Criteria Not Met
Specific and Strategic		
Measurable		
Attainable		
Results Oriented		
Time-bound		

Reliability and Validity

According to Creswell (2015, p. 158) reliability means that scores from an instrument are stable and consistent while validity is the development of sound evidence to demonstrate that the test interpretation matches its proposed use. Though reliability and validity are separate ideas looking at whether measures are reliable, giving consistent results, and valid, measure what they are designed to measure, they often times overlap (Creswell, 2015). Validity and reliability will be determined through the use of triangulation and peer debriefing.

Triangulation

Triangulation involves using multiple data sources in an investigation to produce understanding of the data collected, as a method for corroborating findings, and as a test for validity (Cohen & Crabtree, 2006). Denzin (1978) and Patton (1999) identify four types of triangulation:

1. Methods triangulation looks at the consistency of data collected using different data collection methods (both quantitative and qualitative data).

2. Triangulation of sources looks at data from different sources from either two different points in time, data from differing groups, or comparing differing viewpoints.

3. Analyst triangulation uses multiple analysts to review findings or using multiple observers and analysts to analyze data.

4. Theory/Perspective Triangulation uses "multiple theoretical perspectives to examine and interpret the data."

(Cohen & Crabtree, 2006)

The reliability and validity of this study will be conducted using both methods triangulation and triangulation of sources. Methods triangulation will be utilized to look at teacher responses to question number four which asks teachers to select the an assessment that best describes the instrument they use to measure student growth, and question number six, an open-ended question, that requires teachers to provide their SLO(s). This triangulation will compare responses to similar questions, however, one question will elicit quantitative data while the other elicits quantitative data. Research in this study will also look at both teachers' and administrators' overall perceptions of the SLO process. Triangulation of sources will compare the data collected from these two groups, teachers and administrators. Close-ended items on both surveys are similar and can be easily compared.

Peer Debriefing

Peer debriefing as quoted by Cohen and Crabtree (2006) "is a process of exposing oneself to a disinterested peer in a manner paralleling an analytical sessions and for the purpose of exploring aspects of the inquiry that might otherwise remain only implicit within the inquirer's mind" (Lincoln & Guba, 1985, p. 308). The survey used to gather data in this case study was reviewed by a small group of Connecticut educators who are currently enrolled in an Educational

Leadership program at Sacred Heart University as part of peer debriefing. The two surveys used in this study, one for teachers and one for administrators, were reviewed and discussed.

Subjectivity Statement

I am a 20-year veteran teacher in the state of Connecticut. I am currently certified to teach special education in grades pre-K to 12 (065) and also hold an (092) Intermediate Administrative Certificate. Further, I am in the process of completing a Certificate of Advanced Study (CAS) in Administration. I have a bachelor degree in psychology and a master's degree in special education. I am also a complementary observer within my school district and am a mentor for new teachers.

Throughout my 20 years of teaching, I have experienced a variety of changes to our educational system with the most recent being Connecticut's implementation of SEED. Changes to teacher evaluation have been made with the hope that there would be improvements made to both teaching and learning. While other states have defined how teacher effectiveness is measured, Connecticut has not. Further, in April 2017, Connecticut's State Department of Education declared that the Smarter Balanced Assessment, the state's assessment of students in grades 3-8, will not be used to measure teacher effectiveness in the teacher evaluation model. Over the past few years, I have used curriculum-based assessments and standardized assessments when developing my SLOs. I have also seen may teachers create their own assessments and evaluate student growth in many different ways.

Summary

This study surveyed teachers and administrators and also collected and coded SLOs from teachers in the participating Connecticut school district. The research provided insight to teacher's perceptions relating to the SLO process and sampled SLOs used by Connecticut

teachers. Participants in the study included both teachers and administrators certified in the state of Connecticut across grade levels and content areas. Surveys including both open and close-ended questions were used to collect data for this study. Triangulation and peer debriefing were used to ensure the reliability and validity of the study.

Chapter 4: Results and Discussion

Introduction (The Study and the Researcher)

The researcher for this case study is a graduate student at Sacred Heart University, in Fairfield, Connecticut, enrolled in the Educational Leadership Program. In addition, the researcher is a veteran teacher of 20 years in the state of Connecticut and is pursuing a Certificate of Advance Study in Educational Leadership. The researcher has a BA in Psychology from Quinnipiac University and a MS in Special Education from Southern Connecticut State University.

States across the country restructured their teacher evaluation models to align with policy changes to ensure that their schools would receive federal grant funding (Donaldson, Papay, 2015). Based on the Measure of Effective Teaching Project (MET), funded by the Gates foundation, states developed teacher evaluation models that evaluated teachers using multiple measures that include student achievement, classroom observations, and student feedback (Gates Foundation, 2009). Various measures are being used to measure student achievement and in turn to evaluate teachers. Value added measures (VAM), state tests, standardized assessments, and student learning objectives (SLOs) are examples of ways states are evaluating teacher effectiveness. When looking specifically at SLOs, there are also variations among states, districts, schools, and teachers in how SLOs are developed and the types of measures used to evaluate student progress. Student learning objective measures may include, but are not limited to, standardized assessments, state tests, district developed assessments, classroom assessments, portfolios, and task completion (Makkonen, et. al., 2015).

In 2012, 14 school districts in Connecticut piloted a new teacher evaluation plan known as System for Educator Evaluation and Development (SEED) (Donaldson, et al, 2014).

Connecticut's SEED includes multiple measures to assess teachers. The components of SEED evaluate teachers using modules that include teacher performance and practice, parent feedback, student growth and development, and whole school student learning/student feedback (CT SEED, 2015). Each of these components is weighted. The weights are student growth and development 45%, teacher performance and practice 40%, parent feedback 10%, and whole school student learning/student feedback 5% (CT SEED, 2015). The weights are used annually to calculate an overall performance rating of 1-4, one being Below Standard, two is Developing, three is Proficient, and four is Exemplary. Districts throughout the state have implemented new teacher evaluation plans that include student learning objectives. However, unlike other states, Connecticut has not mandated that statewide assessments, like the SBA, be used for developing student learning objectives. The question then becomes, what measures are Connecticut teachers using to develop SLOs to measure student growth and development? The purpose of this case study was to gain a better understanding of the SLO development process as it relates to teacher evaluation in the State of Connecticut. The following research questions were the focus of this study:

1. When setting their SLOs, what assessments did Connecticut teachers use to measure student growth during the 2016-2017 school year?
2. What are teachers' perceptions of the impact of SLOs on student achievement and teacher evaluation?
3. What are administrators' perceptions of the impact of SLOs on student achievement and teacher evaluation?
4. Do SLOs developed by teachers and administrators meet the SMART goal criteria set forth by CT SEED?

A previous study, conducted in Arizona and Utah, examined the development of SLOs as part of the teacher evaluation model (Makkonen, et al, 2015). Researchers who conducted this study surveyed teachers who participated in SLO pilots. The survey used in the study was developed by Utah's Department of Education. This case study replicated portions of the methodology used in the study conducted in 2015 in Utah and Arizona to examine SLO development in Connecticut. Specifically, this study examined SLO development and the IAGDs teachers in Connecticut schools used to measure student growth. The closed-ended questions used in this survey were adapted from the aforementioned study. In addition, SLOs were collected and coded to determine what assessments were used to assess student growth, and if the SLOs created by teachers and their administrators met the state's SMART goal criteria as outlined in SEED.

A case study research methodology was used to examine one school district in the State of Connecticut. Creswell (2013, p. 469) defines a case study as an "in-depth exploration of a bounded system based on extensive data collection." Teacher evaluation is a systematic procedure at both the district and state level. The stated purpose of the revised teacher evaluation model in the state of Connecticut is to improve teaching, which in turn, will improve learning. The SLO process asks teachers to establish expectations for their students' learning. Additionally, student growth and learning are also the measures used to evaluate teacher effectiveness. The results of this study will provide insight to Connecticut's use of SLOs.

This case study examined the four research questions outlined above and used both qualitative and quantitative data. The use of a mixed methods approach provided a better understanding of the research problem and questions than either method alone (Creswell, 2015).

Description of Sample

To acquire participants for this study, all superintendents, in the state of Connecticut, received an email from the researcher asking if this study would benefit their district's teacher evaluation model. Two school districts expressed interest in the study; however, only one district chose to participate. A Google Form survey was developed that included both open-ended and close-ended questions. The links for both the teacher and administrator surveys were sent to the district's administrator who then forwarded the surveys to teachers and administrators. Fifty-seven teachers, 22% of the district's teaching staff, completed the survey, while only two administrators, 18% of building level administrators, completed the survey.

Teachers who participated in this case study taught at the elementary, middle, and high school level. Table 3 describes illustrates the sample further.

Table 3.

Grade Level Taught by Case Study Participants

Grade Level	N
Elementary (K-5)	29
Middle (6-8)	8
High School (9-12)	20
Total _	57

To further understand the population that participated in this study, please see the table 4 for the content areas taught by participants.

Table 4.

Content Areas Taught by Case Study Participants

Content	N
Art	3
Counseling	2
ELA	3
Elementary Education	18
Math	8
Music	3
PE/Health	2
Science	4
Social Studies	2
Special Education	9
Technology	1
World Language	1
Other	1
Total	57

Lastly, 78% of the teachers who participated in the study reported that they have been teaching for 10 or more years, 15% have been teaching for 5-10 years, while 7% have been teaching for five years or less (Table 5).

Table 5.

Years Taught by Case Study Participants

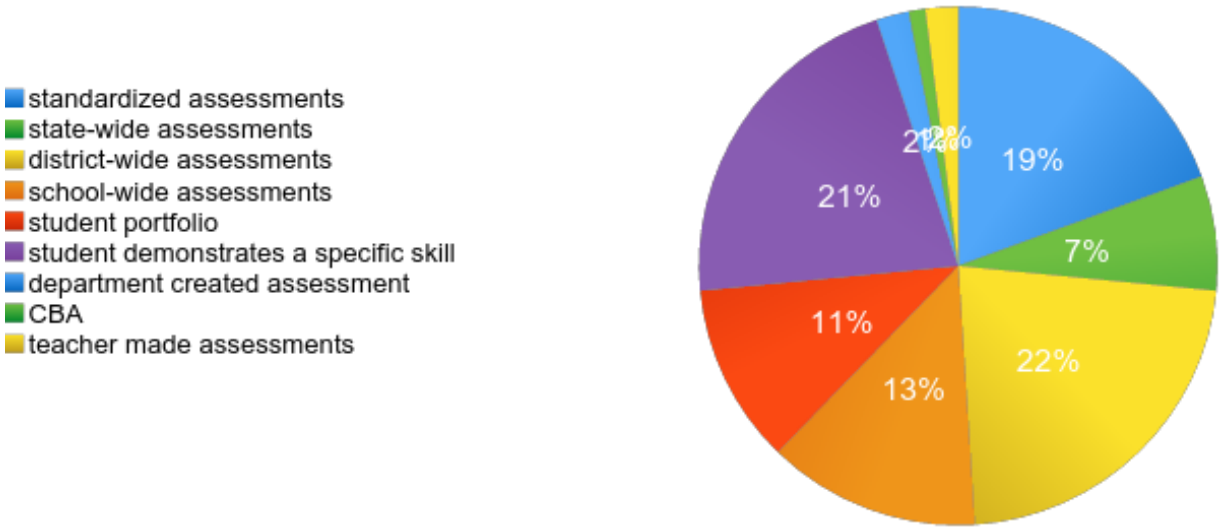
Content	N
Less Than 5 Years	4
5-10 Years	8
More Than 10 Years	43
Did Not Respond	2
Total	57

Results and Discussion

Types of Assessments

Teachers in the sample were provided a list of possible assessments and were asked to select the assessments that best described the IAGDs they used to measure student growth. Table 6 reports the types of IAGD's teachers identified.

Table 6.



Examples of standardized assessments are the STAR and NWEA assessments. These are both computer-based, adaptive assessments that are used as tools to monitor students' progress and growth. Connecticut statewide assessments may include the following: Connecticut Mastery Test (CMT) given in grades 5 and 8, and the Connecticut Academic Performance Test (CAPT) given in grade 10, in the area of science, the Smarter Balanced Assessment is given in grades 3-8 in ELA and math, and the SAT given in grade 11. Districtwide assessments may include writing assessments or content based assessments developed at the district level in areas not tested by standardized assessments. School-wide assessments may include midterm and final exams at the

high school level. CBAs are curriculum based assessments and the other categories are self-explanatory.

The responses were examined to determine if the content area influenced the type of IAGD identified. Content areas were grouped together to examine this data. ELA, math, and science were grouped together as all three content areas are assessed by both state tests and standardized assessments.

Table 7.

Assessments used by ELA, math, and science teachers

Assessments	N
Standardized Assessments	6
Statewide Assessments	1
Districtwide Assessments	4
School-Wide Assessments	3
Student Portfolio	1
Student Demonstrates a Specific Skill	3
Other	3
Total	21

Art, music, PE/health, technology, social studies, counseling, and world language teachers were grouped together as standardized measures are not available in these areas.

Table 8.

Assessments used by art, music, PE/health, technology, social studies, counseling, and world language teachers

Assessments	N
Standardized Assessments	1
Statewide Assessments	2
Districtwide Assessments	0
School-wide Assessments	1
Student Portfolio	3
Student Demonstrates a Specific Skill	8
Other	6
Total	21

For this question, there was also an "other" option. Teachers who selected this option were asked to specify what IAGDs they used. Assessments in the "other" category included department-created assessments as indicated by world language teachers and teacher-made assessments as indicated by music teachers.

Within the survey, teachers were also asked to include their SLOs and IAGDs. Thirty-nine SLO samples were collected and coded inductively using NVivo. Some participants included more than one measure or more than one SLO; a total of 43 SLOs and assessments were coded inductively. Twenty-four of the 43, or 56% of SLOs did not specify the type of assessment used. Examples of these SLOs include, "fitness literacy test", "vocabulary knowledge assessment", and "Students will advance at least 2 TCRR reading levels." Several of these SLOs did not provide enough detail for the researcher to deduce the type of assessment used. Eight of the sample SLOs suggested standardized assessments were used to measure student growth. Based on responses to the open-ended question, it was determined that one of the standardized assessment used in this district was the Northwest Evaluation Association assessment (NWEA). NWEA assesses math, language arts, and science, and is available for students in grades pre-K-12. Several of the SLOs in this category included "MAP," which is directly related to the NWEA and refers to Measures of Academic Progress. Additionally, five SLOs indicated that they require students to demonstrate a specific skill, two suggested school-wide assessments like end of the year exams, one SLO measured student growth using the Smarter Balanced Assessment, one specified using a "teacher developed rubric." Two additional SLOs included other measures like the "Seesaw app" and a "future plan."

The SLO's were examined using NVivo. Seven codes emerged from the data. These seven codes represent the type of IAGD's teacher identify. These codes were standardized

assessments, statewide assessments, school-wide assessments, student demonstrates a specific skill, teacher-made assessments, unknown, and other. Table 9 reports the frequency of the codes.

Table 9.

Assessments used as reported and coded using SLOs provided by teachers

Assessments	N
Standardized Assessments	8
State-Wide Assessments	1
School-Wide Assessments	2
Student Demonstrates a Specific Skill	5
Teacher-Made Assessments	1
Unknown	24
Other	2
Total	43

When looking at the data collected to address research question one, it is evident that within one school district, several assessments are being used to measure student growth. Both qualitative and quantitative data suggest that standardized assessments are most widely used by teachers in this district. In contrast, very few teachers rely on statewide assessments. The content areas taught by teachers seem to impact the assessments used. Teachers, whose content areas are not measured by standardized assessments, are more likely to have students demonstrate a specific skills or these teachers develop their own assessment to measure student growth. At the same time, math, science, and ELA teachers are more likely to use standardized assessments and/or district-wide assessments.

Very few teachers incorporated state tests, like the SBA or CMT, to measure student growth. Only 12% of the teachers surveyed indicated that they used state-wide assessments when developing their SLOs. Unlike other states, Connecticut has not mandated state-wide assessments be used in teacher evaluation (CTSDE, 2015).

Teachers' Perception of the Efficacy of SLOs in Evaluating Teachers

The second research question focused on teachers' perceptions of the impact SLOs have had on teaching, learning, and teacher evaluation. Eleven close-ended questions addressed this research question by requiring participants to record their answers using a five point, Likert Scale with one being *strongly disagree* and five being *strongly agree*. This data was analyzed using SPSS. The mean was reported for each question.

Approximately one-third of responses were neutral (3) on most questions. This is also reflected in the mean for each question. One question differs from this pattern; the last question asks, "Rather than SLOs, I would prefer to be held accountable for school-wide math and/or reading test scores." Approximately 49% of participants strongly disagreed with this statement.

Table 10.

Teachers' perceptions of the impact of SLOs on student achievement and teacher evaluation

Questions	Mean
In theory, I believe the SLO process provides an accurate assessment of my teaching effectiveness.	2.75
Through the SLO process I gained new knowledge about effective ways to assess students.	2.77
I am confident that the assessment(s) used as part of my SLO(s) appropriately measured my students' growth.	3.32
My mid-year check-in with my administrator to discuss my SLO(s) was helpful.	3.37
The SLO process has improved the quality of my conversations with my administrator about instruction.	3.07
The SLO process has improved the quality of my conversations with my administrator about assessment.	3.14
The SLO process helped inform my professional growth.	2.95
I have changed instruction in my classroom as a result of the Implementation of SLOs.	3.32
My participation in the SLO process has benefited my students.	3.14
Overall, the SLO process is worthwhile.	2.69
Rather than SLOs I would prefer to be held accountable for school-wide math and/or reading test scores.	2.00

The data suggests that teachers do not feel the SLO process is effective with respect to measuring their performance by using quantitative metrics to document students' growth. Teachers do not feel the process provides an accurate assessment of their teaching. The frequency of responses to this item was as follows: 6 strongly disagreed, 18 disagreed, 18 were neutral, 14 agreed, and 1 strongly agreed. Teachers do not feel the SLO process has helped them gain new knowledge relating to student assessment. Teachers were divided when asked if the SLO process led to

professional growth; 6 teachers strongly disagreed, 15 disagreed, 14 were neutral, 20 agreed, and 2 strongly agreed. Teachers do feel the SLO process has led them to change instruction (4 strongly disagreed, 7 disagreed, 20 were neutral, 22 agreed, and 3 strongly agreed) and that the SLO process has benefitted students (3 strongly disagreed, 12 disagreed, 19 were neutral, 18 agreed, and 4 strongly agreed). Additionally, teachers feel the SLO process has improved the quality of conversations with their administrators, especially the mid-year check-in. Finally, teachers feel the assessments they are using as part of their SLOs appropriately measure student growth. That being said, it is important to restate that approximately one-third of teachers neither agreed or disagreed with the statements set forth by the survey suggesting they do not have strong feelings for or against the SLO process.

Administrators' Perception of the Efficacy of SLOs in Evaluating Teachers

The third research question for this case study focused on the administrator's perspective on the SLO process and its impact on teaching, learning, and teacher evaluation. The N for this portion of the study was two of a possible 11 building-based administrators. Therefore, a statistical analysis of the results is impractical. However, the qualitative data from the open-ended responses offers additional insight into the SLO process. When looking at the individual responses, the participating administrators appeared to have differing views of the SLO process, one was slightly in favor of the process, while the other was not. One participant indicated they strongly disagreed that the SLO process was worthwhile. Neither participant felt the process has benefited students or has changed teaching and learning in their schools. One participant included the following comment in their survey,

> SLO development can become more of a game rather than a true plan for student growth and achievement. Teachers tend to set goals that are achievable rather than risk what might look like failure and then that reflects on them.

When comparing this statement to teacher responses, this may be an accurate description of what is happening. When teachers were asked if they felt they would meet or exceed their goals for the current school year, 97% answered "yes." It is possible that teachers are selecting assessments and developing SLOs that they know are attainable. Perhaps teachers are not taking risks.

SMART Criteria

The final research question for this case study examined whether SLOs met the SMART goal criteria set forth by CT SEED. The Connecticut State Department of Education, 2014, outlined the criteria as follows:

> Student learning indicators are written in a S.M.A.R.T. goal format, i.e., student learning indicators are SPECIFIC and STRATEGIC (about what is to be learned and by whom), MEASURABLE (identifies the specific measure/assessment and target), ATTAINABLE (target is rigorous but appropriate to improving student learning), RESULTS-ORIENTED (states what results can reasonably be achieved, given available resources), and TIME- BOUND (specifies when the results can be achieved).

Thirty-nine SLOs were gathered and assessed using a rubric highlighting the criteria outlined above (see Table 2). The researcher rated the SLOs using this rubric and an additional rater completed the same task for inter-rater reliability. Of the 39 SLOs, 23% met the SMART criteria, 23% of the SLOs did not meet any of the SMART criteria, and 54% of the SLOs met one or more of the SMART criteria, but not all five.

Table 11.

SMART criteria met by SLOs

Criteria	N
SMART	9
S- SPECIFIC and STRATEGIC	18
M- MEASURABLE	19
A- ATTAINABLE	20
R- RESULTS-ORIENTED	22
T- TIME-BOUND	8
No criteria met	9

The following is an example of an SLO that met the all SMART criteria:

> Students in my 2 Honors Spanish classes will improve their speaking skills in the target language, as assessed by departmental speaking rubrics aligned with those of the AP exam. 1. 90% of students will improve their oral responses, from the baseline speaking assessment given in Quarter 1 to the final speaking assessment given in Quarter 4, not based solely on the score band; however, the quality and content of the oral responses will show growth by the criteria set forth in the rubric. 2. The remaining 10% of students will show growth in one or more areas of the rubric on a more complex speaking assessment.

The teacher indicated that the focus group was the students in their Spanish honors classes and that these students would improve their speaking skills in their target language. This teacher indicated rubrics would be used to measure student growth and specified growth expected by students. The goal seems attainable and is results oriented. Finally, the SLO is time-bound in that the teacher is clear in stating growth will be measured using a baseline assessment in quarter 1 and a final assessment in quarter 4. In contrast, the following is an example of an SLO that does not meet any of the SMART criteria set forth by the state: Students will develop and improve

their written response to literary text. The following is an example of an SLO that met some, but not all SMART criteria:

> Students will make strategic use of digital media in presentations to enhance understanding of their findings, reasoning, and evidence and to add interest. Of the 75 students who are below goal, 85% (64 students) will move up at least one band width. Of the 15% (11 students) of students who do not move up, 100% will maintain their level. Of the 9 students who are at goal, 95% (9 students) students will move to advance.

This SLO is Specific in that it specifies which students are expected to make growth. It is not Measurable because it is not clear what assessment tool is being used to measure growth. It is Attainable and Results-oriented, but it is not Time-bound.

Additional Feedback

To conclude the survey, teachers were given the option to include their comments regarding the SLO process. Ten participants included comments. Statements within the comments were coded inductively, using NVivo. Ten of these statements were coded as negative, two were neutral, and one was positive. Teachers expressed concern regarding the accuracy of test results because students may not put forth effort if they know the results will not impact their grades, while others expressed concern about the amount of time dedicated to assessing students.

One teacher noted:

> Basing teacher effectiveness on assessment results of a student who is overly tested is not ideal. In addition, my classes miss 5 instructional days (that's 2 full weeks in block scheduling) due to standardized testing alone. When you consider the 10-12 classroom

> unit tests and other assessments, you can see that the students spend close to an entire month of instruction completing assessment.

Teachers also felt the process is not equitable for all teachers. One teacher commented:

> The pressure on math and language arts teachers is much greater than the pressure placed on teachers that specialize in other areas. It doesn't seem equitable as standard units of measure are not available for other content areas. It would be more equitable if all teachers were accountable using the same assessments and had to focus on some aspect of math or literacy in their classes. For example, social studies classes could take over non-fiction reading, science could focus on measurement. PE, Health and Art could reinforce and measure student growth on rates, ratios, proportions and percent.

While another teacher noted, "I would prefer a purely art based goal not goals based on literacy." Neutral comments included, "We have math and literacy coaches in each elementary school building. They have been very helpful in suggesting and supporting shifts in instruction." and, "I feel writing SLO's and assessing with them is only a stepping off point for improving teaching to meet the needs of my students." The one positive comment stated, "The SLO process does focus instruction."

Summary

In this chapter, the results for this case study were presented as they related to each research question. The data suggests that teachers do need feel the SLO process is worthwhile. At the same time, teachers do not want to be held accountable for school-wide math and/or reading scores. Results imply about one-third of teachers do not have strong feelings for or against the SLO process. Further, teachers reported that there are positive aspects to the SLO

process, for example, the process accurately assess student growth, it has changed instruction within classrooms, and teachers perceive the process to have benefited student learning.

When analyzing IAGDs used in SLOs it became apparent that standardized assessments, like the NWEA, are widely used within the participating district based on closed-ended responses. Teachers also often require students to demonstrate a specific skill. When analyzing the SLOs provided in response to an open-ended question, the 24 of 43 SLOs could not be coded because not enough detail was provided for the research to deduce the type of assessment used. Further, most SLOs did not meet the state criteria for being SMART goals.

Chapter 5: Conclusions and Implications

Summary

In recent years, teacher evaluation has been a "hot topic" in education. In Connecticut and in other states throughout the country, Departments of Education have been restructuring their teacher evaluation models to meet the demand of federal mandates. With the adoption of the Common Core State Standards and changes to teacher evaluation, teachers and administrators alike, have been changing practice to implement these new initiatives. As part of my internship for my Educational Leadership program, I was asked to look at several teacher evaluation models to provide my district with data to help inform possible changes in teacher evaluation in our district. Having looked at more than ten models from various districts, I was somewhat surprised at the variation in plans from district to district. Connecticut, unlike other states, did not mandate state testing results be tied to teachers' evaluations, I was curious to learn more about what measures teachers were using to develop their SLOs.

As an educator, I have seen an increase in the number of assessments we require of students. In addition to statewide tests, many districts have adopted district-wide assessments like the STAR Assessment or the NWEA. These tests are given at least three times a year and perhaps more frequently if progress monitoring is required. There are also curriculum based assessments and exams for high school students. Further, I have seen many teachers administer additional assessments for their SLOs.

Chapter One discussed the recent changes in teacher evaluation and the changes in federal policies that led to these changes. Federal policies, like No Child Left Behind and Race to the Top, linked grant funding to education reform. As a result, states made changes to their teacher evaluation plans to meet the demands of the newly instated policies. Teacher evaluation

models were greatly influenced by the Measure of Effective Teaching Project (MET). Based on the MET Project, funded by the Gates foundation, states developed teacher evaluation models that evaluated teachers using multiple measures that include student achievement, classroom observations, and student feedback (Gates Foundation, 2009). Connecticut began to make changes to its teacher evaluation model and adopted System for Educator Evaluation and Development (SEED). This plan required 45% of a teacher's overall rating to be based student growth and development which was determined through the development of Student Learning Objectives. The overall purpose of this study was to gain a better understanding of the SLO process, to understand teacher's and administrator's perceptions of the process, to sample SLOs and determine if they meet state criteria, and to examine if the process has had a positive impact on teaching and learning. Chapter One also included the four research questions, definitions of terms, and expected findings.

Chapter Two consisted of a review of literature exploring teacher evaluation with a focus on how teacher effectiveness is measured. The literature review provided an explanation of why teacher evaluation has changed and looked further at the measures currently used to determine a teacher's effectiveness in the classroom. Value added measures (VAM), state tests, standardized assessments, and student learning objectives (SLOs) are examples of ways states are evaluating a teacher's effectiveness. The literature review examined these measures as well as their benefits and weakness. Due to rapid changes in this field, not much research has been done looking specifically at SLOs. To conclude, the literature identified gaps in the research and made recommendations for future research.

Chapter Three restates the purpose of the study and research questions. It also outlines the philosophical framework and the researcher bias. Chapter Three outlines the sample, research

design, and data collection methods. This case study was conducted in a Connecticut school district; both qualitative and quantitative data was collected using an electronic survey. Chapter Three provides an in depth look at how data was collected and analyzed for this case study.

Chapter Four provided research findings as they related to each of the four research questions. Descriptive statistics were used to analyze quantitative data, while qualitative data was coded inductively to better understand the SLO process. Findings suggest that teachers do not have strong feeling for or against the SLO process, however there are components of the process they feel are beneficial and others that are not. Looking at sample SLOs, it is also evident that the majority of SLOs do not meet the state criteria for SMART goals. Chapter Four detailed the sample of participants and reported findings that were also illustrated using tables and pie charts. In conclusion, Chapter Four addressed all four research questions in detail.

In Chapter Five I will discuss the limitations of this study by addressing faults in the study and examining difficulties that arose when conducting the actual research. Chapter Five will also look at implications of this study and how the data could possibly influence current practice relating to SLOs. I will also suggest future research that may provide further data to help fill gaps in the literature and help shape teacher evaluation in the state of Connecticut.

Limitations of the Study

When looking at the limitations of this study, three issues became apparent; the sample size, the number of administrators that participated in the study, and the sensitivity of the topic addressed. All participants in this study were employed by one school district. According to the Connecticut State Department of Education, there are 164 school districts in Connecticut. Connecticut schools are divided into DRGs, as discussed in Chapter 3. The participating district is a DRG D district, DRGs in Connecticut range from A-I. Surveying one out of 164 districts is a

limitation in this study. Creswell (2012) suggested that a small sample size could limit the generalizability of the results of a study. For these reasons, one must be cautious when generalizing results from this study to other districts or the state as a whole.

Another limitation of the study was a lack of participation by school administrators. According to the District Profile and Performance Report for the 2015-2016 school year, the participating district employs 11 building based administrators. Of these 11 administrators, only two, or 18%, chose to participate in this case study. Due to the small sample of administrators, a statistical analysis was not completed as planned.

Teacher evaluation is also a sensitive topic. Therefore, participant may have been reluctant to answer questions openly and honestly. The survey focused on their evaluation process. The survey was also distributed, via email, by district administrators and this may have also influenced participant responses.

Implications for Practice

Looking at both the quantitative and qualitative data in this case study, there are several implications for practice. To begin, we will look at quantitative data. Teachers were asked whether they *strongly agreed* or *strongly disagreed* to 11 statements relating to the SLO process using a five-point Likert Scale. Based on this data, approximately one-third of the participants had neutral feelings toward the SLO process. However, when asked, "Rather than SLOs, I would prefer to be held accountable for school-wide math and/or reading test scores." 67% of participants *disagreed* or *strongly disagreed* with this statement. This finding aligns with results from a study previously conducted in Utah and Arizona where 66% of the teachers surveyed did not want to be held accountable for school-wide test scores in place of SLOs (Makkonen, et al, 2015). The data suggests while teachers may not be in favor of the SLO process, they seem to

feel it is a better alternative than being held accountable for school-wide reading and math scores. The Connecticut State Department of Education (2007) published a document titled *Beyond the Blueprint: Literacy in Grades 4-12 and Across the Content Areas* that suggests literacy should be a focus throughout a child's education and should be taught in all content areas. *Beyond the Blueprint* also provides sample lessons relating to comprehension, vocabulary, spelling, and fluency across curricular areas. Literacy should be a focus of professional development across content areas and grade levels. With such a focus on literacy at the state level, requiring teachers to develop SLOs based on standardized reading scores would not be unreasonable.

Quantitative data also suggests that most teachers did not feel the SLO process has been worthwhile. Teachers feel the SLO process has not accurately assessed their teaching, but more than 46% of teachers reported the assessments used to develop their SLOs and measure student growth were accurate. Additionally, 96.5% of participants indicated that they would meet or exceed their SLOs. Based on the SEED model, many of these teachers would be rated as proficient or exemplary as a result because SLOs account for 45% of a teacher's overall rating (CTSDE, 2014). If students are reaching the academic goals set by teachers and if teachers are meeting or exceeding their goals, one might conclude the teacher's instruction has been effective, therefor accurately assessing their teaching. Forty-five percent of teachers noted that the SLO process has changed their teaching and 39% of teachers feel the process has benefitted students. Though overall feelings suggest the SLO process has not been worthwhile, there are positive outcomes related to the implementation of SLOs. The overall purpose of the SLO process has been to improve teaching and learning and this data suggests the process has had a positive impact on teaching and learning.

The qualitative data asked teachers to provide sample SLOs which were coded to determine the types of assessments used in developing SLOs and they were also rated to determine if the SLOs met the state criteria for being SMART goals. In many cases, the type of assessment used could not be determined looking solely at the SLO. This links to the fact that only 23% of the SLOs met the SMART criteria. Perhaps some teachers did not take the time to convey their actual SLO as some reported their SLO was to simply, "Increase student comprehension." At the same time, other SLOs met some of the SMART criteria, most often neglecting the "T" and not making their SLOs time-bound. Since SLOs are developed and agreed upon by both teachers and administrators, perhaps professional development is needed to assist both teachers and administrators establish SLOs that meet the SMART criteria set forth by the state.

Summary

The data suggests there are mixed opinions when looking at the SLO process. Though some teachers reported they do not feel the SLO process is worthwhile, teachers are reporting the process benefits students. One can infer that teachers feel the SLO process has had a positive impact on teaching and learning, but it is not a meaningful component of teacher evaluation. While standardized measures are not available in all content areas, literacy is a skill that all teachers should be addressing within their classes. Potentially, all teachers can have an SLO relating to standardized reading scores. The second SLO, as required by SEED, can be content based and IAGDs can vary to best assess student skills. Finally, more focus should be placed on the development of SLOs to make certain they are meeting the SMART criteria also outlined in SEED.

Suggestions for Future Research

Student learning objectives (SLOs) are becoming a popular alternative measure because they can be used to evaluate teachers in any grade or subject (Gill, et al, 2013). The SLO process, as defined by Race To the Top Technical Assistance (2010 p. 1), is "a participatory method of setting measurable goals, or objectives, based on the specific assignment or class, such as students taught, the subject matter taught, the baseline performance of the students, and the measurable gain in student performance during the course of instruction." A case study by Lacireno-Paquet, Morgan, and Mello (2014) reviewed state websites and took a close look at the use of SLOs in teacher evaluation plans. Thirty states reported using SLOs in their teacher evaluation models.

Teacher evaluation has been a major focus of our educational system and has faced significant changes. States have been revamping their evaluation systems to meet the new federal requirements. Though student achievement must be a part of a state's teacher evaluation program the federal guidelines do not spell out how student achievement must be measured. Therefore, student achievement has taken many different forms. Research looking specifically at SLOs is limited at this time.

Future research, in the state of Connecticut, should be grounded in the SLO process. One area of focus should be whether SLOs across the state are meeting the SMART criteria as outlined in SEED. In the spring of 2017, the Connecticut State Department of Education announced that the statewide assessment, the SBA, would not be used to evaluate a teacher's effectiveness, but SEED requires at least one SLO be based on standardized assessments. It would be beneficial to learn what measures are being used throughout the state and the impact these assessments are having on teaching and learning across content areas and grades.

Looking beyond the state level, it would be interesting to learn more about how the SLO process is being implemented in other states across the country. I would be curious to learn about specific criteria established by other states for SLO development. Additionally, what assessments do other states require to measure student growth.

As a final note, future research may also want to look at the number of assessments students are subjected to and the amount of time allotted for assessing students. Though assessing students is crucial not only to evaluate a teacher's effectiveness, but also to assess student progress, it seems we are assessing students more than ever. We have statewide tests and standardized assessments, in addition to curriculum based assessments. Data is needed, but is all the data collected meaningful in that it drives instruction? States and districts should be more strategic when determining what assessments are administered to students and assessments should be instructionally sensitive.

References

Baker, B. D., Oluwole, J. O., & Green, P. C. (2013). The legal consequences of mandating high stakes decisions on low quality information: Teacher evaluation in the race-to-the-top era. *EPAA/AAPE*, *21*(5), 1-71. Retrieved from ERIC database.

Connecticut State Department of Education, 2015 SEED Handbook Connecticut's System for Educator Evaluation and Development, 2015, at 1-83 (Conn. 2015). Retrieved from http://www.connecticutseed.org/wp-content/uploads/2015/11/2015_SEED_Handbook_11_24_15.pdf

Connecticut State Department of Education, Connecticut Summative Assessments 2017 - 2018, , at 1 (Conn. 2017). Retrieved from http://www.sde.ct.gov/sde/cwp/view.asp?a=2748&Q=334726

Connecticut State Department of Education Division of Teaching, Learning and Assessment, Research Bulletin: District Reference Groups, 2006, at 8 (Conn. 2006). Retrieved from http://www.sde.ct.gov/sde/lib/sde/PDF/EvalResearch/DRG_2006.pdf

Cosner, S., Kimball, S. M., Barkowski, E., Carl, B., & Jones, C. (2015). Principal roles, work demands, and supports needed to implement new teacher evaluation. *Mid-Western Educational Researcher*, *27*(1), 76-95. Retrieved from ERIC database. (Accession No. EJ1051625)

Department of Education National Comprehensive Center for Teacher Quality. (2011, March). *Measuring teachers' contributions to student learning growth for nontested grades and subjects. Research & policy brief* (L. Geo & L. Holdheide, Authors). Retrieved from ERIC database. (ED520722)

Donaldson, M. (2010). No more valentines. *Educational Leadership*, *67*(8), 54-58. Retrieved from http://www.ascd.org/publications/educational-leadership/may10/vol67/num08/abstract.aspx

Donaldson, M. L., & Papay, J. P. (2015). An idea whose time had come: Negotiating teacher evaluation reform in New Haven, Connecticut. *American Journal of Education, 122*, 39-70.

Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. (2014). Choosing the right growth measure. *Education Next*, *14*(2), 66-71. Retrieved from ERIC database. (Accession No. EJ1032952)

Forman, K., Ph.D., & Markson, C., Ed.D. (2015). Is "effective" the new "ineffective"? A crisis with the New York state teacher evaluation system. *Journal for Leadership and Instruction*, *Fall*, 5-11. Retrieved from ERIC database.

Gabriel, R., & Allington, R. (2012). The MET Project: The wrong 45-million-dollar question. *Educational Leadership*, *70*(3), 44-49. Retrieved from ERIC database.

Jiang, J. Y., Sartain, L., Sporte, S. E., & Steinberg, M. P. (2014). The impact of teacher evaluation reform on student learning: Success and challenges in replicating experimental findings with non-experimental data. *SREE*, (Spring). Retrieved from ERIC database.

Kane, T. (2012). Capturing the dimensions of effective teaching. *Education Next*, *Fall*, 34-41. Retrieved from ERIC database.

Lacireno-Paquet, N. (n.d.). How states use student learning objectives in teacher evaluation systems: a review of state websites. *Washington DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands*. Retrieved from ERIC database.

Makkonen, R., Tejwani, J., & Rodriguez, F. (n.d.). A descriptive study of the pilot implementation of student learning objectives in Arizona and Utah. *Washington DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West*. Retrieved from ERIC database.

Marshall, K. (2012). Fine-tuning teacher evaluation. *Educational Leadership*, *70*(3), 50-53. Retrieved from ERIC database.

Megan, K. (2016, February 2). State releases some evaluation ratings for teachers. *Hartford Courant*. Retrieved from http://www.courant.com/education/hc-teacher-evaluation-data-0202-20160202-story.html

Sarah, W. L., & Donaldson, M. L. (2015). District leaders' framing of educator evaluation policy. *Educational Administration Quarterly*, *52*(1), 110-143. http://dx.doi.org/10.1177/0013161x15616661

Warring, D. F. (2015). Teacher evaluations: Use or misuse? *Universal Journal of Educational Research*, *3*(10), 703-709. http://dx.doi.org/10.13189/ujer.2015.031007

Woulfin, S. L., & Donaldson, M. L. (2016). District leaders' framing of educator evaluation policy. *Educational Administration Quarterly*, *52*(1), 110-142. Retrieved from ERIC database. (Accession No. EJ1085966)

Appendix A

Survey Distributed to Teachers

(1) Please select the grade level you teach.

a. Elementary K-2

b. Elementary 3-5

c. Middle School 6-8

d. High School

(2) Please indicate the content area that best describes what you teach.

a. Art

b. Counseling

c. ELA

d. Elementary Education

e. Math

f. Music

g. PE/Health

h. Science

i. Social Studies

j. Special Education

k. Technology

l. World Language

m. Other

(3) How many years have you been teaching?

a. Less than 5 years

b. 5-10 years

c. More than 10 years

(4) What best describes the assessments you use to measure student achievement as related to your SLO(s)?

a. Standardized Assessments (STAR, NWEA, etc.)

b. Statewide Assessments (SBA, CMT, CAPT, SAT)

c. Districtwide Assessments

d. School-wide Assessments

e. Student Portfolio

f. Student Demonstrates a Specific Skill

g. Other

(5) Do you believe you will meet or exceed your SLO(s) for this school year?

a. Yes

b. No

(6) Please include your SLO (and IAGD if applicable) below.

(7) In theory, I believe the SLO process provides an accurate assessment of my teaching effectiveness.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Strongly Agree

(8) Through the SLO process I gained new knowledge about effective ways to assess students.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Strongly Agree

(9) I am confident that the assessment(s) used as part of my SLO(s) appropriately measured my students' growth.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Strongly Agree

(10) My mid-year check-in with my administrator to discuss my SLO(s) was helpful.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Strongly Agree

(11) The SLO process has improved the quality of my conversations with my administrator about instruction.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Strongly Agree

(12) The SLO process has improved the quality of my conversations with my administrator about assessment.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Strongly Agree

(13) The SLO process helped inform my professional growth.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Strongly Agree

(14) I have changed instruction in my classroom as a result of the implementation of SLOs.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Strongly Agree

(15) My participation in the SLO process has benefited my students.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Strongly Agree

(16) Overall, the SLO process is worthwhile.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Strongly Agree

(17) Rather than SLOs I would prefer to be held accountable for school-wide math and/or reading test scores.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Strongly Agree

Appendix B

Survey Distributed to Administrators

(1) Please select the grade level you supervise.

   a. Elementary K-2
   b. Elementary 3-5
   c. Middle School 6-8
   d. High School

(2) What best describes the assessments you use to measure student achievement as related to your SLO(s)?

   a. Standardized Assessments (STAR, NWEA, etc.)
   b. Statewide Assessments (SBA, CMT, CAPT, SAT)
   c. Districtwide Assessments
   d. School-wide Assessments
   e. Student Portfolio
   f. Student Demonstrates a Specific Skill
   g. Other

(3) In theory, I believe the SLO process provides an accurate assessment of the teaching effectiveness of the staff members I supervise.

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

(4) Through the SLO process, teachers have gained new knowledge about effective ways to assess students.

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

(5) I am confident that the assessment(s) used by staff members as part of their SLO(s) appropriately measured students' growth.

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

(6) Mid-year check-ins with teachers to discuss SLO(s) are beneficial to teaching and learning.

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

(7) The SLO proves has improved the quality of my conversations with teachers about instruction.

   a. Strongly Disagree
   b. Disagree

c. Neutral

d. Agree

e. Disagree

(8) The SLO process improved the quality of my conversations with teachers about assessment.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Disagree

(9) The SLO process helped inform professional development opportunities for teachers.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Disagree

(10) There has been a change in teaching and learning as a result of the implementation of SLOs.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Disagree

(11) Participation in the SLO process has benefited students.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Disagree

(12) Overall, the SLO process is worthwhile.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Disagree

(13) Rather than SLOs, I would prefer teachers to be held accountable for school-wide math and/or reading test scores.

a. Strongly Disagree

b. Disagree

c. Neutral

d. Agree

e. Disagree

(14) Please feel free to add any additional information or comments.