Fall 2020

A Mixed Methods Study of Student Learning Objectives (SLOS) in Evaluating Teacher Effectiveness

Xumei Fan

A MIXED METHODS STUDY OF STUDENT LEARNING OBJECTIVES (SLOS) IN EVALUATING TEACHER EFFECTIVENESS

by

Xumei Fan

Bachelor of Arts
Shanxi Normal University, 1998

Master of Arts
Shanxi University, 2005

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Educational Psychology and Research

College of Education

University of South Carolina

2020

Accepted by:

Michael Seaman, Major Professor

Christine DiStefano, Committee Member

Melissa Duffy, Committee Member

Lucy Spence, Committee Member

Leigh D'Amico, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

ACKNOWLEDGEMENTS

I would not have completed my doctoral study without the substantial support academically and personally from many wonderful people in my life.

I would like to thank my committee members Dr. Seaman, Dr. DiStefano, Dr. Spence, Dr. D'Amico, and Dr. Duffy for their valuable advice on my dissertation. My advisor and chairman of my dissertation committee Dr. Seaman has been very professional, knowledgeable, and supportive. Without his guidance and encouragement, I would not have completed my dissertation study. I want to thank Dr. DiStefano for her insightful feedback on my dissertation and opportunities to teach EDRM 423. I took my very first course at USC with Dr. Spence, and she has been helping me grow academically and personally. I appreciate the opportunity of working with Dr. D'Amico, and I value her permission for me to use some data and her input in my dissertation study. I am grateful for all the valuable advice from Dr. Duffy who helped improve the quality of my dissertation.

My heartful thanks go to Dr. Johnson who gave me opportunities to teach and work on his research projects, and I learned so much from him. Dr. Thompson has always been very positive and encouraging, and it meant so much to me. I am truly thankful to my work supervisors Drs. Dickenson, D'Amico, and Lewis at the REM Center, and working with them on various projects helps me become a better educator, researcher, evaluator, collaborator, and learner.

I also want to thank my dear friends Tim, Jane, and Qiguang. They welcomed me when I was fresh off the plane, they introduced me to American culture, southern culture, and many other adventures, and they treated me like part of their family.

Last but not the least, my very special thanks go to my mother Jine and my daughter Amy. My mother never had an opportunity to attend school, but she has never had any doubt or hesitation in supporting me for a better education. My daughter Amy has been supporting me unconditionally, and her love and care has meant so much in my journey of this doctoral study.

ABSTRACT

Student Learning Objectives (SLOs), as a measure of students' academic growth, have been broadly used in the evaluation of teacher effectiveness. In this research I employed a mixed-methods sequential explanatory design and incorporated four studies to investigate using SLOs in teacher evaluation in South Carolina (SC). For Study 1 and Study 2, I used surveys to explore educators' perspectives of SLOs before and after the full implementation of the teacher evaluation system in SC. For Study 3, I used interviews with teachers to explore in-depth the impact and implementation of SLOs. In Study 4, I investigated the relationship between teachers' SLO scores and classroom observation scores. An analysis of survey scale questions and open-ended questions, interviews, and evaluation scores from a total of 1,020 participants revealed important findings about using SLOs in the evaluation of teacher effectiveness.

First, most educators agreed on the impact of SLOs before the full implementation of the evaluation system but disagreed on the impact of SLOs after the full implementation of the system. Second, most educators agreed on the impact of classroom observations both before and after the full implementation of the evaluation system. Third, in comparison with teachers, administrators reported more positive views of both SLOs and classroom observations. Fourth, after the full implementation of the evaluation system, early career teachers reported more positive views of both SLOs and classroom observations in comparison with career teachers. In addition, there were various issues reported regarding the implementation of SLOs in teacher evaluation.

Finally, teachers who had positive views of classroom observations tended to have more positive views of SLOs. Teachers who had a higher classroom observation score tended to have a higher SLO score, and the SLO scores could better differentiate teacher performance in comparison with their classroom observation scores.

These findings provide important information about using SLOs in teacher evaluation in South Carolina. They can be used to improve teacher evaluation system, inform policy making, promote teacher professional development, enhance classroom instruction and assessment, and support student learning.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

ADEPT..... Assisting, Developing, and Evaluating Professional Teaching

AIR.....American Institutes for Research

ESEA..... Elementary and Secondary Education Act

ESSA..... Every Student Succeeds Act

FFT.....Framework for Teaching

JCSEE ..... Joint Committee on Standards for Educational Evaluation

MET .....Measure of Effective Teaching

NAEP .....National Assessment of Educational Progress

NCLB.....No Child Left Behind Act

NCTQ..... National Council on Teacher Quality

NEA .....National Education Association

NIET ..... National Institute for Excellence in Teaching

PD ..... Professional Development

RTT..... Race to the Top

SCTS.....South Carolina Teaching Standards

SLOs .....Student Learning Objectives

TAP.....Teacher Advancement Program

VAMs..... Value-Added Models

CHAPTER 1

INTRODUCTION

The impact of a teacher on student learning has been documented in numerous studies (e.g., Rockoff, 2004), and teacher effectiveness was found to have a significant association with student learning outcomes (Aaronson et al., 2007; Danielson, 2001; Darling-Hammond, 2000, 2012; Hanushek, 1992, 2011; Heck, 2009; Stronge et al., 2008; Tucker & Stronge, 2005). The experience level of a teacher could predict the amount of student learning in a classroom (Muñoz et al., 2011). Students taught by the most effective teachers outperformed their peers taught by the least effective teachers by as much as one grade level (Hanushek, 1992). Differences in teacher quality resulted in a difference of 7.5 percentage points in student achievement (Rivkin et al., 2005). In addition, studies (e.g., Darling-Hammond, 2000) found that the impact of teachers on student learning outcomes was stronger than the impact of student background factors, including poverty, language, and minority status (p. 39).

A national priority in education is to improve student academic achievement. As part of an ongoing effort to improve student leaning, educational legislation and policies emphasize students' equal access to educational opportunities, quality instruction, and student success. In 1965, the Elementary and Secondary Education Act (ESEA) was signed into law and promoted equal access to education and educational standards. This act was amended and reauthorized as the No Child Left Behind Act (NCLB) in 2002. The NCLB highlighted the importance of teacher quality and required schools to address the

achievement gap among students. In response to the educational legislation and laws, states and school districts nationwide enacted policies to develop and reform educator evaluation systems to recognize and reward effective teachers to ensure teacher quality (National Council on Teacher Quality, 2017a).

Another federal program the Race to the Top (RTT) was passed in 2009. The RTT initiative encouraged states to implement educational policies in six major areas in which teacher evaluation was an important aspect. Participating states and school districts were encouraged to redesign their educator evaluation systems by employing multiple measures and multiple rating categories in evaluating teacher effectiveness. The RTT emphasized teachers' contribution to student learning outcomes and provided school districts the capacity for using the educator evaluation results to "inform professional development, compensation, promotion, retention, tenure, and removal" (U.S. Department of Education [USDOE], 2010a, p. 34).

Correspondingly, states and school districts reformed their teacher evaluation systems by taking teachers' accountability for student learning into consideration. An increasing number of states and school districts started to include measures of students' academic growth as one component of the evaluation system of teacher effectiveness. According to the National Council on Teacher Quality (NCTQ) (2015), as of 2015, 43 states required objective measures of student achievement in their teacher evaluation systems, which was a large increase from 2009 when only 15 states had the requirement. Among the 43 states, 17 states (e.g., GA, NC, NY, TN) included student achievement and growth measures as the preponderant criterion and 18 states (e.g., AZ, FL, IL, OH) included it as a significant criterion in teacher evaluation. Steinberg and Donaldson

(2016) reported that 36 of 46 states, and 20 of 23 large districts were using student test score data in evaluating teacher performance.

The Every Student Succeeds Act (ESSA) was passed as the latest reauthorization of the ESEA in 2015. The ESSA emphasized equity for high needs students. Specifically, the ESSA removed the requirement for states to develop and implement teacher evaluation systems that were established under NCLB waivers. Instead, ESSA provided states and school districts with greater flexibility in developing and implementing their teacher evaluation systems. Since the passage of ESSA, some states (e.g., Indiana, North Carolina) have modified their laws/regulations about teacher evaluation systems. For example, Utah removed measures of student achievement growth based on state standardized tests as evidence in evaluating teacher effectiveness.

Educational legislation and policies appear to have great impact on the development and implementation of teacher evaluation systems in states and school districts. The purpose of evaluation also plays an important role in the evaluation of teacher effectiveness. One major purpose is to identify effective teachers and ineffective teachers for making decisions on recruitment, retention, promotion, and compensation (Hanushek, 2009). Another major purpose is to provide teachers with feedback and professional development to improve their instructional practices. Darling-Hammond (2013) described the teacher evaluation system as a "teaching and learning system" that supports improvement for teachers throughout their career. She described:

> [I]t is important to link both formal professional development and job-embedded learning opportunities to the evaluation system. Evaluation alone will not improve practice. Productive feedback must be accompanied by opportunities to learn.

> Evaluations should trigger continuous goal-setting for areas teachers want to work on, specific professional development supports and coaching, and opportunities to share expertise, as part of recognizing teachers' strength and need. (p. 99)

The study of educator evaluation systems provides important information about school performance, classroom instruction, and student learning. Steinberg and Sartain (2015) studied the relationship of teacher evaluation and school performance and found that higher-achieving and lower-poverty schools are the primary beneficiaries, and the teacher evaluation program is most successful in advantaged schools. Studies also found that principals' leadership characteristics and their attitudes have an impact on teacher evaluation (Peterson & Peterson, 2006; Tuytens & Devos, 2010). To evaluate teacher effectiveness, the assessment and evaluation system must employ a carefully constructed set of multiple measures (NEA, 2010, p. 9).

In the evaluation of teacher effectiveness, a commonly used method is classroom observations. School administrators observe teachers' classroom instruction and rate the teachers using an observational rubric. This method focuses on teachers' classroom instructional practices and interactions with students. Traditionally, states and school districts used binary evaluation systems, and teachers were rated as either satisfactory or unsatisfactory, either met or not met. By 2017, almost all states (43) required that teacher evaluation instruments had at least three rating categories rather than being binary judgement (National Council on Teacher Quality, 2017a). Classroom observations were found to be able to provide significant, useful information about a teacher's practice if used thoughtfully, and were viewed as credible by stakeholders (Goe et al., 2008). Teachers and principals in Arizona had positive views of the reformed teacher evaluation

systems and considered the standards-based teacher observations as the most credible method of evaluation (Ruffini et al., 2014). However, there are some limitations of using classroom observation in teacher evaluation. One major limitation is that the evaluation results based on classroom observations could not differentiate teacher effectiveness. For example, only about 1% of teachers were rated in the category of unsatisfactory (Brill, 2009; Weisberg et al., 2009).

With the increased emphasis on teachers' accountability for students' academic achievement (e.g., Hanushek, 2011), students' learning outcomes are taken into consideration in evaluating teacher effectiveness. Value-Added Models (VAMs), a measure of student growth, is employed in the evaluation of teacher effectiveness. VAMs use statistical methods to examine students' academic growth based on standardized tests. The method is considered objective and cost-efficient (Goe et al., 2008), and could predict students' long-term outcomes (Chetty et al., 2014a, b; Sanders, 2000). However, approximately 70 percent of teachers who teach in subject areas that have no standardized test scores may not be evaluated using VAMs. Teachers' value-added ratings are not stable across time, vary substantially with different types of achievement tests, and depend on different groups of students (Amrein-Beardsley & Collins, 2012; Morgan et al., 2014; Papay, 2011).

Another student growth measure Student Learning Objectives (SLOs) is used as a component in teacher evaluation system. Student learning objectives are content- and grade/course-specific learning objectives that educators can validly measure to document student learning over a defined period of time (Marion et al., 2012). Teachers can select subject area, student groups, time span, and curriculum standards. They also set growth

targets for students and use instructional strategies and assessment methods to measure student growth. One major advantage of using SLOs lies in its flexibility of measuring student growth based on various types of assessment (e.g., standardized tests, performance assessment) on all subject areas. Students at the schools where teachers used SLOs had higher growth rates in reading and math than those at the schools where teachers did not use SLOs (Slotnik et al., 2013). Teachers indicated that the SLOs process is beneficial to students and teachers' professional development (Makkonen et al., 2015). Garrett and Steinberg (2015) suggested that "while practical limitations on using value-added scores across all teachers will remain, efforts by states and districts to develop SLOs for all grades and subjects may prove useful for use in conjunction with classroom observation measures, for both accountability and development purposes" (p. 239).

SLOs appears to be valued as a measure of student growth and has potentials while being used in conjunction with classroom observations in evaluating teacher effectiveness. Researchers have studied SLOs in the evaluation of teacher effectiveness in states and school districts including the Denver Public Schools (Slotnik et al., 2004), Charlotte-Mecklenburg in North Carolina (Slotnik et al., 2013), and Arizona and Utah (Makkonen et al., 2015). In particular, researchers (e.g., Balch & Springer, 2015) studied teachers' performance pay, test scores, and student learning objectives. Although states and school districts have been using SLOs as a component in their teacher evaluation system, they implement SLOs differently.

This study is intended to explore how educators in South Carolina perceive using SLOs in evaluating teacher effectiveness and whether teachers' SLO scores could better differentiate teacher performance in comparison with their classroom observation scores.

This study of SLOs is significant for the following reasons. First, teacher evaluation systems should employ multiple measures to ensure reliable and valid evaluation of teacher effectiveness. Using classroom observation alone is not adequate for evaluating teacher effectiveness. Second, federal programs (e.g., RTT) emphasize teachers' accountability for student learning, so student growth should be used as a measure of teachers' contribution to student learning in evaluating teacher effectiveness. Third, the student growth measures obtained with VAMs are not reliable, valid, or fair, and cannot measure the effectiveness of about 70 percent of teachers who are teaching subjects that have no standardized tests (Darling-Hammond, 2012; NEA, 2010). Finally, the SLOs study findings can help improve the implementation of the teacher evaluation system, inform teacher professional development, and enhance classroom instruction and assessment.

This study employed a mixed-methods sequential explanatory design and incorporated four studies to investigate using SLOs in teacher evaluation in South Carolina (SC). Data were collected at different stages. For Study 1 and Study 2, I used surveys to explore educators' perspectives of SLOs before and after the full implementation of the teacher evaluation system in SC. For Study 3, I used interviews with teachers to explore in-depth the impact and implementation of SLOs. In Study 4, I investigated the relationship between teachers' SLO scores and classroom observation scores. Through analyzing multiple sources of data collected, I sought to understand educators' views of the impact and implementation of SLOs, their understanding and knowledge about SLOs, their need for support in implementing SLOs, the successes and challenges of implementing SLOs, assessment methods in measuring student growth, and

using SLOs as an additional method in teacher evaluation. Specifically, I compared educators' views of SLOs before and after the full implementation of the Expanded ADEPT evaluation system in South Carolina. Further, in this study I also examined the relationship of teachers' SLO scores and classroom observation scores and investigated whether SLOs could better differentiate teacher performance.

A perception study is of great value in educational research. For an innovation to be successfully implemented, four characteristics including need, clarity, complexity and practicality are especially important (Fullan, 2001). The policy's practicality refers to whether the proposed change is considered practical and feasible to be implemented. In order to examine the practicality of an innovation, it is especially valuable to understand the perceptions of the relevant personnel. Weatherley and Lipsky (1977) emphasized the importance of perceptions by indicating that individuals' interpretation, engagement, and response to a policy will facilitate policy adoption, reshaping, and transformation. To examine the practicality of educational policy change, it is important to explore the insight embedded in teachers' response to change (Gitlin & Margonis, 1995). Understanding teachers' experiences and perceptions of policy change can inform the successes and challenges of the policy change (Datnow & Castellano, 2000).

In implementing a government-regulated policy in education, it is essential to understand teachers' perceptions of the characteristics of the policy (Tuytens & Devos, 2009, p. 925). Researchers van den Berg et al. (1999) argued that teachers' constructions of their systems of knowledge, skills, and attitudes towards their work can shape their professional behavior, thus impact their attitude and actions in dealing with innovations in education, and ultimately influence the implementation of the innovations. A teacher's

perception of and reaction to a new policy is influenced by his or her preexisting knowledge and worldview. The level of a teacher's engagement and buy-in to the policy change is shaped by their perceptions of the policy (Coburn, 2005; Spillane, 1999). When the reform of a policy is ambiguous or misaligned with teachers' beliefs, knowledge, and values, they often react with intense and negative emotions, and resist any changes (Muncey & McQuillan, 1996; Schmidt & Datnow, 2005). Some policy changes may increase stress, and when the stress is more pronounced, it can suppress teachers' job commitment and performance (Collie et al., 2012; Skaalvik & Skaalvik, 2011). In addition, the social and structural conditions of the schools where the teachers are affiliated, the networks used at the school, and teachers' relationships with school leaders can impact teachers' perceptions of educational policy change (Spillane, 1998).

Educator evaluation systems have been reformed to better measure teacher effectiveness. Understanding how educators perceive the coupling of student growth measures with the teacher evaluation system can inform the implementation of the educator evaluation system. As Rogers's (1995) theory of perceived attributes described, when individuals perceive an initiative as positive, they tend to adopt it. The Joint Committee on Standards for Educational Evaluation (JCSEE, 2009) also stated that legally defensible evaluation programs must provide teachers both procedural and substantive due process. Therefore, investigating educators' perceptions of the component, process, and consequences of teacher evaluation is an initial step for improving an evaluation system.

Some studies focused on the investigation of educators' perceptions of the educator evaluation systems. Cherasaro et al. (2016) indicated that teachers' views of

evaluator credibility, fairness of the evaluation, and quality of the feedback from evaluators have an impact on teachers' support of the evaluation system and their decisions for using the evaluation results to improve their teaching practice. Finster and Milanowski (2018) found that teacher perceptions of the Performance Evaluation System are interrelated and linked to perceptions of changes in teaching practices. Delvaux et al. (2013) indicated that teachers with fewer than five years of teaching experience reported greater impact of the evaluation system on their professional development than teachers with more experience. Jiang et al. (2015) found that teachers' perceptions of the educator evaluation system are dependent on teacher characteristics, and teachers' perceptions about the school leadership and professional community are positively associated with their perceptions of the evaluation system.

Educators' views of the evaluation systems seem to have an association with their overall job satisfaction. Ford et al. (2018) suggested that there is a small and positive association between teachers' views of supportive teacher evaluation experience and job satisfaction. Teachers who considered that the evaluation had positive impact on their practices reported more satisfaction. Similarly, Koedel et al. (2017) examined the impact of performance ratings on job satisfaction for public school teachers in Tennessee and found that teachers who received higher ratings tended to be more satisfied with their profession of teaching.

Although various studies investigated teacher evaluation systems, classroom observations, and VAMs (e.g., Bell et al., 2012), there is limited research focusing on using SLOs in teacher evaluation. There is a need to explore using SLOs in the evaluation of teacher effectiveness. When SLOs are used to measure student growth in evaluating

teacher effectiveness, it is important to understand educators' perceptions of SLOs. A positive perception might suggest a more successful implementation of SLOs in the teacher evaluation system. A negative perception might suggest there is a gap between educator perceptions of the system and the actual evaluation system. Understanding educators' perceptions can inform professional development. It is important to understand whether teachers and administrators perceive SLOs similarly or differently. Teachers and administrators play different roles in teacher evaluation. Generally, teachers are classroom instructors who have a direct impact on student learning. Administrators are generally the evaluators of teacher effectiveness, and they are also decision makers. Understanding the perceptions of teachers and administrators on SLOs can inform professional development for teachers and administrators, and it can further enhance communication and collaboration between teachers and their evaluators.

In investigating teachers' perceptions of SLOs, their educational background, including academic degrees and years of experience, should be taken into consideration. Their educational and professional experience might help shape their views of an educational policy. Understanding the impact of their educational experience on their perceptions of the teacher evaluation can inform professional development in implementing the evaluation system. Other factors, including training of SLOs, experience of using SLOs, participation in teacher effectiveness programs, grade levels taught, and their perceptions of observational rubrics, are also important. It can inform professional development in SLOs and teacher evaluation in general. For example, if teachers' years of experience in education have significant impact on teachers' perceptions of SLOs, the professional development should provide differentiated training

sessions for veteran teachers and early career teachers. In addition, it is very important to take school context into consideration in the investigation of teacher effectiveness. Garcia and Weiss (2019) indicated that high-poverty schools suffered the most from the shortage of high-effective and credentialed teachers. Comparing the effectiveness of teachers between advantageous and disadvantageous schools would inform educational policy making and funding opportunities for school and teacher improvement.

To examine SLOs in evaluating teacher effectiveness. I conducted four studies. In the first three studies I focused on exploring educators' views of using SLOs in teacher evaluation. In the fourth study I examined teachers' SLO scores in comparison with their classroom observation scores. In Study 1 and Study 2 I employed surveys to examine educators' perceptions of SLOs in teacher evaluation. In both survey studies I focused on educators' views of the impact of SLOs and classroom observations, their understanding and knowledge about SLOs, and the support needed to implement SLOs. I conducted the Study 1 survey before the full implementation of the teacher evaluation system in South Carolina and the Study 2 survey after the full implementation of the system. In Study 3 I employed structured interviews to explore in-depth teachers' views of the impact and implementation of SLOs. Specifically, the interview questions focused on teachers' experience of using SLOs, the impact of SLOs on teaching and learning, successes and challenges in implementing SLOs, assessment methods in measuring student growth, and the role of SLOs in teacher evaluation system. I conducted the interviews about one year after the full implementation of evaluation system in South Carolina. In addition to exploring how educators perceive using SLOs in evaluating teacher effectiveness, I also conducted Study 4 on the relationship between teachers' SLO scores and their classroom

observation scores. I intended to explore whether teachers' SLO scores could better differentiate teacher performance in comparison with their classroom observation scores.

The key purpose of the four studies is to investigate the perceptions of educators about using SLOs in teacher evaluation, and the associations of teachers' SLO scores and classroom observation scores. The findings of the studies should provide important information about SLOs as a component in the teacher evaluation system that can be used to improve the teacher evaluation system, inform teacher professional development, enhance classroom instruction and assessment, and improve student learning. The four studies collectively addressed the following questions:

1. How do school administrators and teachers perceive SLOs in evaluating teacher effectiveness before the full implementation of the teacher evaluation system?
2. Are teachers' academic degrees, years of experience in education, SLOs training, TAP participation, and their perceptions of classroom observations associated with their perceptions of SLOs before the full implementation of the teacher evaluation system?
3. How do school administrators and teachers perceive SLOs in evaluating teacher effectiveness after the full implementation of the teacher evaluation system?
4. Are teachers' academic degrees, years of experience in education, experience of using SLOs, grade levels taught, and their perceptions of classroom observations associated with their perceptions of SLOs after the full implementation of the teacher evaluation system?
5. Does using SLOs in teacher evaluation have an impact on teachers' instructional practices and students' learning outcomes?

6. What are the successes and challenges of implementing SLOs in evaluating teacher effectiveness?
7. How do teachers view the SLOs assessment methods used to measure student growth in teacher evaluation?
8. Do teachers consider SLOs as an additional reliable method in the evaluation of teacher effectiveness?
9. Can teachers' SLO scores better differentiate their performance in comparison with their classroom observation scores?
10. What are the associations between teachers' SLO scores and their classroom observation scores?

CHAPTER 2

REVIEW OF LITERATURE

The review of related literature consists of three sections. The first section reviews teacher effectiveness focusing on the characteristics of teacher effectiveness, two commonly used frameworks of teaching, and the impact of effective teaching on student learning. The second section describes the evaluation of teacher effectiveness. It introduces the purposes of teacher evaluation, teacher evaluation systems in the United States, and commonly used methods in teacher evaluation including classroom observations, student growth measures (VAMs and SLOs), and student perception survey. The third section provides information about the South Carolina teacher evaluation system the Expanded ADEPT and SLOs.

2.1 TEACHER EFFECTIVENESS

Teachers play a very important role and have significant impact on student learning (Hattie, 2009; Odden et al., 2004). To be qualified for the position of a teacher, teachers generally need to obtain certain certification and licensure, receive an academic degree, have educational experience, and have sufficient knowledge of the subject, students, and teaching strategies. These characteristics of teachers are considered as teacher quality. Studies found an association between student learning outcomes and teacher quality/characteristics including having a teaching certificate or license (Darling-Hammond & Young, 2002), years of teaching experience (Goldhaber & Brewer, 1997), and academic degrees obtained (Rowan et al., 1996).

However, obtaining these teacher characteristics and/or qualifications does not guarantee effective teaching. Teacher effectiveness focuses on specific teaching practices that involve the interactions between teachers and students in classrooms, teachers' course design and lesson planning, teachers' management of classroom learning environment, classroom assessment, and teachers' other professional activities. In addition to these instructional practices, teacher effectiveness is often linked to student learning outcomes. Goe et al. (2008) reviewed multiple research, reports, standards, and policy documents and summarized five aspects that define teacher effectiveness. They include: 1) Effective teachers have high expectations for all students and help students learn; 2) Effective teachers contribute to positive, attitudinal, and social outcomes for students; 3) Effective teachers use diverse resources to plan and structure engaging learning opportunities, monitor student progress formatively, and evaluate learning using multiple sources of evidence; 4) Effective teachers contribute to the development of classrooms and schools that value diversity and civic-mindedness; and 5) Effective teachers collaborate with other teachers, administrators, parents, and education professionals to ensure student success, particularly the success of students with special needs and those of high risk for failure (p. 8).

Regarding the characteristics and qualifications of effective teachers, there are different frameworks, among which the standards developed by the National Board for Professional Teaching Standards (NBPTS) and the Framework for Teaching (FFT) developed by Charlotte Danielson (1996) are widely used in evaluating teacher effectiveness. The NBPTS was established in 1987, with a mission of advancing the quality of teaching and learning. The NBPTS uses a rigorous, authentic, performance-

based assessment program (e.g., classroom instruction videos, student work) to measure and recognize teacher effectiveness, and certifies teachers who meet the standards and have the qualifications to teach. There are five core propositions and standards describing what accomplished teachers should know and be able to do: 1) Teachers are committed to students and their learning; 2) Teachers know the subjects they teach and know how to teach the subject to students; 3) Teachers are responsible for monitoring and organizing student learning; 4) Teachers think systematically about their practice and learn from their experience; and 5) Teachers are members of learning communities. The five core propositions and standards have been widely adopted as a framework to measure teacher performance and teacher effectiveness.

Another well-known framework in teaching, the Framework for Teaching (FFT), was developed by Charlotte Danielson in 1996. The FFT was later revised in 2007, recognizing state curriculum standards and incorporating additional research. The FFT Evaluation Instrument was developed to evaluate teacher effectiveness in 2011, and was updated and released in 2013 (Danielson, 2013). The FFT Evaluation Instrument has been widely used in teacher evaluation, school coaching and mentoring, and teacher professional development in different states (Danielson, 2013). The major goal of the FFT is to promote clear and meaningful conversation about effective teaching practices, define effective teaching, and build a strong profession. The FFT consists of 22 components of effective instruction that are clustered in four major domains: planning and preparation, classroom environment, instruction, and professional responsibilities. The FFT is generally used to rate teachers at four performance levels including

unsatisfactory, basic, proficient, and distinguished, and can be used across grade levels and subject matter.

Since the development of the FFT, researchers have tested the framework in many studies. The observational instrument FFT was validated by examining the relationship between teachers' observational scores based on FFT and students' achievement (Kane et al., 2013). These researchers found that students taught by a teacher in the top quartile scored 0.10 standard deviations higher in math and 0.13 standard deviations higher in reading than students taught by a teacher in the bottom quartile. Similarly, measure of effective teaching (MET) researchers found that the teachers' observational scores using FFT had a .19 correlation with students' math achievement and a .11 correlation with students' ELA achievement (Bill & Melinda Gates Foundation, 2012). There was a small to moderate positive association between teachers' FFT scores and student learning, with some variation by grade level and subject matter (Gallagher, 2004; Milanowski, 2004).

2.2 EVALUATION OF TEACHER EFFECTIVENESS

As the impact of effective teaching on student learning is evident, it is critically important to identify effective teachers using certain reliable and valid evaluation methods. The Joint Committee on Standards for Educational Evaluation (JCSEE, 2009) defined personnel evaluation as "the systematic assessment of a person's performance and/or qualifications in relation to a professional role and some specified and defensible institutional purpose" (p. 3). JCSEE (2009) developed standards for personnel evaluation that include: 1) The primary use of evaluations is to provide effective services to student; 2) The evaluation practices must be free of needless threatening or demoralizing characteristics; 3) The use of the personnel evaluations must adhere to culturally

competent practices; 4) Sound professional development and training experiences must result from the personnel evaluation; 5) Although disagreements may arise about what constitutes good teaching, good administration, or good research, these disagreements are necessary; and 6) Evaluations will vary in complexity and importance.

Understanding the purposes of teacher evaluation is an essential step in developing evaluation systems. Marzano (2012) identified two major purposes of teacher effectiveness evaluation: measurement and development. According to Marzano (2012), "Measuring teacher effectiveness and developing teachers are different purposes with different implications. An evaluation system designed primarily for measurement will look quite different from a system designed primarily for development" (p. 15). Similarly, the National Education Association (NEA, 2010) also explicitly stated that a comprehensive teacher assessment and evaluation system should have two distinct components: 1) ongoing, consistent, formative assessments of performance for the sole purpose of fostering professional growth and improved practice, and 2) periodic summative evaluations of teacher performance for the purpose of approving continued employment (p. 5).

Establishing the key purposes of teacher evaluation before developing an evaluation system is essential. The overarching purpose of teacher evaluation is to identify effective teachers. On one hand, teacher evaluation is formative, and the evaluation process and results are intended to inform professional development and improve classroom instruction and student learning. On the other hand, teacher evaluation is summative and linked to a teacher's contribution to student learning, and the evaluation results are used to identify effective and ineffective teachers for making

decisions about employment, promotion, and compensation. In many cases, teacher evaluation serves both purposes. For example, Donaldson (2009) and Welsh (2011) indicated that states used their teacher evaluation results to help make high-stake decisions about employment, retention, promotion, and compensation. States also use teacher evaluation results to inform teacher professional development.

To recognize effective and ineffective teachers, some states employed the credential provided by the NBPTS as an important standard, and other states developed their own teacher evaluation systems. Regarding the quality of teacher evaluation system, Darling-Hammond (2012) indicated that a high-quality teacher evaluation system should include five key element: 1) common statewide standards for teaching that are related to meaningful student learning and are shared across the profession; 2) performance assessments, based on these standards, guiding state functions such as teacher preparation, licensure, and advanced certification; 3) local evaluation systems aligned to the same standards, for evaluation on-the-job teaching based on multiple measures of teaching practice and student learning; 4) support structures to ensure trained evaluators, mentoring for teachers who need additional assistance, and fair decisions about personnel actions; and 5) aligned professional learning opportunities that support the improvement of teachers and teaching quality (p. 4-5).

Regarding the evaluation of teacher effectiveness, researchers (e.g., Brophy & Good, 1986; Campbell et al., 2004; Goe et al., 2008) emphasized that effective teaching should be evaluated through the classroom experiences that teachers create, and be associated with student achievement and students' social and emotional development as well. Therefore, states and school districts have been striving to design, develop, and

reform teacher evaluation systems that incorporated multiple methods of teacher effectiveness evaluation. By 2015, 46 states in the United States reported they had reformed their teacher evaluation systems. Generally, the newly developed educator evaluation systems are different from the previous systems in two important aspects. First, the new systems employ standards-based classroom observation protocols (standards-based observational rubrics) that are related to improving teachers' classroom instructional practice. Second, the new evaluation systems include student performance measures (e.g., VAMs, SLOs). Some states also consider feedback from students and/or their parents (Steinberg & Donaldson, 2016).

Steinberg and Donaldson (2016) summarized the components and their corresponding weights of the new teacher evaluation systems in 46 states and 23 large school districts. All these states and school districts incorporated classroom observation in their new educator evaluation systems. Thirty-six out of 46 states, and 20 out of 23 large school districts, used student growth measures in their evaluation systems. In addition to using classroom observations and student growth measures, some states also used measures of professional conduct, school-wide achievement indicators, student surveys, parent/caregiver surveys, and peer surveys. Regarding the weights of the evaluation methods, between 50 percent and 60 percent of a typical teacher's summative rating depends on classroom observation. About 20 percent (specifically 15.8 percent based on the states and 21.7 percent based on districts for VAMs, and 21.5 percent based on the states and 13.7 percent based on the districts for SLOs) of a typical tested teacher's summative rating depends on student growth measure.

Regarding the teacher evaluation systems with multiple components, researchers provided different findings. Kane et al. (2013) suggested that the combination of classroom observation scores using the FFT framework, teachers' VAM scores, and student surveys of teacher performance can identify teacher effectiveness. However, Kraft and Gilmour (2017) indicated that new evaluation systems with multiple rating categories had not necessarily resulted in more differentiated ratings (p. 237). In particular, they compared teacher rating distributions among 24 states and concluded that the weighted average of teachers rated unsatisfactory/ineffective was less than half percent (0.48 %), and only two states (Maryland and New Mexico) rated more than one percent of teachers in the lowest category. The median percentage of teachers rated above proficient varied from 6% in Georgia to 62% in Tennessee. This suggests that there is much variability in teacher ratings across states.

Studies also explored the perceptions of teachers and principals in terms of teacher evaluation systems. Moran (2017) drew on ethnographic research procedure and explored the views of first-grade teachers on a high-stake teacher evaluation system in Tennessee and found that teachers understand the importance of being held accountable, but they seem lack knowledge about the score computing process. Liu et al. (2019) investigated teachers' perceptions of the System for Educator Evaluation and Development (SEED) in Connecticut and found that the majority of teachers did not consider the evaluation feedback as effective in improving their instruction. However, these teachers did value specific, frequent, evidence-based feedback that was related to professional development opportunities. Similarly, Paufler (2018) examined principals' views of evaluating teachers based on professional practice and

student achievement by analyzing data collected from a large urban school district in Texas. The findings revealed that principals had concerns regarding the evaluation systems' negative impact on morale, their lack of autonomy in decision making on evaluating teachers and staffing, and their perceived lack of value as professionals. In addition, Marshall et al. (2016) indicated the challenges of measuring teacher effectiveness due to different content areas, grade levels, and groups of students.

Table 2.1 describes various methods used in evaluating teacher effectiveness. Classroom observation is the most commonly used method because evaluators can directly observe a teacher's performance in a classroom. School administrators are often the evaluators and observe teachers' classroom instruction and rate their performance based on an observational rubric. Student growth measures are major components in measuring teachers' contribution to student learning, and two commonly used methods are VAMs and SLOs. Teachers' self-evaluation generally adopts teacher surveys and logs to evaluate their effectiveness. Some popular surveys include the National Assessment of Educational Progress (NAEP) and the Schools and Staffing Survey's Teacher Follow-Up Survey. These surveys require teachers to self-evaluate their own teaching practices. Using logs is another way for teachers' self-evaluation, through which teachers document their instruction. Instructional logs were found to be valid, reliable, and cost-effective (Rowan et al., 2009). A student survey is used by some states to collect students' views of their teachers' performance. In addition, other methods including portfolios and analysis of classroom artifacts (e.g., student work, lesson plans, assessment) are also used in the evaluation of teacher effectiveness.

Table 2.1 *Methods Used in Evaluating Teacher Effectiveness*

Focus	Method	Brief Introduction
Teaching	Classroom Observation	This method focuses on teachers' classroom instruction, preparation, classroom management, etc. School administrators (e.g., principals), as evaluators, observe teachers for multiple times, and grade teachers' performance using an observational instrument/rubric.
Student Learning	VAMs	This method considers teachers' contribution to student learning outcomes (growth) measured by standardized tests.
	SLOs	This method considers teachers' contribution to student learning outcomes (growth) measured by various types of assessment (exam, performance assessment, etc.) to evaluate whether students have achieved the learning goals established by individual teachers or teams of teachers.
Views of Teachers and Students	Teacher Self-evaluation	This method employs teachers' self-report evaluation. Surveys and journals/log are generally used to collect and document teachers' teaching practice.
	Student Survey	This method generally employs a survey to gather students' views of their teachers' teaching practice.

2.2.1 Classroom Observations

Classroom observations are one of the most commonly used teacher evaluation system (Goe et al., 2008). States and school districts develop their observational instrument, and school administrators and/or trained evaluators observe teachers' classroom instruction and rate their performance using an observational rubric. The observational rubric is generally developed based on some types of teaching standards (e.g., FFT). It consists of multiple domains with indicators, and there are typically three to five performance levels for each indicator within the domains. For example, the National Institute for Excellence in Teaching (NIET) developed a 4.0 classroom observational rubric that is used by some states (e.g., South Carolina).

Steinberg and Donaldson (2016) summarized teacher evaluation systems in 46 states and 23 large school districts and concluded that all of them used classroom observation as part of their new educator evaluation systems. According to the teacher-level component weights, about 53.2 percent (based on the states) and 56.0 percent (based on the districts) of a typical teacher's summative rating depends on classroom observation. Thus, classroom observation scores account for the largest share of a teacher's summative rating.

Regarding the use of classroom observations in the evaluation of teacher effectiveness, studies have revealed both positive and negative findings. Classroom observations are considered as the most accepted measures of teacher effectiveness because school administrators, as evaluators, can directly observe teachers' instructional practice and class dynamics (Heneman et al., 2006). School administrators obtain feedback from teachers, students, and parents, and have a comprehensive view of teachers' contribution to their schools (Harris et al., 2014). School administrators have substantial experience of conducting teacher evaluation (Liu & Johnson, 2006), and their evaluation scores of teachers are considered as stable and credible (Weisberg et al., 2009). Researchers (e.g., Garrett & Steinberg, 2015) suggested that teachers' evaluation scores from classroom observations are more straightforward and transparent for educators to connect to their actual work, and stakeholders should feel more confident about the use of classroom observation for teacher performance evaluation (p. 225).

There are limitations of using classroom observations to measure teacher effectiveness. First, the criteria in the observational rubrics are inadequate (Danielson, 2001; Darling-Hammond, 2013; Marzano, 2012), and observational systems should be

grounded in standards-based evidence of instruction with multiple observations (Steinberg & Sartain, 2015). Second, school administrators who serve as evaluators may not have content knowledge expertise in the teachers' subject area and are incapable of accurately evaluating teacher effectiveness. Third, evaluators spend a very short time observing a class, and they may not capture a whole picture of teachers' instruction. For example, researchers Callahan and Sadeghi (2015) indicated that although multiple teacher observations are conducted under the new evaluation systems, the value of these observations are diminished due to the short time spent for each observation. Fourth, principals seem to be more focused on entering information on tablets than in actually observing (Callahan & Sadeghi, 2015, p. 56). Fifth, evaluators spent a substantial amount of time to score and tag teacher practice on multiple elements, and it is difficult to coordinate with multiple observers (Strunk et al., 2014).

In addition to the problems occurred in the evaluation process, the observational outcomes are questionable. Morgan et al. (2014) studied teachers' observational ratings across four years and indicated that observational ratings of teacher effectiveness are not stable across time. Teachers' observational scores demonstrate considerable variability based on raters and lessons taught (Rowan et al., 2013). Teachers' observational scores do not discriminate between effective teaching and ineffective teaching. For example, Brill (2009) reported that only 1.8 percent of teachers were rated in the category of unsatisfactory in New York City. Another research study revealed that 99.7 percent of teachers in Chicago were evaluated as satisfactory to distinguished (Rich, 2012). Similarly, the National Council on Teacher Quality (2015) reported that 97% of teachers in New Jersey (2013-2014), 97.7% of the teachers in Florida (2013-2014), 95% of the

teachers in New York (2012-2013), 98% of the teachers in Michigan (2012), and 98% of the teachers in Tennessee (2013) were rated as highly effective or effective in teacher evaluation (p. 12).

According to Donaldson (2009) and Varlas (2009), a single-method evaluation system could not adequately inform professional development for teachers. Garrett and Steinberg (2015) cautioned that high-stake decisions about teacher employment and promotion should not rely solely on the measures based on classroom observations. Danielson and McGreal (2000) indicated that most instruments used to assess teacher effectiveness have notable design flaws and demonstrate weak correlations with student achievement. In addition, classroom observations focus on teachers' classroom performance and interactions with a given group of students, and the measures based on observation could not meet the evaluation purpose of accountability (Welsh, 2011). Therefore, employing multiple measures of teacher effectiveness has become a common attribute of teacher evaluation systems. Many states are now using both classroom observation rubrics and student academic growth data to evaluate teacher effectiveness (Steinberg & Donaldson, 2016).

2.2.2 Student Growth Measures Using VAMs

Classroom observation has been valued as a major component in evaluating teacher effectiveness due to its capability of observing teachers' instructional practices directly. Educators do have concerns about using observations as the only method to make decisions about teacher effectiveness. It is especially questionable when the observational results are used to make high-stake decisions about teacher employment, retention, promotion, and compensation. In addition, the federal program RTT required

that states and school districts should employ multiple methods and include teachers' contribution to student learning in measuring teacher effectiveness (U.S. Department of Education [USDOE], 2010a), In response to the federal program RTT that emphasized teachers' accountability and contribution to student learning, states and school districts (e.g., New York, Florida) incorporated student growth measures as a component in their teacher evaluation systems. Student growth measures use a baseline achievement measure and project the amount of growth that students are expected to gain in a given time period. Two commonly used methods are VAMs and SLOs.

VAMs use statistical methods to examine students' academic growth measured by students' score change on standardized tests over a period of time. The statistical model uses students' prior achievement to predict their achievement the next year measured by standardized tests on a specific subject. When most students perform better than their predicted performance based on standardized test scores, the teacher is evaluated to be effective. When most students perform worse than their predicted performance based on standardized test scores, the teacher is evaluated to be ineffective. For the teachers who are teaching grades and subjects that can be assessed using standardized tests, the classroom VAMs have been often used as a common measure for student growth.

Using VAMs to evaluate teacher effectiveness has some advantages. VAMs examine students' academic growth based on standardized tests, and the measurement is objective and cost-efficient (Goe et al., 2008), and can predict students' long-term outcomes including college attendance and adult earnings (Chetty et al., 2014a, b; Sanders, 2000). According to Harris (2011), value-added growth models, "[h]old people

accountable for what they can control," which is a central tenant that must be incorporated in educator evaluation systems (p. 4).

Despite some benefits of using the VAMs in evaluating teacher effectiveness, researchers (e.g., Collins, 2014) found many limitations of using VAMs. First, VAMs can only be used to measure students' academic growth based on standardized tests. However, according to the NEA (2010), approximately 70 percent of teachers teach in subject areas that have no standardized test scores. For those teachers who teach in a subject area (e.g., music, visual arts, ESL, keyboarding) that does not have a standardized test, there will be a lack of data needed for calculating their VAM score. Second, VAMs are based on statistical theory behind inferences that assumes random assignments. However, in educational settings, students are generally not randomly assigned to the classrooms. Using VAMs to analyze the standardized test scores of students who are not randomly assigned to classrooms is problematic. This might cause unintended consequences that do not result in improving teacher performance or the educational system (Morganstein & Wasserstein, 2014, p. 109).

In addition, studies suggested that teachers' VAM scores demonstrate substantial year-to-year variability (Corcoran, 2010; Newton et al., 2010). In particular, Morgan et al. (2014) studied teachers' value-added ratings across four years, and they found that nearly two-thirds of the teachers have value-added ratings that differ two or more points, which further suggests that using value-added ratings of teacher effectiveness might not be stable across time. Similarly, Papay (2011) and Lockwood et al. (2007) found that teachers' VAM ratings vary substantially with different types of achievement tests that students take. Teachers' VAM ratings also depend on the group of students they have.

Some teachers who are rated as effective with one group of students might not be rated as effective with another group of students (Brophy & Good, 1986). The same teacher might be more effective when teaching more advanced students than when teaching students who are in special education programs or who are English language learners (Amrein-Beardsley & Collins, 2012). Researchers (e.g., Welsh, 2011) argued that the amount of gain that the students who start out to be high-performing would be smaller than those who start out to be low-performing students, which would make the teachers of the high-performing students appear to be less effective. In addition, Moran (2017) found that the use of VAMs is particularly problematic in primary grades.

Researchers also investigated the relationships between teachers' VAM scores and their classroom observation scores. Studies (e.g., Bell et al., 2012; Grossman et al., 2013) found moderate to relatively low correlations between teachers' observational scores and their VAM scores. Strunk et al. (2014) indicated that there are moderate correlations between teachers' value-added and observation-based measures, and teachers receive similar but not entirely consistent signals from each performance measure. Specifically, Harris et al. (2014) conducted a study of 30 schools about their evaluation systems and found a positive weak association between teachers' ratings by their principal and their value-added measures. Grossman et al. (2014) examined how the relationships between teachers' observation scores and their VAM scores change based on different tests used in measuring student achievement. They further indicated that unlike the classroom observations, VAMs do not provide diagnostic information for teachers to improve their instrument (Grossman et al., 2014). A recent study by Basileo and Toth (2019) investigated the association of teachers' observation scores and their

value-added scores in Florida, and found teachers' observation ratings and their value-added measures are small, positive, and statistically significant, and teachers' observation scores are the largest level one predictor of value-added measures accounting for student, teacher, observation systems, and school characteristics (p. 11).

Regarding whether VAMs should be included in the teacher evaluation system, The American Educational Research Association (AERA) suggested that states and districts need to acknowledge that "there are considerable risks of misclassification and misinterpretation in the use of VAMs" (AERA, 2015, p. 4). Similarly, Rockoff and Speroni (2010) suggested that value-added measures of effectiveness are biased so that it is important to use other information to help achieve more stability and accuracy in teacher evaluations.

Despite the concerns that educators have about using VAMs in evaluating teacher evaluation, many states use student growth ratings for evaluating individual teachers. Some states even associate classroom value-added growth to the effectiveness of the entire school. In these cases, the scoring system incorporate the school-wide component into the individual teachers' summative evaluation as a portion of the total score (Lacireno-Paquet et al., 2014). The Southern Regional Education Board (SREB, 2013) reported that most of the 16 states in the region were using a value-added model to assess student growth in their teacher evaluation systems, and they attributed between 35% and 50% of teachers' overall ratings to this measure. Steinberg and Donaldson (2016) summarized the teacher-level component weights of the new teacher evaluation systems in states and large school districts. They concluded that about 15.8 percent (based on the

46 states) and 21.7 percent (based on 23 districts) of a typical tested teacher's summative rating depend on VAMs.

2.2.3 Student Growth Measures Using SLOs

VAMs have various limitations, and one major limitation is that teachers who are teaching a subject area or a grade level that has no standardized test would not have a VAM score. Therefore, use of another student growth measure, SLOs, is gradually becoming common practice. Student learning objectives are learning targets that teachers set for their students to reach through a period of instruction (e.g., one academic year, or one semester). SLOs are defined as a set of goals that measure teachers' progress in achieving student growth targets that focus on students' expected learning at the end of the instructional period (Lachlan-Haché et al., 2012, p. 1). The American Institutes for Research (AIR, 2014) defined student learning objectives as "a measurable, long-term, academic goal informed by available data that a teacher or teacher team sets at the beginning of the year for all students or for subgroups of students".

In developing SLOs, teachers often consider six major components: student groups to be included, time span (e.g., one semester, one year), curriculum standards to be addressed, growth targets set for students, instructional strategies, and assessment methods. The process of using SLOs to measure student growth consists of 1) developing SLOs, usually constructed by an individual teacher or a team of teachers; 2) submitting SLOs for the approval from trained evaluators; 3) checking-in through midcourse conversations between teachers and evaluators; 4) reviewing SLO attainment and scoring by both teachers and evaluators to determine if the student growth targets are achieved; and 5) completing the summative rating of the teachers and reflecting on the lessons

learned from the process (Lachlan-Haché et al., 2012). There are various forms of SLOs in measuring student growth, and they could be at a course-level, at a class-level, at content-level, or focus on a specific subgroup of students. The time period for assessing SLOs also differs based on the course structure. It could cover one school year, one semester, or even one quarter of a school year (Lachlan-Haché et al., 2012).

The SLOs process is an inseparable part of instruction, and it connects curriculum standards, classroom instructions, and assessment. To understand student baselines, teachers use archival student data or diagnostic assessment results. Teachers use curriculum standards and set differentiated learning goals/objectives/targets for their students. To help students achieve these learning goals, teachers employ appropriate instructional strategies. During the instructional process, teachers use formative assessment results to track student progress and adjust instructional strategies. At the completion of the instruction, teachers use appropriate assessment methods to evaluate whether students have achieved the learning goals. Ultimately, teachers are evaluated and given an overall score based on how well their students have achieved the goals. Therefore, the SLOs process is part of instruction, and SLOs results can be used to individualize learning for students and inform instruction for teachers. As Schneider and Johnson (2019) indicated, "The SLO process is not a template that teachers complete at the beginning of the year and return to at the end of the year. It is a formative assessment process of understanding where students are in their learning and where they need to go next in regard to the SLO learning goal" (p. 142).

SLOs were initially used in Denver Public Schools in 1999, aiming at measuring student academic growth in making decisions on teachers' contributions to the growth

and pay-for-performance (CTAC, 2008). The school districts that are early adopters of SLOs also include the Austin Independent School District (Texas) and Charlotte-Mecklenburg Schools (North Carolina), and both districts started to pilot and use SLOs in evaluating teacher performance in 2008 (CTAC, 2013). States including Rhode Island, Georgia, and New York are among the early adopters of SLOs in their evaluation systems of teacher effectiveness. As of 2012, SLOs were required, recommended, or encouraged as an example of student achievement in nearly half of the states in the US (Lachlan-Haché et al., 2012, p. 1).

One major advantage of using SLOs in the evaluation of teacher effectiveness lies in its adaptability to all subject areas and grade levels. Unlike VAMs that can only be used to measure student growth based on standardized tests, SLOs can be used to measure student growth based on various types of assessment (e.g., standardized tests, performance assessment). In addition, according to Lachlan-Haché (2015), it is necessary to thoughtfully analyze data in assessing student growth in order to create meaningful learning goals for students, and SLOs data analysis of student growth also considers teaching assignment and contextual factors (e.g., school conditions, student experience). Through SLOs, teachers employ more evidence-based practices (Slotnik et al., 2013).

Multiple studies found an association between well-developed SLOs and increases in student learning outcomes. A four-year (1999-2003) study of SLOs conducted in Denver Public Schools revealed that students whose teachers have high quality SLOs perform better than their peers on standardized tests (Slotnik et al., 2004). Similarly, another five-year (2007-2012) study of SLOs conducted in Charlotte-Mecklenburg showed that students at the schools where teachers use SLOs have higher

growth rates in reading and math than the students at the schools where teachers do not use SLOs (Slotnik et al., 2013). Similarly, Makkonen et al. (2015) examined the implementation of SLOs in Arizona and Utah and found that the end-of-year SLO scores of teachers in Arizona differentiate between high- and low-performing teachers. Teachers' SLOs score have a statistically significant association with teachers' ratings based on classroom observations and student surveys of teachers. The researchers found that the SLO scores of teachers in Utah have little variation, and most teachers are rated as meeting expectations. Researchers (e.g., Balch & Springer, 2015) studied teachers' performance pay, test scores, and student learning objectives and found that teachers' SLO scores are not statistically significantly associated with their VAM scores.

Other studies focused on educators' perceptions of SLOs. Teachers hold mixed views regarding the implementation of SLOs. Some have positive views of SLOs, indicating that SLOs provides them with opportunities to use data, and they have more active engagement in their evaluation after the SLOs implementation (Donaldson, 2012; New Teacher Project, 2012). Teachers in Utah consider the SLOs process as beneficial to students and teachers' professional development. However, they do not consider that the implementation of SLOs has positive impact on instruction or their understanding of effective ways to assess students (Makkonen et al., 2015). Similarly, only fewer than half of the surveyed teachers in Connecticut reported that SLOs are useful to them as professionals, though the majority agreed that analyzing student data is valuable (Donaldson et al., 2014). Interestingly, a series of studies of SLOs in Austin Independent School District (Texas) revealed that teachers report more and more positive views of SLOs along with longer time of SLOs implementation. Teachers considered the SLOs

process as frustrating and time-consuming when SLOs were first piloted in 2007-2008 (Schmitt et al., 2008). About 48% of teachers in 2008-2009 and 68% of teacher in 2009-2010 agreed or strongly agreed that using SLOs had improved their teaching. From 2009-2010 to 2013-2014, teachers' views of SLOs' impact on their teaching had stayed stable with about two-thirds of the teachers agreeing or strongly agreeing that using SLOs had improved their teaching (Courtemanche et al., 2014; Schmitt et al., 2013). By 2013-2014, a majority of teachers indicated that that they often considered SLOs when planning and conducting their daily work, and the student achievement results of using SLOs were worth the extra work (Courtemanche et al., 2014).

There are some limitations of using SLOs in evaluating teacher effectiveness. It is challenging to develop high-quality SLOs and fully implement SLOs (Slotnik et al., 2004). As the study in Denver public schools indicated, SLOs quality had a large improvement over the four years of development and implementation (CTAC, 2008). Another limitation lies in the validity, reliability, and accuracy of teachers' SLO scores due to various factors including the quality of the assessment designed by teachers and the quality of evaluators (Crouse et al., 2016). In addition, surveys of teachers and principals in Maryland revealed that about 50% of the educators reported needing support to have access to data and analyze student data (Slotnik et al., 2014).

The U.S. Department of Education's Institute of Education Sciences published a report in 2014 that presented information on SLO implementation in 30 states (Lacireno-Paquet et al., 2014). Among these states, 21 required the approval of SLOs by an outside evaluator, such as a school principal or district master teacher. The NCTQ (2015) indicated that 22 states required or allowed the use of SLOs as student growth measures

in teacher evaluations. However, only nine out of the 22 states required that the learning objectives be reviewed and approved. Steinberg and Donaldson (2016) summarized the new teacher evaluation systems in states and large school districts. Twenty-four out of 46 states and nine out of 23 large school districts use SLOs in their educator evaluation systems. Among those states and school districts that used SLOs, nine states and one district use SLOs as their only measure of student growth. According to the teacher-level component weights, about 21.5 percent (based on the states) and 13.7 percent (based on the districts) of a typical teacher's summative rating depend on teachers' SLO scores.

2.2.4 Student Perception Survey

In the evaluation of teacher effectiveness, classroom observations and student growth measures are major components in the teacher evaluation systems. Another measure of teacher effectiveness that has been adopted in some states and school districts (e.g., Georgia, Denver Public Schools) is the student perception survey. Students are invited to complete a survey with questions about their teachers' classroom instructions, classroom management, professional responsibilities, etc. Students attend classes, observe teachers' classroom instruction, participate in classroom activities, and interact with teachers and peers on a daily basis. Students are the direct recipients of instruction and have extensive experience with their teachers (Follman, 1992), and their views of their teachers are considered as one of the important components in evaluating teacher effectiveness (Geo et al., 2008).

To collect students' views of their teachers, survey instruments are developed. One commonly used student survey instrument in evaluating teacher effectiveness in K-12 education is the Tripod Project survey developed by Cambridge Education. The

Tripod survey is designed to measure seven constructs (i.e., Seven Cs) of teaching practices including Care, Control, Clarify, Challenge, Captivate, Confer, and Consolidate. Under each construct, students are asked to rate their teachers on multiple agreement statements on a five-point Likert scale ranging from "Totally Untrue" to "Totally True." The Tripod survey is applicable to three grade bands: K-2, Grades 3-5, and Grades 6-12.

The Tripod survey has been broadly used to collect students' views of their teachers and contributes to a balanced view of teacher performance and effectiveness (Ferguson, 2010; MET Project, 2012). The Measures of Effective Teaching (MET) project sponsored by the Gates Foundation in 2009 worked to identify the measures of teacher effectiveness that could predict student achievement gains. The MET used the Tripod survey to explore student views of teacher effectiveness. Based on the study of 3,000 teacher participants in seven school districts in the United States, the MET researchers found that students perceive clear differences among teachers, and students' views of their teachers are predictive of their achievement gains. They concluded that the Tripod survey is a reliable measure of teacher effectiveness (MET Project, 2012).

Other student survey instruments used in evaluating teacher effectiveness include My Student Survey developed by Ryan Balch at Vanderbilt University, iKnowMyClass developed by Russell Quaglia at the Quaglia Institute for Student Aspiration (QISA), and the Panorama Student Survey developed by the Harvard Graduate School of Education. In particular, My Student Survey consists of six constructs: Presenter, Manager, Counselor, Coach, Motivator, and Content Expert. Under each construct, students are asked to rate their teachers on some frequency statements on a five-point Likert scale ranging from "Never" to "Every Time."

Although student surveys have been broadly employed in evaluating faculty performance in higher education, they are comparatively new in evaluating PK-12 teachers. It can be traced back to as early as 1896 when students in a city in Iowa were invited to provide views on effective teacher characteristics (Follman, 1995). Educators and policy makers are increasingly becoming aware of the importance of including student surveys in teacher evaluation. As of 2015, at least 23 states required or encouraged school districts to include student surveys as a measure of teacher effectiveness in teacher evaluation (Center on Great Teachers and Leaders, 2015). Though some states include student surveys in their educator evaluation systems, the weights given to student surveys are comparatively small. Based on the component weights of the new teacher evaluation systems in 46 states and 23 large school districts, between 1.0 percent and 2.4 percent of a teacher's summative rating is accounted for by student surveys (Steinberg & Donaldson, 2016).

Studies of using student surveys in evaluating teacher effectiveness revealed that students are capable of distinguishing effective teachers from ineffective teachers (Follman, 1992, 1995), high student survey ratings of teachers correlate with high academic achievement, engagement, and self-efficacy (Balch, 2012), and student responses of teacher performance are reliable, valid, and stable over time at the classroom level (Ferguson, 2010). In particular, Wilkerson, Manatt, Rogers, and Maughan (2000) conducted a study involving 2,000 K-12 students and found that student ratings of teachers significantly predict student achievement, but ratings by principals and teachers themselves are not significant predictors of student achievement. They further suggested that student ratings are the best predictors of their achievement across all subjects.

Similarly, a most recent study by Kearney and Garfiend (2019) also indicated that students' perceptions of teacher effectiveness significantly contribute to the variance in middle grade mathematics achievement.

Goe et al. (2008) indicated that using student surveys is cost- and time-efficient, can be collected anonymously, and requires minimal training, thus they recommended that student ratings of teachers were worth considering for inclusion in teacher evaluation systems (p. 40). Similarly, Peterson et al. (2000) also indicated that student surveys are valid and reliable for the evaluation of teacher effectiveness. English et al. (2016) suggested that "student survey instruments can be a valuable component of a comprehensive teacher evaluation system" (p. 11), and student surveys "can and should" be used as one of the measures of teacher effectiveness (Marzano & Toth, 2013, p. 75).

While many researchers (e.g., Goe et al., 2008) acknowledged the advantages of using student surveys in evaluating teacher effectiveness, some (e.g., Follman, 1992, 1995; Popham, 2013) cautioned that students' lack of knowledge, issues of confidentiality, and reliability of the student survey might lead to an inaccurate evaluation of teacher effectiveness. In particular, Goe et al. (2008) indicated that students might not be qualified to evaluate teachers regarding curriculum, content knowledge, classroom management, and collegiality. Other researchers (Lamb et al., 2013) indicated that teachers have concerns about student surveys, arguing that some students might not understand the survey, or might not take the survey seriously. Therefore, researchers (e.g., Goe et al., 2008) suggested that student ratings of teacher effectiveness should not be a primary measure of teacher effectiveness.

2.3 SOUTH CAROLINA TEACHER EVALUATION SYSTEM

In line with many other states, the South Carolina Department of Education (SCDE) is dedicated to developing and implementing the evaluation of teacher effectiveness. In 1998, the ADEPT (Assisting, Developing, and Evaluating Professional Teaching) system was first designed as an educator evaluation initiative. All teachers in South Carolina were required to complete the ADEPT requirements to be eligible for a professional teaching certificate. In 2012, the SCDE was granted a waiver from the ESEA requirements and redesigned its educator evaluation system the Expanded ADEPT and includes student growth measures. In 2015, ESSA was passed as the latest reauthorization of the ESEA, allowing states greater flexibility in educator evaluation systems. Correspondingly, the SCDE reformed its educator evaluation system to be Expanded ADEPT Support and Evaluation System.

South Carolina Teaching Standards (SCTS) 4.0 is the primary formal evaluation model for classroom-based teachers (Table 2.2) and SLOs data are collected as an artifact that supports ratings of teachers within professional practice domains. The SCTS 4.0 rubric is based on the performance standards designed and validated by the National Institute for Excellence in Teaching (NIET). The SCTS 4.0 includes four domains: instruction, planning, environment, and professionalism. There are 12 indicators of instruction, three indicators of planning, four indicators of environment, and four indicators of professionalism. Each indicator is rated using a 4-point scale (1-Unsatisfactory; 2-Needs Improvement; 3-Proficient; 4-Exemplary). South Carolina used the classroom observational rubric in the evaluation of teacher effectiveness in some pilot

schools and started to fully implement the SCTS 4.0 in all schools in the 2018-2019 academic year.

Table 2.2 *The SCTS 4.0 Domains and Indicators*

Instruction	Planning	Environment	Professionalism
Standards and Objectives	Instructional Plans	Expectations	Growing and Developing Professionally
Motivating Students	Student Work	Managing Student Behavior	Reflecting on Teaching
Presenting Instructional Content	Assessment	Environment	Community Involvement
Lesson Structure and Pacing		Respectful Culture	School Responsibilities
Activities and Materials			
Questioning			
Academic Feedback			
Grouping Students			
Teacher Content Knowledge			
Teacher Knowledge of Students			
Thinking			
Problem Solving			

In South Carolina the Expanded ADEPT educator evaluation system, the SLOs are required for all classroom teachers and are used as an artifact to support teachers' ratings based on the SCTS indicators (Expanded ADEPT Support and Evaluation System Guidelines, 2018). The SLOs focuses on measuring teachers' ability to set appropriate targets for student learning, accurately measure and analyze student growth, plan, implement, and adjust instructions, and ensure student growth. Table 2.3 describes SLOs scoring criteria. Based on the holistic rubric, there are four performance levels ranging from 1 (Unsatisfactory) to 4 (Exemplary). For example, if a teacher sets up rigorous goals for students, uses appropriate assessments to monitor student progress, strategically revises instruction, and between 90% and 100% of his/her students meet their growth

targets, the teacher obtains 4 points (Exemplary). If a teacher inconsistently uses assessments, fails to monitor progress or adjust instruction based on progress monitoring data, and 0% - 50% of students meet their growth targets, this teacher obtains 1 point (Unsatisfactory).

Table 2.3 *South Carolina SLOs Scoring Rubric*

Rating	Criteria
Exemplary 4	• 90% - 100% of students have met their growth target. • Educator set up rigorous, superior goal(s); skillfully used appropriate assessments, continuously monitored progress; and strategically revised instruction in response to ongoing progress monitoring.
Proficient 3	• 75% - 89% of students have met their growth target. • Educator set up attainable goal(s); used appropriate assessments, consistently monitored progress; adjusted instruction in response to progress monitoring.
Needs Improvement 2	• 51% - 74% of students have met their growth target. • Educator set up goal(s); used assessments that were not appropriate for the goal, inconsistently monitored progress; inconsistently or inappropriately adjusted instruction.
Unsatisfactory 1	• 0% - 50% of students have met their growth target. • Educator inconsistently used assessments, failed to monitor progress; failed to adjust instruction based on progress monitoring data.

Teachers' SLO scores are used as a modifier for the teacher's overall evaluation ratings. If a teacher earns an SLOs score of 4 points, there will be an increase of 0.25 points in the teacher's overall evaluation rating. If a teacher obtains an SLO scores of 2 or 3 points, there will be no change on the teacher's overall evaluation ratings. If a teacher earns an SLO score of 1 point, there will be a decrease of 0.25 points in the teacher's overall evaluation rating. The SCDE requires that the SLOs must be completed as a part of the evaluation process. If a teacher fails to complete the SLOs, the teacher will score 1 point on SLOs, which will result in a decrease of 0.25 points in his/her overall rating.

Teachers' overall rating is based on a 4-point composite score scale. A teacher obtains a performance level of Unsatisfactory with a composite score of 1.24 points or below. A teacher obtains a performance level of Needs Improvement with a composite score ranging between 1.25 and 2.25 points. A teacher obtains a performance level of Proficient with a composite score ranging between 2.26 and 3.75 points. A teacher obtains a performance level of Exemplary with a composite score of 3.76 or above. The final evaluation results have two categories: Not Met and Met. A performance level of Unsatisfactory or Needs Improvement results in an overall effectiveness rating of Not Met. A performance level of Proficient or Exemplary results in an overall effectiveness rating of Met. All districts are required to implement the Expanded ADEPT Support and Evaluation system starting from 2018-2019 school year. According to the Expanded ADEPT Support and Evaluation System Guidelines (2018), school districts report evaluation data to the State Board of Education (SBE) including, but not limited to, overall effectiveness ratings, observation results, and student learning objective scores annually.

CHAPTER 3

METHODS

This research employed a mixed-methods sequential explanatory design (Creswell & Creswell, 2018) to investigate SLOs as a measure of student growth in the evaluation of teacher effectiveness. In the first phase, I collected survey data in 2017, which is more than one year before the full implementation of the teacher evaluation system in South Carolina (2018-2019). In the second phase, I collected both the survey data and the interview data in 2019-2020, which is more than one year after the full implementation of the teacher evaluation system. Data from the two surveys were used to compare educators' views of SLOs before and after the full implementation of the evaluation system. The interview data were collected to gain an in-depth understanding about the impact and implementation of SLOs in the evaluation of teacher effectiveness through the lens of teachers. The qualitative data collected through the interviews with teachers were used to help explain, illustrate, and elaborate the results from analyzing the quantitative data. In addition, teachers' evaluation scores were used to examine the relationship of their SLO scores and classroom observation scores.

A major advantage of using quantitative data from surveys lies in its capacity of collecting responses from a large number of participants, and the results tend to be generalized to a large population. However, quantitative data could not provide detailed information about respondents' explanation, elaboration, or reasoning for selecting certain items in the survey. Therefore, interview data are generally collected to

complement the survey findings. In addition, results from qualitative data through interviews could provide detailed and specific information that could potentially help improve the intervention programs and inform policy making.

This research incorporated four studies using both quantitative and qualitative data collected at different stages. The first two studies employed surveys to explore how educators including teachers and administrators perceive using SLOs and classroom observations in the evaluation of teacher effectiveness. The views of administrators and teachers were compared both before and after the full implementation of the teacher evaluation system in the state. In addition, the two studies explored teachers' views of the impact of SLOs and classroom observations based on their personal and professional background information (e.g., academic degree, years of experience in education, SLOs training, TAP participation, SLOs experience, grade levels taught).

Study 3 used interviews to investigate teachers' views of the impact and implementation of SLOs. Study 3 focused on the impact of SLOs on teaching and learning, the successes and challenges of implementing SLOs at schools, assessment methods used to measure student academic growth, whether SLOs are an additional reliable method in the evaluation of teacher effectiveness. Study 4 examined associations of teachers' SLO scores and their classroom observation scores, and it sought to understand whether SLO scores could better differentiate teacher performance in comparison with their classroom observation scores. To protect the privacy of the teachers who participated in the studies, I used pseudonyms for teachers, schools, and school districts in the studies. Table 3.1 provides an outline of the four studies.

Table 3.1 *Outline of the Four Studies*

	Study 1	Study 2	Study 3	Study 4
Research Questions	• How do school administrators and teachers perceive SLOs in evaluating teacher effectiveness before the full implementation of the evaluation system? • Are teachers' academic degrees, years of experience in education, SLOs trainings, TAP participation, and their perceptions of classroom observations associated with their perceptions of SLOs before the full implementation of the evaluation system?	• How do school administrators and teachers perceive SLOs in evaluating teacher effectiveness after full implementation of the evaluation system? • Are teachers' academic degrees, years of experience in education, experience of using SLOs, grade levels taught, and their perceptions of classroom observations associated with their perceptions of SLOs after the full implementation of the evaluation system?	• Does using SLOs in teacher evaluation have an impact on teachers' instructional practices and students' learning outcomes? • What are the successes and challenges of implementing SLOs in evaluating teacher effectiveness? • How do teachers view the SLOs assessment methods used to measure student growth in teacher evaluation? • Do teachers consider SLOs as an additional reliable method in the evaluation of teacher effectiveness?	• Can teachers' SLO scores better differentiate teachers' performance in comparison with their classroom observation scores? • What are the associations between teachers' SLO scores and their classroom observation scores?
Participants	438 educators	289 educators	18 teachers	275 teachers
Instrument	Survey	Survey	Interviews	Evaluation scores
Analysis	Quantitative	Quantitative & Qualitative	Qualitative	Quantitative

3.1 STUDY 1: SURVEY (BEFORE FULL IMPLEMENTATION)

Study 1 explored educators' perceptions of SLOs before the full implementation of the educator evaluation system in South Carolina. Data were collected through the Research, Evaluation, and Measurement (REM) Center at the University of South Carolina. This study was intended to address two major research questions: 1) How do school administrators and teachers perceive SLOs in evaluating teacher effectiveness before the full implementation of the teacher evaluation system? 2) Are teachers' academic degrees, years of experience in education, SLOs trainings, TAP participation, and perceptions of classroom observations associated with their perceptions of SLOs before the full implementation of the teacher evaluation system?

3.1.1 Participants

Participating schools were involved in two programs: The Teacher Advancement Program (TAP) and the South Carolina Department of Education (SCDE) Partnership Program. The TAP is a performance-based compensation system in South Carolina as part of a federal grant titled Teacher Incentive Fund (TIF). TAP encourages schools to recruit, evaluate, and compensate teachers based on their performance. The SCDE Partnership Program involves four partner school districts for a professional learning initiative. Schools in the four districts are either a priority or a focus school that has a high need to improve. The four districts created professional learning plans in collaboration with the SCDE to build partnership structures, provide support to administrators, coaches, and teachers, and ultimately to be prepared for full implementation of the educator evaluation system.

Survey participants consisted of 438 educators from 36 schools in 13 districts in South Carolina (Table 3.2). Twenty-four schools within nine school districts are in the TAP, and 12 schools within four school districts are in the SCDE Partnership Program. Among the 438 participants, the majority (about 95%) are teachers, and 5% are school administrators. About 63% of the participants have a master's degree or above, and about 37% have a bachelor's degree or below. About 87% the participants are career teachers with more than three years of experience in education, and about 13% are early career teachers who have three or fewer years of experience in education.

Table 3.2 *Study 1 Survey Participants*

Variable	Level	N	%
Position	Teacher	416	95.0
	Administrator	22	5.0
Degree	Bachelors or Below	162	37.0
	Masters or Above	276	63.0
Experience in Education	Early Career (0-3)	57	13.1
	Career (3+)	379	86.9

3.1.2 Instrument

A survey was used as the instrument of this study (Appendix A). The survey is intended to measure three dimensions of educators' perceptions of SLOs. The first dimension includes four questions related to the impact of SLOs on educator effectiveness, instruction, student leaning, and teachers' professional development. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). The second dimension includes nine questions focusing on educators' knowledge about SLOs. The questions are on a 4-point scale (1-No Knowledge; 2-Limited Knowledge; 3-Some Knowledge; 4-Substantial Knowledge). The third dimension includes six questions about educators' need for support to successfully implement SLOs.

The questions are on a 3-point scale (1-Need No Support; 2- Need Some Support; 3-Need A Lot of Support). The survey also includes four additional questions about educators' views of the impact of classroom observations. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). In addition, participants' educational background information including highest academic degree, years of experience in education, SLO training received, and TAP participation is included in the survey.

This instrument is a revised version of the survey that was previously used in a project at the REM Center. The original instrument was designed to be administered to evaluators of teacher effectiveness. I made revisions in language to make these items applicable for the teachers and administrators. I developed additional questions to assess the elements and implementation procedures associated with SLOs. Based on the 175 responses in the previous project, the reliability of the first dimension is .83 (Cronbach's alpha), and the reliability of the second dimension is .76 (Cronbach's alpha). These values of Cronbach's alphas are acceptable according to Nunnally and Bernstein (1994) who suggested an alpha coefficient of .70 as an acceptable level. To ensure the validity of the revised instrument used for this study, I invited five experts in the fields of educator evaluation, survey design, and classroom instruction to review the instrument. Revisions were made based on feedback from the reviewers, and the revised survey instrument was used for this study.

I calculated the reliability coefficients for the four dimensions based on the survey respondents in this study (Table 3.3). All four dimensions including the impact of SLOs, knowledge about SLOs, support needed in implementing SLOs, and the impact of

classroom observations have good reliability coefficients with Cronbach's alpha coefficients ranging from .88 to .96. The Cronbach's alpha coefficients are acceptable according to Nunnally and Bernstein (1994).

Table 3.3 *Study 1 Reliability of the Survey Subscales*

Category	Valid Responses	Number of Items	Cronbach's Alpha	Scale
Impact of SLOs	366	4	0.92	1-4
Knowledge about SLOs	397	9	0.96	1-4
Support Needed in Implementing SLOs	381	6	0.88	1-3
Impact of Observational Rubric	270	4	0.94	1-4

3.1.3 Data Collection and Analysis

The survey was administrated online using SurveyMonkey. An email invitation was sent to the principals to facilitate responses within the participating schools. The email message explained the purpose of the survey and provided a clickable button to begin the on-line survey. Each week, for four weeks, a reminder email was sent to those who had not completed the survey. In addition, principals were contacted by phone to encourage distribution of the surveys within their schools. Data collection took approximately five weeks in 2017.

The survey data are ordinal, and the assumptions of normal distribution using parametric tests (e.g., t-test, ANOVA) are generally violated. However, statisticians (e.g., de Winter & Dodou, 2010) suggested that parametric tests are valid with non-normal data when a large sample is used for data analysis. This study involved 438 educators, and I used parametric tests in the analyses. Considering the small sample of administrators, I also used non-parametric methods (e.g., Mann-Whitney tests) to check the results from the parametric tests.

To answer the first research question, I used descriptive statistics about educators' perceptions of the impact of SLOs, their knowledge/understanding about SLOs, support needed to implement SLOs, and the impact of classroom observations. The percentages of educators who agreed or strongly agreed on the statements within each dimension were reported. I calculated means of educators' responses on the items within each dimension to better understand educators' views on each specific item within the dimension. I compared administrators' views and teachers' views on these aspects. Inferential statistics were calculated to understand whether teachers and administrators had statistically significant differences on their views of the aspects. I also constructed 95% confidence intervals for the measures of central tendency. Within each school, I examined whether administrators and teachers had consistent views.

To answer the second research question, I examined the associations of teachers' views of SLOs and their educational background factors including academic degrees, years of experience in education, SLOs trainings, and TAP participation. Considering the factors are categorical, I compared central tendency and variation of perceptions within each factor. Inferential statistics were used to understand whether teachers who had different educational background had statistically significant differences on their views of the SLOs and classroom observations. I also constructed 95% confidence intervals for the measures of central tendency. In addition, I calculated correlations between teachers' views of SLOs and classroom observations. Finally, I tested models using multiple regression analysis to assess the unique impact of teachers' educational background variables and their views of classroom observations on their perceptions of SLOs. In the analysis, I was interested in understanding teachers' views of the overall impact of SLOs

that could possibly be predicted by their views of the overall impact of classroom observations, their overall knowledge about SLOs, and the overall support they needed. Therefore, I calculated the means of the dimensions based on educators' responses to the items within the dimensions.

3.2 STUDY 2: SURVEY (AFTER FULL IMPLEMENTATION)

Study 2 examined South Carolina educators' perceptions of SLOs in teacher evaluation. It focused on the impact of SLOs on evaluating teaching effectiveness, improving classroom instructional practices, promoting student learning, and informing professional development. It also explored educators' reported knowledge about SLOs and their need for support in successfully implementing SLOs. This study was intended to address two major research questions: 1) How do school administrators and teachers perceive SLOs in evaluating teacher effectiveness after the full implementation of the teacher evaluation system? 2) Are teachers' academic degrees, years of experience in education, SLOs experience, grade levels taught, and perceptions of classroom observations associated with their perceptions of SLOs after the full implementation of the teacher evaluation system?

3.2.1 Participants

I used a stratified random sampling method to recruit study participants. For all the school districts in South Carolina, I used the poverty index and enrollment as two criteria and divided the school districts into eight groups. I randomly selected one school district from each group, and my initial plan was to include educators from the eight selected school districts. However, three school districts (Riverview, Bloom, and Glover) agreed to participate in this study. Glover school district has a low poverty and a medium

enrollment, Riverview has a medium poverty and a large enrollment, and Bloom has a high poverty and a small enrollment.

To examine whether these three school districts are representative of all school districts in South Carolina, I compared them on multiple indicators including school poverty levels, school location, and school enrollment. The South Carolina school report card data from 2018-2019 school year was used for retrieving data about school poverty and school enrollment. The South Carolina Department of Education E-rate data file from 2017-2018 school year was used for the information about school location. In comparing the school information, only elementary, middle, and high schools are included. Table 3.4 describes the comparison between three selected districts and all school districts in South Carolina.

Table 3.4 *Study 2 Sample School Districts vs All School Districts*

Indicator	Level	Three Selected School Districts	All School Districts in SC
School Poverty Levels	Low (50% or Below)	34.6%	21.3%
	Medium (50%-75%)	46.2%	40.1%
	High (75% or Above)	19.2%	38.6%
School Location	Rural	42.3%	48.4%
	Urban	57.7%	51.6%
School Enrollment	Small (500 or Below)	19.2%	39.7%
	Medium (501-1000)	55.8%	47.1%
	Large (1001 or Above)	25.0%	13.2%

The percentages of schools are reported within each category. For the school poverty levels, it appears that the three selected school districts have higher percentages of schools that have low poverty levels. For the school location, the three selected school districts have slightly higher percentages of schools that are in urban areas. For school enrollment, the three selected school districts have higher percentages of schools that

have large enrollment. Although the three school districts might not perfectly represent all school districts in South Carolina based on the three indicators, the educators' views collected from the three districts may still provide valuable information about the teacher evaluation system in South Carolina.

Participants in this study consisted of 289 educators from three school districts in South Carolina (Table 3.5). Among the 289 respondents, the majority (about 89.7%) are teachers, 5.9% are school administrators, and 4.4% are others including coaches, school counselors, and media specialists. More than two-thirds (68.7%) of the participants reported that their highest degree was an educational specialist (Ed.S), masters, or Ph.D., and fewer than one-third (31.3%) have a bachelor's degree or below. Many educators (88.4%) reported having more than three years of experience in education. About three-quarters of educators (74.0%) reported having been using SLOs for more than three years. Among the teacher participants, 63.3% teach PK-5, 15.9% teach Grades 6-8, 17.9% teach Grades 9-12, and 2.8% teach students across grade levels.

Table 3.5 *Study 2 Survey Participants*

Variable	Level	N	%
Position	Teachers	244	89.7
	Administrators	16	5.9
	Others	12	4.4
Degree	Bachelor's or Below	83	31.3
	Ed.S., Master's, Ph.D.	182	68.7
Experience in Education	0-3 Years	31	11.6
	3+ Years	236	88.4
Experience in SLOs	0-3 Years	68	26.0
	3+ Years	194	74.0
Grades Level	PK-5	159	63.3
	Grades 6-8	40	15.9
	Grades 9-12	45	17.9
	Across levels	7	2.8

3.2.2 Instrument

A survey was used as the instrument of the study (Appendix B). The survey is very similar to the survey used for Study 1, and both measure three dimensions of educators' perceptions of SLOs. The first dimension is about perceptions of the impact of SLOs on educator effectiveness, instruction, student achievement, and teachers' professional development. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). The second dimension assesses educators' understanding/knowledge about SLOs. The questions are on a 4-point scale (1-No Knowledge; 2-Limited Knowledge; 3-Some Knowledge; 4-Substantial Knowledge). The third dimension assesses educators' need for support to successfully implement SLOs. The questions are on a 3-point scale (1-Need No Support; 2-Need Some Support; 3-Need A Lot of Support). In addition, four questions about teachers' perceptions of the impact of classroom observations are included. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). Furthermore, participants' demographic information including current position (administrator or teacher), highest academic degree, years of experience in education, years of experience in using SLO, and grade levels taught were collected in the survey. These variables were considered in the analysis of teachers' views of SLOs and classroom observations.

The Cronbach's Alpha coefficients were calculated to examine the reliability of the four subscales (Table 3.6). All four dimensions including the impact of SLOs, knowledge about SLOs, support needed in implementing SLOs, and the impact of classroom observations had high reliability with Cronbach's alpha coefficients ranging

from .92 to .93. These Cronbach's alpha coefficients suggest good reliability for the subscales (Nunnally & Bernstein, 1994).

Table 3.6 *Study 2 Reliability of the Survey Subscales*

Category	Valid Responses	Number of Items	Cronbach's Alpha	Scale
Impact of SLOs	273	4	0.92	1-4
Knowledge about SLOs	280	9	0.93	1-4
Support Needed in Implementing SLOs	283	6	0.92	1-3
Impact of Observational Rubric	264	4	0.93	1-4

3.2.3 Data Collection and Analysis

To collect survey responses, I solicited assistance from three district leaders to help disseminate survey links to the school administrators and teachers within their districts. As appreciation of the effort from the three school districts, I developed reports based on the responses of the educators in the districts. I used SurveyMonkey to collect responses from participants. An email invitation was sent to district leaders to facilitate responses from teachers and administrators within the district. The email message explained the purpose of the survey and provided a clickable button to begin the on-line survey. During the process of six weeks, three reminder emails were sent to those who had not completed the survey.

The survey data are ordinal, and the assumptions of normal distribution using parametric tests (e.g., t-test, ANOVA) are generally violated. However, statisticians (e.g., de Winter & Dodou, 2010) suggested that parametric tests are valid with non-normal data when a large sample is used for data analysis. This study involved 289 educators, and I used parametric tests in the analyses. Considering the small sample of administrators, I

also used non-parametric methods (e.g., Mann-Whitney tests) to check the results from the parametric tests.

To answer the first research question in this study, I used descriptive statistics about educators' perceptions of the impact of SLOs, their knowledge/understanding about SLOs, the support needed to implement SLOs, and the impact of classroom observations. The percentages of educators who agreed or strongly agreed to the statements within each dimension were reported. I calculated means of educators' responses to the items within each dimension to better understand educators' views on each specific item within the dimension. I compared administrators' views and teachers' views on these aspects. Inferential statistics were used to examine whether teachers and administrators had statistically significant differences on their views of the aspects. I also constructed 95% confidence intervals for the measures of central tendency.

To answer the second research question in this study, I examined the associations of teachers' views of SLOs and their educational background factors including academic degrees, years of experience in education, experience of using SLOs, and grade levels taught. Considering the factors are categorical, I compared central tendency and variation of perceptions within each factor. Inferential statistics were used to understand whether teachers who had different educational backgrounds had statistically significant differences on their views of the SLOs and classroom observations. I also constructed 95% confidence intervals for the measures of central tendency. In addition, I calculated correlations between teachers' views of SLOs and classroom observations. Finally, I tested models using multiple regression analysis to assess the unique impact of teachers' educational background variables and their views of classroom observations on their

perceptions of SLOs. In the analysis, I was interested in understanding teachers' views of the overall impact of SLOs that could possibly be predicted by their views of the overall impact of classroom observations, their overall knowledge about SLOs, and the overall support they needed. Therefore, I calculated the means of the dimensions based on educators' responses to the items within the dimensions.

3.3 STUDY 3: INTERVIEWS

This study employed interviews to explore teachers' views of the impact and implementation of SLOs in the evaluation of teacher effectiveness. The interviews were conducted about one year after the full implementation of the Expanded ADEPT teacher evaluation system in South Carolina. SLOs were a required component in the evaluation system. Teachers had at least one year of experience of using SLOs in teacher evaluation. This study was intended to address four research questions: 1) Does using SLOs in teacher evaluation have an impact on teachers' instructional practices and students' learning outcomes? 2) What are the successes and challenges of implementing SLOs in evaluating teacher effectiveness? 3) How do teachers view the SLOs assessment methods used to measure student growth in teacher evaluation? 4) Do teachers consider SLOs as an additional reliable method in the evaluation of teacher effectiveness? In addition, teachers' experiences of using SLOs, the professional development regarding SLOs, their confidence about using SLOs, and the support needed for implementing SLOs were also discussed in the interviews.

3.3.1 Participants

Participants consisted of 18 teachers who had some experience of using SLOs in the evaluation of teacher effectiveness (Table 3.7). Participation was voluntary. The 18

teachers were from 18 different schools within 10 school districts in South Carolina. This study used a purposeful sampling method to select teachers for the interviews. During participants recruitment, school poverty, school location, educational levels, subjects taught, and teaching experience were taken into consideration.

Table 3.7 *Study 3 Teacher Information*

Teacher	Subjects Taught	Gender	Teaching Experience (Years)	SLOs Experience (Years)
Maci	All subjects	Female	24	7
Tina	All subjects	Female	25	3
Vanassa	All subjects	Female	7	5
Mary	Music	Female	38	5
Susan	Visual Arts	Female	12	4
Lisa	All Subjects	Female	14	5
Hedi	World Languages	Female	24	4
Kori	Science	Female	19	5
Daniel	Social Studies	Male	4	4
Camilia	ELA	Female	8	3
Candice	Music	Female	26	5
Lisa	ELA	Female	19	7
Olivia	Mathematics	Female	3	3
Adde	ELA	Female	19	9
Jane	Science	Female	22	5
Kara	Health Science	Female	5	5
Katie	US History	Female	15	4
David	AP Government	Male	8	4

Among the 18 teachers, six teach in elementary schools, six teach in middle schools, and six teach in high schools. They teach various subjects including all subjects (elementary), ELA, mathematics, science, social studies, music, visual arts, world languages, US history, AP government, and health science. Sixteen teachers are female and two are male. The teacher participants have various years of teaching experience with veteran teachers who have more than 20 years of teaching experience and early career teachers who have fewer than five years of teaching experience. Teachers have between three and nine years of experience of using SLOs.

The 18 teachers were from 18 different schools located at different areas with different poverty levels in South Carolina (Table 3.8). Four schools are located in city areas, nine schools at suburban areas, and five schools in rural areas. The school location is based on the National Center for Education Statistics (NCES) Education Demographic and Geographic Estimates (EDGE) program. Some teachers were from high-poverty schools and some were from low-poverty schools. The school poverty indexes range from 23.88 to 84.68, with a higher number indicating higher poverty. The poverty index is based on the school poverty information provided by the South Carolina Department of Education in 2019.

Table 3.8 *Study 3 School Information*

School Name	District	School Location	Poverty Range
Lone Oak Elementary School	Lake	Suburban	60-80
Waterville Elementary School	Littlewood	Rural	60-80
Clear Lake Elementary School	Deer Valley	City	20-60
Bear Valley Elementary School	Hillside	Suburban	60-80
Pleasant valley Elementary School	Ocean	Suburban	20-60
Bayshore Elementary School	Springhill	City	80-99
Mountainview Middle School	Richmond	Suburban	20-60
Freedom Middle School	Horizon	Rural	60-80
Waterfalls Middle School	Richmond	City	60-80
Summers Middle School	Richmond	Rural	80-99
Rainbow Middle School	Littlewood	Rural	60-80
Littlerock Middle School	Lakeview	Suburban	60-80
Garden Grove High School	Richmond	Suburban	20-60
Maple Leaf High School	Greenland	Suburban	60-80
Sun Valley High School	Richmond	City	20-60
Eastview High School	Ocean	Suburban	20-60
Apple Valley High School	Hillside	Suburban	20-60
Pinewood High School	Greenland	Rural	20-60

3.3.2 Instrument

The instrument of this study is an interview protocol with nine questions about using SLOs in teacher evaluation (Appendix D). The nine questions focus on teachers' views of the impact of SLOs on teaching and learning, the successes and challenges in implementing SLOs, the assessment methods used to measure student growth, and whether SLOs is an additional reliable method in evaluating teacher effectiveness. In addition, the interview questions also include teachers' experience of using SLOs, the professional development about SLOs, and the support needed in implementing SLOs. The interview protocols were developed based on the basic procedures of implementing SLOs. Five experts in the field of teacher evaluation, SLOs, qualitative studies, and teacher education were invited to review the interview protocol to ensure the validity of the instrument. The initially developed protocol was revised based on comments and feedback from the reviewers.

3.3.3 Data Collection and Analysis

The interviews were conducted through phone calls in the fall of 2019. Participants were recruited through recommendations from teachers, parents, colleagues, and friends. The interviews were between 15 minutes and 35 minutes. The interviews were recorded after obtaining permission from each participant. Each participant was paid $30 for their time and input. I utilized a qualitative analysis method for this study. R for Qualitative Data Analysis (RQDA) was used as a software for the data analysis. In data analysis, I followed four steps. First, I transcribed the 18 interview recordings. Second, I read the transcripts and identified responses that are relevant to the specific research questions, and the responses that were irrelevant to the research questions were

excluded. Third, I used RQDA to code the transcripts and reviewed the codes for patterns, and constructed themes based on the patterns of the codes within each question. Finally, I summarized and interpreted the themes. In addition, some representative responses were identified as quotes to help better understand the findings of the study.

3.4 STUDY 4: TEACHERS' EVALUATION SCORES

Study 4 employed a quantitative method to examine teachers' evaluation scores based on both SLOs and classroom observations. The data set was obtained from the Teacher Advancement Program (TAP) in South Carolina. The TAP program is a performance-based compensation system that was developed when South Carolina received a Teacher Incentive Fund (TIF) grant. TAP system encourages schools to recruit, evaluate, and compensate teachers, and ensure effective teaching and improve student academic achievement. To evaluate teacher effectiveness, the TAP system uses multiple methods including classroom observation and student growth measures. In this study, the SLO scores and classroom observation scores of 275 teachers who were involved in the fourth year of the TAP program were used in the analysis.

One purpose of this study was to examine whether teachers' SLO scores could better differentiate teacher performance compared with their classroom observation scores. The other purpose was to investigate the relationships between teachers' SLO scores and their classroom observation scores. Factors including school level, school poverty, and teacher type were considered in comparing teachers' SLO scores and their observational scores. Study 4 was intended to answer two research questions: 1) Can teachers' SLO scores better differentiate their performance in comparison with their

classroom observation scores? 2) What are the associations between teachers' SLO scores and their classroom observation scores?

3.4.1 Participants

Participants in this study consisted of 275 teachers who were evaluated using both the SLOs and classroom observations (Table 3.9). The teachers are from 16 TAP schools in four school districts in South Carolina. School enrollment ranged from 120 to 1200, and school poverty index ranged from 42% to 94%. The teachers taught at different grade levels ranging from Pre-K to 12th grade. Among the 275 teachers, 67 (24.4%) taught prekindergarten to Grade 5, 19 (6.9%) taught middle school grade levels, 103 (37.5%) taught high school grade levels, and 86 (31.3%) taught across grade level (i.e., elementary and middle, middle and high). Among the 275 teachers, 229 (83.3%) were career teachers, 23 (8.4%) were master teachers, and 23 (8.4%) were mentors. The teachers had between one and four years of experience of using SLOs at the time when they were evaluated.

Table 3.9 *Study 4 Participants*

Variable	Level	N	%
Type	Career teachers	229	83.3
	Master teachers	23	8.4
	Mentors	23	8.4
Grades Level	PK-5	67	24.4
	Grades 6-8	19	6.9
	Grades 9-12	103	37.5
	Across levels	86	31.3

3.4.2 Data Source

Data were collected through the TAP program. Both classroom observations and student growth measures are used to evaluate teacher effectiveness for the TAP schools. For the measure of student growth, teachers are either evaluated using the VAMs or

SLOs. In this study, teachers who were evaluated using classroom observations and SLOs were included because the purpose this study was to examine the relationships of teachers' SLO scores and their classroom observation scores. To maintain the confidentiality, TAP school teachers' personal information was removed from the data file. School administrators observed teachers teaching in the classroom and scored their teaching using an observational rubric during the implementation of the TAP. At the same time, these teachers were required to use SLOs to measure their students' growth as part of the evaluation of their teaching effectiveness. Teachers either individually or collaboratively established learning objectives/goals/targets for their students. The learning targets were approved by evaluators who were either school or district administrators. At the end of one semester or one school year, the teachers were evaluated based on the number/percentages of their students who achieved the learning goals. Therefore, each TAP school teacher had an SLOs score and a classroom observation score. Both assessments used a 5-point scale with a higher score point indicating more effectiveness.

3.4.3 Data Analysis

This study employed a quantitative method to investigate teachers' SLO scores and their classroom observation scores. To explore whether teachers' SLO scores could better differentiate teacher performance, I calculated frequencies of each score point for teachers' SLO scores and classroom observation scores. By comparing the percentages of teachers who obtained each score point ranging from 1 to 5, I was able to identify the distribution of teachers' SLO scores and their classroom observation scores. To examine the relationship between teachers' SLO scores and their classroom observation scores, I

conducted a correlation analysis. I used Spearman's rho correlation considering the ordinal data features of the scores. Additionally, I calculated the means and standard deviations of teachers' evaluation scores based on district, school type, school poverty, and teacher type to examine whether these factors have an association with teachers' evaluation scores. The data are ordinal, and the assumptions of normal distribution using parametric tests (e.g., t-test, ANOVA) are generally violated. However, statisticians (e.g., de Winter & Dodou, 2010) suggested that parametric tests are valid with non-normal data when a large sample is used for data analysis. This study involved 275 teachers, and I used both parametric tests (e.g., ANOVA) and non-parametric tests (e.g., Kruskal-Wallis H tests) depending on the cell counts in the analyses.

CHAPTER 4

RESULTS

The results are presented based on the four individual studies. Study 1 results focus on educators' views of SLOs and classroom observations before the full implementation of the educator evaluation system in South Carolina. Study 2 results focus on educators' views of SLOs and classroom observations after the full implementation of the educator evaluation system in South Carolina. Study 3 results focus on teachers' views on the impact and implementation of SLOs. Study 4 results are related to the relationship between teachers' SLO scores and their classroom observation scores.

4.1 STUDY 1 RESULTS

Based on the survey responses of 438 educators from South Carolina, I present the following results. First, I present teachers' and administrators' views of SLOs, their knowledge about SLOs, and support needed to implement SLOs. Second, I present teachers' perceptions of SLOs based on academic degrees, years of experience in education, SLOs trainings, and TAP participation. Third, I present educators' views of classroom observations, compare administrators' and teachers' views of classroom observations, and present teachers' views based on academic degrees, years of experience in education, SLOs trainings, and TAP participation. In addition, I summarize educators' views of the impact of SLOs and classroom observations. Finally, I present a multiple

regression analysis to better understand which variables might have the most impact on educators' views of SLOs.

4.1.1 Educators' Perceptions of SLOs

I analyzed educators' perceptions of SLOs using both percentages and means (Table 4.1). About two-thirds of the educators agreed or strongly agreed that using SLOs evaluates teacher performance effectively. About three-quarters of the educators agreed or strongly agreed that using SLOs improves teachers' instructional practice, promotes student learning, and informs teachers' professional development. On average, between 90% and 95% of the administrators agreed or strongly agreed on the four aspects about the impact of SLOs, and between 64.5% and 78.3% of the teachers agreed or strongly agreed on the four aspects about the impact of SLOs.

I calculated the average score for each item on the SLO impact, and I also calculated the average score based on all items to understand the overall impact of SLOs. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement.

On average, educators agreed on the impact of SLOs (M = 2.83), administrators reported more positive views (M = 3.19) than teachers (M = 2.82). Independent t-tests were conducted to help understand whether the views of teachers and administrator are significantly different. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The

alpha for the overall impact of SLOs was set to be .05. Based on the analysis results, there was a statistically significant difference between teachers' views and administrators' views of the overall impact of SLOs (p = .001, Cohen's d = 0.65), and SLOs' evaluating teacher performance effectively (p < .001, Cohen's d = 0.62) (Cohen, 1988). On average, in comparison with teachers, administrators had statistically significantly higher agreement with the statement that SLOs evaluate teacher performance effectively and the overall impact of SLOs.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the item indicating belief in whether SLOs can be used to evaluate teacher performance effectively is likely between 0.19 and 0.59 points higher for administrators than for teachers, indicating that administrators hold a slightly higher belief in this statement than teachers. The average score of the four items regarding the beliefs in the impact of SLOs is likely between 0.17 and 0.57 points higher for administrators than for teachers, indicating that administrators hold a higher belief in the overall impact of SLOs than teachers.

To check the analysis, I also used Mann-Whitney Tests considering the sample size of administrators is small. Based on the analysis results, there was a statistically significant difference between teachers' views and administrators' views of the overall impact of SLOs (Z = -2.84, p = .005) and SLOs' impact on student learning (Z = -2.53, p = .012). Overall, teachers and administrators demonstrated statistically significantly different views of the impact of SLOs, with administrators holding more positive views of the impact of SLOs than teachers.

Table 4.1 *Educators' Perceptions of SLOs*

Impact of SLOs	Percentage			Mean				
	All	Teacher	Administrator	All	Teacher	Administrator	*P*-value	95% CI
Evaluating teacher performance effectively	66.0	64.5	95.0	2.68	2.66	3.05	.000	[0.19, 0.59]
Improving teachers' instructional practice	75.1	74.0	90.5	2.86	2.84	3.25	.024	[0.06, 0.76]
Promoting student learning	79.1	78.3	95.0	2.94	2.92	3.35	.014	[0.09, 0.77]
Informing teacher PD	74.8	73.9	90.0	2.86	2.85	3.10	.142	[-0.09, 0.59]
Overall	73.8	72.7	92.6	2.83	2.82	3.19	.001	[0.17, 0.57]

Notes: Respondents include 365-373 teachers and 20 administrators.

I analyzed educators' knowledge about SLOs using both percentages and means (Table 4.2). Between 81.7% and 92.4% of teachers and between 81.0% and 95.3% administrators reported having some or substantial knowledge on the nine aspects of SLOs. It appears that most of the teachers and administrators seemed to have some to substantial knowledge about SLOs and SLOs implementation. Comparatively, large percentages of teachers reported having knowledge about the purpose of SLOs (91.4%) and content to be included in SLOs (92.4%). Large percentages of administrators reported having knowledge about the purpose of SLOs (95.2%), student groups to be included in SLOs (95.3%), content to be included in SLOs (95.2%), and setting growth targets for SLOs (95.3%). In comparison with teachers, larger percentages of administrators reported having some or substantial knowledge on seven out of the nine aspects of SLOs. In comparison with administrators, larger percentages of teachers reported having some or substantial knowledge on two out of the nine aspects of SLOs. Overall, administrators reported slightly more knowledge than teachers.

I calculated the average score for each item on the SLO knowledge, and I also calculated the average score based on all items to understand the overall knowledge about SLOs. The questions are on a 4-point scale (1-No Knowledge; 2- Limited Knowledge; 3-Some Knowledge; 4- Substantial Knowledge). A higher score point indicates more knowledge. An average score ranging from 3.5 to 4 points indicates substantial knowledge, a score ranging from 2.5 to 3.49 points indicates some knowledge, a score ranging from 1.5 to 2.49 points indicates limited knowledge, and a score ranging from 1 to 1.49 points indicates no knowledge.

On average, educators reported having some knowledge about SLOs (M = 3.23), and administrators had slightly more knowledge (M = 3.29) than teachers (M = 3.23). For all nine items, administrators had slightly larger means than teachers on six items, and teachers had slightly larger means on three items. To understand whether teachers and administrator reported significantly different levels of knowledge about SLOs, I used independent t-tests. With 9 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .006 (i.e., .05/9). The alpha for the overall knowledge about SLOs was set to be .05. On average, there were no statistically significant differences of reported knowledge between teachers and administrators.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.56 points) of the knowledge about SLOs between administrators and teachers on each of the items, as well as the average of all items.

To check the analysis, I also used Mann-Whitney Tests considering the sample size of administrators is small. The results were consistent with those by independent t-tests, and teachers and administrators did not report significantly different levels of their overall knowledge about SLOs or any of the nine aspects of SLOs. It appears that most of the teachers and administrators felt that they had some or substantial knowledge about the SLOs and SLOs implementations. The differences of the reported knowledge between teachers and administrators were very small.

Table 4.2 *Educators' Knowledge about SLOs*

Knowledge about SLOs	Percentage			Mean				
	All	Teacher	Administrator	All	Teacher	Administrator	*P*-value	95% CI
Purpose of SLOs	91.6	91.4	95.2	3.33	3.33	3.43	.498	[-0.20, 0.40]
Student groups to be included in SLOs	88.4	88.0	95.3	3.25	3.25	3.38	.404	[-0.18, 0.45]
Content to be included in SLOs	92.6	92.4	95.2	3.34	3.34	3.43	.526	[-0.19, 0.38]
Implementation of SLOs in the district	82.2	81.7	90.5	3.08	3.07	3.29	.206	[-0.12, 0.56]
Developing high quality SLOs	85.2	85.6	81.0	3.17	3.18	3.10	.606	[-0.40, 0.23]
Selecting appropriate assessments	87.1	87.2	85.7	3.21	3.21	3.24	.868	[-0.28, 0.33]
Setting growth targets for SLOs	86.9	86.4	95.3	3.19	3.19	3.24	.763	[-0.26, 0.35]
Instructional strategies to meet SLOs targets	89.3	89.3	90.4	3.25	3.25	3.24	.951	[-0.31, 0.29]
Analyzing student assessment data	89.4	89.3	90.4	3.25	3.25	3.24	.934	[-0.31, 0.28]
Overall	88.1	87.9	91.0	3.23	3.23	3.29	.670	[-0.21, 0.32]

Notes: Respondents include 382-384 teachers and 21 administrators.

I analyzed educators' support needed to implement SLOs using both percentages and means (Table 4.3). Between 39.9% and 64.5% of the teachers and between 38.1% and 85.0% of the administrators reported needing some or a lot of support in implementing different aspects of SLOs. Overall, slightly more than half of the teachers and administrators reported that they needed some or a lot of support in implementing SLOs. Comparatively, large percentages of teachers and administrators reported needing support in setting growth targets, analyzing assessment data, and developing assessments. small percentages of teachers and administrators reported needing some or a lot of support in understanding and implementing standards in SLOs. In comparison with teachers, larger percentages of administrators reported needing some or a lot of support in five out of the six aspects of SLOs.

I calculated the average score for each item on the support needed, and I also calculated the average score based on all items to understand the overall support needed. The questions are on a 3-point scale (1-Need No Support; 2-Need Some Support; 3-Need A Lot of Support). A higher score point indicates more support needed. An average score ranging from 2.5 to 3 points indicates need for a lot of support, a score ranging from 1.5 to 2.49 points indicates need for some support, a score ranging from 1 to 1.49 points indicates need for no support.

On average, educators reported needing some support in implementing SLOs (M = 1.62), and administrators reported slightly more support needed (M = 1.71) than teachers (M = 1.62). For all six aspects of support in implementing SLOs, administrators reported needing more support than teachers in implementing standards in SLOs, understanding the cognitive levels of the standards, developing assessment for SLOs, and

setting growth targets for SLOs. Administrators and teachers reported the same level of support needed regarding understanding standards of SLOs and analyzing assessment data for SLOs.

To understand whether teachers and administrator reported significantly different support needed in implementing SLOs, I used independent t-tests. With 6 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .008 (i.e., .05/6). The alpha for the overall support needed to implement SLOs was set to be .05. Based on the analysis results, there were no statistically significant differences between teachers' and administrators' overall support needed (p = .369) or any of the six aspects of SLOs. It further suggests that teachers and administrators reported similar levels of support needed in implementing SLOs.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.49 points) of the average support needed between administrators and teachers on each of the items, as well as the average of all items.

To check the analysis, I also used Mann-Whitney Tests considering the sample size of administrators is small. The results are consistent with those by the independent t-tests, and there were no statistically significant differences between teachers' and administrators' overall support (Z = -0.98, p = .326) or any of the six aspects of SLOs. The consistent findings suggest that teachers and administrators needed similar levels of support in implementing SLOs.

Table 4.3 *Educators' Need for Support in Implementing SLOs*

Need for Support in Implementing SLOs	Percentage			Mean				
	All	Teacher	Administrator	All	Teacher	Administrator	*P*-value	95% CI
Understanding standards in SLOs	39.8	39.9	38.1	1.43	1.43	1.43	.983	[-0.25, 0.24]
Implementing standards in SLOs	42.9	42.7	47.7	1.48	1.47	1.52	.710	[-0.21, 0.31]
Understanding cognitive levels of standards	59.3	58.7	70.0	1.66	1.66	1.75	.511	[-0.18, 0.37]
Developing assessments for SLOs	61.3	60.3	80.9	1.69	1.68	1.90	.107	[-0.05, 0.49]
Setting growth targets for SLOs	65.6	64.5	85.0	1.76	1.75	1.95	.162	[-0.08, 0.48]
Analyzing assessment data in SLOs	62.9	62.4	71.4	1.71	1.71	1.71	.971	[-0.26, 0.27]
Overall	55.3	54.8	65.5	1.62	1.62	1.71	.369	[-0.11, 0.30]

Notes: Responses of 371-375 teachers and 20-21 administrators.

To explore whether teachers and administrators have consistent views within each school, I calculated the means based on schools. Among 36 schools that were involved in this study, 11 schools had both administrators and teachers who responded to the survey. Regarding the impact of SLOs, most schools (nine out of 11) had administrators who reported more positive views than teachers. As for the knowledge about SLOs, fewer than half (five out of 11) of the schools had administrators who reported more knowledge than teachers. Concerning the need for support in implementing SLOs, eight out of the 11 schools had administrators who reported needing more support than teachers.

In addition, to examine whether teachers and administrators within the same school have consistent views, I conducted Pearson correlation analysis. The relationship between teachers' and administrators' views of the impact of SLOs was small and not statistically significantly different from 0 (r = -.303, p = .365). The relationship between teachers' and administrators' knowledge about SLOs was small and not statistically significantly different from 0 (r = .119, p = .728). The relationship between teachers' and administrators' need for support was also small and not statistically significantly different from zero (r = .101, p = .767). Therefore, I concluded that teachers and administrators did not report consistent views of the impact of SLOs, did not report similar levels of knowledge about SLOs, and did not report similar levels of support needed in implementing SLOs within the same school. It further suggests that teachers and administrators within he same school did not necessarily hold similar views on the impact of SLOs, have similar levels of knowledge about SLOs, or needed similar levels of support in implementing SLOs.

Table 4.4 *Educators' Perceptions Within Schools (Means)*

School	SLOs Impact		SLOs Knowledge		SLOs Support	
	Teachers	Administrators	Teachers	Administrators	Teachers	Administrators
Fairview High	2.75	3.17	3.04	3.44	1.67	1.50
Bent Primary	2.44	3.00	3.05	3.11	1.58	1.83
Central Middle	2.86	3.00	2.98	2.11	1.81	2.00
Legacy School	2.88	3.50	3.36	3.06	1.48	1.67
Moon High	2.78	3.63	3.21	3.28	1.59	1.50
Mountainview Elementary-Middle	2.88	3.25	3.29	3.26	1.63	2.00
Mountainview High	2.59	3.00	3.38	3.11	1.54	2.00
Summer High	3.00	3.00	3.00	2.89	1.17	1.67
Summer Middle	2.28	3.75	3.40	4.00	1.58	1.33
Victory Elementary	3.38	3.25	2.50	3.67	1.50	1.67
Wall Elementary-Middle	3.23	3.00	3.54	3.44	1.42	2.00

Notes: Number of teachers within each school ranges from 1-49, number of administrators within each school ranges from 1-3.

4.1.2 Teachers' Perceptions of SLOs Based on Variables

This section focuses on teachers' views of the impact of SLOs, their knowledge about SLOs, and support needed to implement SLOs. Specifically, I examined whether there was an association between teachers' perceptions of SLOs and their academic degrees, years of experience in education, SLOs training, and school participation in the TAP. I calculated the average score for each item on the SLO impact, and I also calculated the average score based on all items to understand the overall impact of SLOs (Table 4.5). The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, teachers who had a master's degree or above reported the same or slightly more positive views of the impact of SLOs than those who had a bachelor's degree or below. To understand whether the differences were statistically significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of SLOs was set to be .05. On average, there were no statistically significant differences of views between teachers who had a master's degree or above and those who had a bachelor's degree or below.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.25 points) of the perceptions of SLOs between teachers with a

bachelor's degree or below and those with a master's degree or above on each of the items, as well as the average of all items.

Table 4.5 *Teachers' Perceptions of SLOs Based on Degree*

Impact of SLOs	Bachelor or Below	Master or Above	P-value	95% CI
Evaluating teacher performance effectively	2.63	2.68	.603	[-0.12, 0.21]
Improving teachers' instructional practices	2.82	2.86	.666	[-0.13, 0.20]
Promoting student learning	2.92	2.92	.970	[-0.16, 0.16]
Informing teacher PD	2.79	2.88	.286	[-0.07, 0.25]
Overall	2.79	2.83	.520	[-0.10, 0.19]

Notes: 140-145 teachers had a degree of bachelor or below, and 224-229 had a master's or above.

Table 4.6 describes teachers' reported knowledge about SLOs based on their degrees. I calculated the average score for each item on the SLO knowledge, and I also calculated the average score based on all items to understand the overall knowledge about SLOs. The questions are on a 4-point scale (1-No Knowledge; 2- Limited Knowledge; 3-Some Knowledge; 4- Substantial Knowledge). A higher score point indicates more knowledge. An average score ranging from 3.5 to 4 points indicates substantial knowledge, a score ranging from 2.5 to 3.49 points indicates some knowledge, a score ranging from 1.5 to 2.49 points indicates limited knowledge, and a score ranging from 1 to 1.49 points indicates no knowledge. On average, teachers reported having some knowledge about SLOs, with means ranging from 3.0 to 3.49. Teachers who had a master's degree or above reported more knowledge about SLOs than those who had a bachelor's degree or below on all nine aspects of SLOs. To understand whether these differences were statistically significant based on teachers' degree, I used independent t-tests. With 9 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .006 (i.e., .05/9). The alpha for the

overall knowledge about SLOs was set to be .05. Based on the analysis results, there were no statistically significant differences of knowledge between teachers who had a master's degree or above and those who had a bachelor's degree or below.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.29 points) of the knowledge about SLOs between teachers with a bachelor's degree or below and those with a master's degree or above on each of the items, as well as the average of all items.

Table 4.6 *Teachers' Knowledge about SLOs Based on Degree*

Knowledge about SLOs	Bachelor or Below	Master or Above	P-value	95% CI
Purpose of SLOs	3.28	3.35	.319	[-0.07, 0.21]
Student groups to be included in SLOs	3.18	3.29	.145	[-0.04, 0.26]
Content to be included in SLOs	3.31	3.35	.517	[-0.09, 0.18]
Implementation of SLOs in the district	3.00	3.11	.171	[-0.05, 0.27]
Developing high quality SLOs	3.09	3.23	.071	[-0.01, 0.28]
Selecting appropriate assessments	3.13	3.26	.082	[-0.02, 0.27]
Setting growth targets for SLOs	3.12	3.23	.130	[-0.03, 0.26]
Instructional strategies to meet SLOs targets	3.19	3.28	.222	[-0.05, 0.23]
Analyzing student assessment data	3.16	3.31	.038	[0.01, 0.29]
Overall	3.16	3.27	.094	[-0.02, 0.23]

Notes: 147-149 teachers had a degree of bachelor or below, and 233-235 had a master's or above.

Table 4.7 describes teachers' need for support in implementing SLOs based on their degrees. I calculated the average score for each item on the support needed, and I also calculated the average score based on all items to understand the overall support needed. The questions are on a 3-point scale (1-Need No Support; 2-Need Some Support; 3-Need A Lot of Support). A higher score point indicates more support needed. An average score ranging from 2.5 to 3 points indicates need for a lot of support, a score

ranging from 1.5 to 2.49 points indicates need for some support, a score ranging from 1 to 1.49 points indicates need for no support. On average, teachers reported needing some support in implementing SLOs, with means ranging from 1.42 to 1.76. Teachers reported very similar support needed to implement SLOs regardless of their academic degrees. To understand whether the differences of support needed were significant, I used independent t-tests. With 6 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .008 (i.e., .05/6). The alpha for the overall support needed to implement SLOs was set to be .05. Based on the analysis results, there were no statistically significant differences of support needed between teachers who had a master's degree or above and those who had a bachelor's degree or below.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.14 points) of the support needed between teachers with a bachelor's degree or below and those with a master's degree or above on each of the items, as well as the average of all items.

Table 4.7 *Teachers' Need for Support Based on Degree*

Need for Support in Implementing SLOs	Bachelor or Below	Master or Above	P-value	95% CI
Understanding standards in SLOs	1.42	1.44	.813	[-0.10, 0.13]
Implementing standards in SLOs	1.49	1.47	.767	[-0.14, 0.10]
Understanding cognitive levels of standards	1.70	1.64	.339	[-0.19, 0.07]
Developing assessments for SLOs	1.68	1.69	.912	[-0.12, 0.14]
Setting growth targets for SLOs	1.76	1.74	.788	[-0.15, 0.11]
Analyzing assessment data in SLOs	1.76	1.68	.175	[-0.22, 0.04]
Overall	1.64	1.61	.601	[-0.13, 0.07]

Notes: 142-144 teachers had a degree of bachelor or below, and 229-233 had a master's or above.

In the analysis, teachers' years of experience in education was considered. Table 4.8 describes teachers' views of the impact of SLOs based on their years of experience in education. I calculated the average score for each item on the SLO impact, and I also calculated the average score based on all items to understand the overall impact of SLOs. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, career teachers who had more than three years of experience in education reported notably more positive views of the impact of SLOs than the early career teachers who had three or fewer years of experience. To understand whether the differences were significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of SLOs was set to be .05. On average, career teachers reported statistically significantly more positive views than early career teachers regarding the overall impact of SLOs with a small to medium effect size (p = .001, Cohen's d = 0.45), impact of SLOs on evaluating teacher performance effectively with a small to medium effect size (p = .003, Cohen's d = 0.44), improving teachers' instructional practices with a small to medium effect size (p = .006, Cohen's d = 0.38), and promoting student learning with a small to medium effect size (p = .011, Cohen's d = 0.36) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the item indicating belief in whether SLOs can be used to evaluate teacher performance effectively is likely between 0.12 and 0.60 points higher for career teachers than for early career teachers. The average score on the item indicating belief in whether SLOs can be used to improve teachers' instructional practices is likely between 0.09 and 0.55 points higher for career teachers than for early career teachers. The average score on the item indicating belief in whether SLOs can be used to promote student learning is likely between 0.07 and 0.51 points higher for career teachers than for early career teachers. The average score on the four items indicating belief in the overall impact of SLOs is likely between 0.13 and 0.53 points higher for career teachers than for early career teachers. These results indicate that career teachers hold higher belief in these statements than early career teachers.

Table 4.8 *Teachers' Perceptions of SLOs Based on Experience*

Impact of SLOs	Early Career (0-3)	Career (3+)	P-value	95% CI
Evaluating teacher performance effectively	2.35	2.71	.003	[0.12, 0.60]
Improving teachers' instructional practices	2.57	2.89	.006	[0.09, 0.55]
Promoting student learning	2.67	2.96	.011	[0.07, 0.51]
Informing teacher PD	2.63	2.88	.025	[0.03, 0.48]
Overall	2.53	2.86	.001	[0.13, 0.53]

Notes: 51-53 teachers had 0-3 years of experience, 312-320 had 4 or more years of experience.

Table 4.9 describes teachers' reported knowledge about SLOs based on their years of experience in education. I calculated the average score for each item on the SLO knowledge, and I also calculated the average score based on all items to understand the overall knowledge about SLOs. The questions are on a 4-point scale (1-No Knowledge; 2- Limited Knowledge; 3- Some Knowledge; 4- Substantial Knowledge). A higher score

point indicates more knowledge. An average score ranging from 3.5 to 4 points indicates substantial knowledge, a score ranging from 2.5 to 3.49 points indicates some knowledge, a score ranging from 1.5 to 2.49 points indicates limited knowledge, and a score ranging from 1 to 1.49 points indicates no knowledge. On average, career teachers who had more than three years of experience in education reported more knowledge about SLOs than early career teachers who had three or fewer years of experience on all nine aspects. To understand whether the differences were significant, I used independent t-tests. With 9 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .006 (i.e., .05/9). The alpha for the overall knowledge about SLOs was set to be .05. On average, career teachers reported statistically significantly more knowledge than early career teachers regarding their overall knowledge with a small to medium effect size (p = .006, Cohen's d = 0.38), the purposes of SLOs with a small to medium effect size (p = .001, Cohen's d = 0.45), developing high quality of SLOs with a small to medium effect size (p = .003, Cohen's d = 0.41), and setting growth targets for SLOs with a small to medium effect size (p = .003, Cohen's d = 0.42) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the knowledge about the purpose SLOs is likely between 0.13 and 0.53 points higher for career teachers than for early career teachers. The average score on the knowledge about developing high quality SLOs is likely between 0.11 and 0.52 points higher for career teachers than for early career teachers. The average score on the knowledge about setting growth targets for SLOs is likely between 0.11 and 0.51 points higher for career teachers than for early career teachers.

The average score based on the nine items on the knowledge about SLOs is likely between 0.07 and 0.42 points higher for career teachers than for early career teachers. These results indicate that career teachers have more knowledge than early career teachers regarding the purpose SLOs, developing high quality SLOs, setting growth targets for SLOs, and overall knowledge.

Table 4.9 *Teachers' Knowledge about SLOs Based on Experience*

Knowledge about SLOs	Early Career (0-3)	Career (3+)	P-value	95% CI
Purpose of SLOs	3.04	3.37	.001	[0.13, 0.53]
Student groups to be included in SLOs	3.06	3.27	.039	[0.01, 0.43]
Content to be included in SLOs	3.21	3.36	.126	[-0.04, 0.34]
Implementation of SLOs in the district	2.89	3.09	.070	[-0.02, 0.43]
Developing high quality SLOs	2.91	3.22	.003	[0.11, 0.52]
Selecting appropriate assessments	2.98	3.25	.011	[0.06, 0.47]
Setting growth targets for SLOs	2.92	3.23	.003	[0.11, 0.51]
Instructional strategies to meet SLOs targets	3.08	3.27	.047	[0.00, 0.39]
Analyzing student assessment data	3.06	3.28	.025	[0.03, 0.42]
Overall	3.02	3.26	.006	[0.07, 0.42]

Notes: 52-53 teachers had 0-3 years of experience, 327-329 had 4 or more years of experience.

Table 4.10 describes the teachers' need for support in implementing SLOs based on their years of experience in education. I calculated the average score for each item on the support needed, and I also calculated the average score based on all items to understand the overall support needed. The questions are on a 3-point scale (1-Need No Support; 2-Need Some Support; 3-Need A Lot of Support). A higher score point indicates more support needed. An average score ranging from 2.5 to 3 points indicates need for a lot of support, a score ranging from 1.5 to 2.49 points indicates need for some support, a score ranging from 1 to 1.49 points indicates need for no support. On average, teachers

reported very similar support needed to implement SLOs regardless of their years of experience in education. To understand whether the differences of support needed were significant, I used independent t-tests. With 6 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .008 (i.e., .05/6). The alpha for the overall support needed was set to be .05. Based on the analysis results, there were no statistically significant differences of support needed between career teachers and early career teachers.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.25 points) of the support needed between early career teachers and career teachers on each of the items, as well as the average of all items.

Table 4.10 *Teachers' Need for Support Based on Experience*

Need for Support in Implementing SLOs	Early Career (0-3)	Career (3+)	P-value	95% CI
Understanding standards in SLOs	1.37	1.44	.393	[-0.10, 0.24]
Implementing standards in SLOs	1.45	1.48	.759	[-0.15, 0.20]
Understanding cognitive levels of standards	1.76	1.64	.184	[-0.30, 0.06]
Developing assessments for SLOs	1.63	1.69	.504	[-0.12, 0.25]
Setting growth targets for SLOs	1.71	1.75	.609	[-0.14, 0.24]
Analyzing assessment data in SLOs	1.75	1.70	.649	[-0.23, 0.14]
Overall	1.61	1.62	.921	[-0.13, 0.15]

Notes: 49-51 teachers had 0-3 years of experience, 319-322 had 4 or more years of experience.

SLOs training that teachers had received was considered as a factor. Table 4.11 describes teachers' views of the impact of SLOs based on the training they had received. I calculated the average score for each item on the SLO impact, and I also calculated the average score based on all items to understand the overall impact of SLOs. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A

higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, teachers who received SLOs training reported more positive views of the impact of SLOs than those who did not receive SLOs training. To understand whether the differences were significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of SLOs was set to be .05. Based on the analysis results, teachers who received trainings reported statistically significantly more positive views than those who did not regarding the overall impact of SLOs with a small to medium effect size (p = .009, Cohen's d = 0.34) and the impact of SLOs on teachers' instructional practices with a small to medium effect size (p = .002, Cohen's d = 0.40) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the item indicating belief in whether SLOs can be used to improve teachers' instructional practices is likely between 0.12 and 0.54 points higher for teachers who received training than those who did not receive training, indicating that teachers who received training hold a slightly higher belief in this statement than those who did not. The average score on the four items indicating belief in the overall impact of SLOs is likely between 0.06 and 0.43 points higher for teachers who received training than those who did not receive training, indicating that teachers who

received training hold a slightly higher belief in the overall impact of SLOs than those who did not.

Table 4.11 *Teachers' Perceptions of SLOs Based on Training*

Impact of SLOs	No Training	Training	P-value	95% CI
Evaluating teacher performance effectively	2.45	2.71	.019	[0.04, 0.47]
Improving teachers' instructional practices	2.57	2.90	.002	[0.12, 0.54]
Promoting student learning	2.72	2.97	.016	[0.05, 0.45]
Informing teacher PD	2.69	2.89	.057	[-0.00, 0.40]
Overall	2.61	2.86	.009	[0.06, 0.43]

Notes: 63-65 teachers did not receive training and 297-305 received training.

Table 4.12 describes teachers' reported knowledge about SLOs based on training. I calculated the average score for each item on the SLO knowledge, and I also calculated the average score based on all items to understand the overall knowledge about SLOs. The questions are on a 4-point scale (1-No Knowledge; 2- Limited Knowledge; 3- Some Knowledge; 4- Substantial Knowledge). A higher score point indicates more knowledge. An average score ranging from 3.5 to 4 points indicates substantial knowledge, a score ranging from 2.5 to 3.49 points indicates some knowledge, a score ranging from 1.5 to 2.49 points indicates limited knowledge, and a score ranging from 1 to 1.49 points indicates no knowledge. On average, teachers who received SLOs training reported more knowledge about SLOs than those who did not on all nine aspects. To understand whether the differences were significant, I used independent t-tests. With 9 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .006 (i.e., .05/9). The alpha for the overall knowledge about SLOs was set to be .05. Based on the analysis results, teachers who received SLOs training reported statistically significantly more knowledge than those who did not

regarding their overall knowledge about SLOs with a small to medium effect size (p = .001, Cohen's d = 0.47), the purpose of SLOs with a small to medium effect size (p = .005, Cohen's d = 0.36), content to be include in SLOs with a small to medium effect size (p = .001, Cohen's d = 0.45), implementation of SLOs in the district with a medium effect size (p < .001, Cohen's d = 0.58), developing high quality of SLOs with a small to medium effect size (p = .001, Cohen's d = 0.45), and setting growth targets for SLOs with a small to medium effect size (p = .002, Cohen's d = 0.41) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the knowledge about the purpose SLOs is likely between 0.08 and 0.44 points higher for teachers who received training than for those who did not. The average score on the knowledge about the content to be included in SLOs is likely between 0.11 and 0.46 points higher for teachers who received training than for those who did not. The average score on the knowledge about implementation of SLOs in the district is likely between 0.25 and 0.65 points higher for teachers who received training than for those who did not. The average score on the knowledge about developing high quality SLOs is likely between 0.13 and 0.51 points higher for teachers who received training than for those who did not. The average score on the knowledge about setting growth targets for SLOs is likely between 0.11 and 0.48 points higher for teachers who received training than for those who did not. The average score based on the nine items is likely between 0.12 and 0.44 points higher for teachers who received training than for those who did not. These results indicate that teachers who received training have more knowledge than those who did not receive training regarding these aspects.

Table 4.12 *Teachers' Knowledge about SLOs Based on Training*

Knowledge about SLOs	No Training	Training	P-value	95% CI
Purpose of SLOs	3.11	3.37	.005	[0.08, 0.44]
Student groups to be included in SLOs	3.03	3.29	.008	[0.07, 0.45]
Content to be included in SLOs	3.09	3.38	.001	[0.11, 0.46]
Implementation of SLOs in the district	2.69	3.14	.000	[0.25, 0.65]
Developing high quality SLOs	2.90	3.23	.001	[0.13, 0.51]
Selecting appropriate assessments	3.02	3.25	.016	[0.04, 0.42]
Setting growth targets for SLOs	2.94	3.23	.002	[0.11, 0.48]
Instructional strategies to meet SLOs targets	3.05	3.28	.012	[0.05, 0.41]
Analyzing student assessment data	3.06	3.28	.018	[0.04, 0.40]
Overall	2.99	3.27	.001	[0.12, 0.44]

Notes: 63-64 teachers did not receive training and 313-315 received training.

Table 4.13 describes teachers' need for support in implementing SLOs based on training. I calculated the average score for each item on the support needed, and I also calculated the average score based on all items to understand the overall support needed. The questions are on a 3-point scale (1-Need No Support; 2-Need Some Support; 3-Need A Lot of Support). A higher score point indicates more support needed. An average score ranging from 2.5 to 3 points indicates need for a lot of support, a score ranging from 1.5 to 2.49 points indicates need for some support, a score ranging from 1 to 1.49 points indicates need for no support. On average, teachers who received SLOs training reported slightly less support needed to implement SLOs than those who did not receive SLOs training. To understand whether the differences of support needed were significant, I used independent t-tests. With 6 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .008 (i.e., .05/6). The alpha for the overall support needed was set to be .05. On average, teachers who received SLOs training reported significantly less overall support with a small to medium effect

size (p = .019, Cohen's d = 0.33), and less support needed in analyzing assessment data in SLOs with a small to medium effect size (p = .003, Cohen's d = 0.40) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the support in analyzing assessment data in SLOs is likely between 0.08 and 0.38 points lower for teachers who received training than for those who did not, indicating that teachers who received training need slightly less support in analyzing assessment data than those who did not receive training.

Table 4.13 *Teachers' Need for Support Based on Training*

Need for Support in Implementing SLOs	No Training	Training	P-value	95% CI
Understanding standards in SLOs	1.47	1.42	.531	[-0.20, 0.10]
Implementing standards in SLOs	1.55	1.45	.253	[-0.25, 0.07]
Understanding cognitive levels of standards	1.77	1.64	.104	[-0.30, 0.03]
Developing assessments for SLOs	1.84	1.65	.017	[-0.35, -0.04]
Setting growth targets for SLOs	1.92	1.72	.019	[-0.37, -0.03]
Analyzing assessment data in SLOs	1.90	1.67	.003	[-0.38, -0.08]
Overall	1.75	1.59	.019	[-0.28, -0.03]

Notes: 62-64 teachers did not receive training and 302-308 received training.

School participation in TAP was considered in the analysis of teachers' perceptions of SLOs. Table 4.14 describes teachers' views of the impact of SLOs based on their school participation in TAP. I calculated the average score for each item on the SLO impact, and I also calculated the average score based on all items to understand the overall impact of SLOs. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, teachers whose schools did not participate in TAP reported

more positive views of the impact of SLOs than the teachers whose schools participated in TAP. To understand whether the differences were significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of SLOs was set to be .05. Based on the analysis results, teachers did not have significantly different views of the impact of SLOs based on their schools' TAP participation.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.45 points) of the perceptions of SLOs between teachers whose schools participated in TAP and the teachers whose schools did not participate in TAP on each of the items, as well as the average of all items.

Table 4.14 *Teachers' Perceptions of SLOs Based on TAP*

Impact of SLOs	Non-TAP	TAP	P-value	95% CI
Evaluating teacher performance effectively	2.75	2.65	.421	[-0.34, 0.14]
Improving teachers' instructional practice	3.00	2.81	.105	[-0.41, 0.04]
Promoting student learning	3.12	2.89	.030	[-0.45, -0.02]
Informing teacher PD	2.89	2.84	.602	[-0.27, 0.16]
Overall	2.94	2.79	.140	[-0.34, 0.05]

Notes: 51-57 teachers did not participate in TAP and 308-317 participated in TAP.

Table 4.15 describes teachers' reported knowledge about SLOs based on their school participation in TAP. I calculated the average score for each item on the SLO knowledge, and I also calculated the average score based on all items to understand the overall knowledge about SLOs. The questions are on a 4-point scale (1-No Knowledge; 2- Limited Knowledge; 3- Some Knowledge; 4- Substantial Knowledge). A higher score point indicates more knowledge. An average score ranging from 3.5 to 4 points indicates

substantial knowledge, a score ranging from 2.5 to 3.49 points indicates some knowledge, a score ranging from 1.5 to 2.49 points indicates limited knowledge, and a score ranging from 1 to 1.49 points indicates no knowledge. On average, teachers whose schools participated in TAP reported more knowledge about SLOs than those whose schools did not participate in TAP on all nine aspects. To understand whether the differences were significant, I used independent t-tests. With 9 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .006 (i.e., .05/9). The alpha for the overall knowledge about SLOs was set to be .05. On average, teachers whose schools participated in TAP reported statistically significantly more knowledge than those whose schools did not participate in TAP regarding their overall knowledge of SLOs with a small to medium effect size (p = .008, Cohen's d = 0.37), and implementing SLOs in the district with a small to medium effect size (p = .003, Cohen's d = 0.41) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the knowledge about implementation of SLOs in the district is likely between 0.11 and 0.53 points higher for teachers whose schools participated in TAP than for those whose schools did not participate in TAP, indicating that teachers who whose schools participated in TAP have slightly more knowledge about implementation of SLOs in the district than those whose schools did not participate in TAP. The average score based on the nine items on the knowledge about SLOs is likely between 0.06 and 0.39 points higher for teachers whose schools participated in TAP than for those whose schools did not participate in TAP, indicating that teachers who whose

schools participated in TAP have slightly more knowledge about SLOs than those whose schools did not participate in TAP.

Table 4.15 *Teachers' Knowledge about SLOs Based on TAP*

Knowledge about SLOs	Non-TAP	TAP	P-value	95% CI
Purpose of SLOs	3.17	3.36	.048	[0.00, 0.38]
Student groups to be included in SLOs	3.07	3.28	.032	[0.02, 0.41]
Content to be included in SLOs	3.12	3.38	.006	[0.08, 0.45]
Implementation of SLOs in the district	2.80	3.12	.003	[0.11, 0.53]
Developing high quality SLOs	2.98	3.22	.020	[0.04, 0.43]
Selecting appropriate assessments	3.00	3.25	.010	[0.06, 0.45]
Setting growth targets for SLOs	3.02	3.23	.035	[0.02, 0.41]
Instructional strategies to meet SLOs targets	3.13	3.27	.140	[-0.05, 0.32]
Analyzing student assessment data in SLOs	3.10	3.28	.053	[-0.00, 0.37]
Overall	3.04	3.27	.008	[0.06, 0.39]

Notes: 59-60 teachers did not participate in TAP and 321-323 participated in TAP.

Table 4.16 describes teachers' need for support in implementing SLOs based on their school participation in TAP. I calculated the average score for each item on the support needed, and I also calculated the average score based on all items to understand the overall support needed. The questions are on a 3-point scale (1-Need No Support; 2-Need Some Support; 3-Need A Lot of Support). A higher score point indicates more support needed. An average score ranging from 2.5 to 3 points indicates need for a lot of support, a score ranging from 1.5 to 2.49 points indicates need for some support, a score ranging from 1 to 1.49 points indicates need for no support. On average, teachers whose schools participated in TAP reported less support needed to implement SLOs than those whose schools did not participate in TAP. To understand whether the differences of support needed were significant, I used independent t-tests. With 6 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .008 (i.e., .05/6). The alpha for the overall support needed was set to be .05. On average, there was a statistically significant difference of the overall support needed

between teachers whose schools participated in TAP and those whose schools did not participate in TAP with a small to medium effect size (p = .034, Cohen's d = 0.30) (Cohen, 1988). There were no statistically significant differences of support needed between teachers whose schools participated in TAP and those whose schools did not participate in TAP regarding the six aspects.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score based on the six items on support needed in implementing SLOs is likely between 0.01 and 0.28 points lower for teachers whose schools participated in TAP than those whose schools did not participate in TAP, indicating that teachers whose schools participated in TAP need slightly less support than those whose schools did not participate in TAP. In addition, the confidence intervals show that I am confident that there is very little difference (at most, 0.39 points) of support needed in implementing SLOs between teachers whose schools participated in TAP and the teachers whose schools did not participate in TAP on each of the items.

Table 4.16 *Teachers' Need for Support Based on TAP*

Need for Support in Implementing SLOs	Non-TAP	TAP	P-value	95% CI
Understanding standards in SLOs	1.53	1.42	.177	[-0.27, 0.05]
Implementing standards in SLOs	1.56	1.46	.218	[-0.28, 0.06]
Understanding cognitive levels of standards	1.84	1.63	.015	[-0.39, -0.04]
Developing assessments for SLOs	1.79	1.66	.146	[-0.31, 0.05]
Setting growth targets for SLOs	1.85	1.73	.165	[-0.31, 0.05]
Analyzing assessment data in SLOs	1.87	1.68	.028	[-0.37, -0.02]
Overall	1.74	1.60	.034	[-0.28, -0.01]

Notes: 53-56 teachers did not participate in TAP and 317-319 participated in TAP.

4.1.3 Educators' Perceptions of Classroom Observations

Educators' views of classroom observations were analyzed using both percentages and means (Table 4.17). Between 77.4% and 81.1% of the educators agreed

or strongly agreed that using classroom observations evaluates teacher performance effectively, improves teachers' instructional practice, promotes student learning, and informs teacher professional development. Overall, all administrators (100%) and between 76.0% and 79.9% of the teachers agreed or strongly agreed on the four aspects of the impact of classroom observations.

I calculated the average score for each item on the impact of classroom observations, and I also calculated the average score based on all items to understand the overall impact of classroom observations. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, educators agreed on the impact of classroom observations (M = 2.88), administrators reported more positive views (M = 3.42) than teachers (M = 2.84). Independent t-tests were conducted to help understand whether the views of teachers and administrator are significantly different. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of classroom observations was set to be .05. On average, administrators had significantly more positive views than teachers regarding the overall impact of classroom observations with a large effect size ($p < .001$, Cohen's $d = 1.11$), and the impact of classroom observations on evaluating teacher performance effectively with a large effect size ($p < .001$, Cohen's $d = 1.04$), improving teachers' instructional practice with a large effect size ($p = .002$,

Cohen's $d = 0.91$), promoting student learning with a large effect size ($p = .001$, Cohen's $d = 0.98$), and informing teachers' professional development with a large effect size ($p = .001$, Cohen's $d = 0.98$) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the item indicating belief in whether classroom observations can be used to evaluate teacher performance effectively is likely between 0.28 and 0.97 points higher for administrators than for teachers. The average score on the item indicating belief in whether classroom observations can be used to improve teachers' instructional practice is likely between 0.20 and 0.88 points higher for administrators than for teachers. The average score on the item indicating belief in whether classroom observations can be used to promote student learning is likely between 0.25 and 0.92 points higher for administrators than for teachers. The average score on the item indicating belief in whether classroom observations can be used to inform teacher professional development is likely between 0.25 and 0.89 points higher for administrators than for teachers. The average score based on the four items on the impact of classroom observations is likely between 0.27 and 0.88 points higher for administrators than for teachers. These results suggest that administrators hold higher beliefs in these statements about the impact of classroom observations than teachers.

To check the analysis, I also used Mann-Whitney Tests considering the sample size of administrators is small. The results are consistent, and there was a statistically significant difference between teachers' and administrators' views of the overall impact of classroom observations ($Z = -4.29$, $p < .001$) and all four aspects.

Table 4.17 *Educators' Views of Classroom Observations*

Impact of Classroom Observations	Percentage			Mean				
	All	Teachers	Administrators	All	Teacher	Administrator	*P*-value	95% CI
Evaluating teacher performance effectively	77.4	76.0	100.0	2.85	2.81	3.44	<.001	[0.28, 0.97]
Improving teachers' instructional practice	80.4	79.3	100.0	2.87	2.84	3.38	.002	[0.20, 0.88]
Promoting student learning	79.2	78.0	100.0	2.89	2.85	3.44	.001	[0.25, 0.92]
Informing teacher PD	81.1	79.9	100.0	2.90	2.87	3.44	.001	[0.25, 0.89]
Overall	79.5	78.3	100.0	2.88	2.84	3.42	<.001	[0.27, 0.88]

Notes: Respondents include 263-265 teachers and 16 administrators.

4.1.4 Teachers' Perceptions of Classroom Observations Based on Variables

Classroom observations are a major mode of teacher evaluation. This section focused on the associations of teachers' perceptions of the impact of classroom observations and their educational background. Teachers' academic degrees, years of experience in education, classroom observations training received, and school participation in the TAP were considered in the analysis.

Teachers' highest academic degree was analyzed (Table 4.18). I calculated the average score for each item on the impact of classroom observations, and I also calculated the average score based on all items to understand the overall impact of classroom observations. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, teachers who had a master's degree or above reported slightly more positive views of the impact of classroom observations than those who had a bachelor's degree or below on three out of the four aspects. To understand whether the differences were significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of classroom observations was set to be .05. On average, there were no statistically significant differences of the views of teachers who had a master's degree and those who had a bachelor's degree.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.23 points) of the perceptions of classroom observations between teachers with a bachelor's degree or below and those with a master's degree or above on each of the items, as well as the average of all items.

Table 4.18 *Teachers' Views of Classroom Observations Based on Degree*

Impact of Classroom Observations	Bachelor or Below	Master or Above	P-value	95% CI
Evaluating teacher performance effectively	2.79	2.83	.714	[-0.14, 0.20]
Improving teachers' instructional practice	2.84	2.83	.918	[-0.18, 0.16]
Promoting student learning	2.81	2.88	.468	[-0.11, 0.23]
Informing teacher PD	2.85	2.88	.694	[-0.13, 0.19]
Overall	2.83	2.85	.715	[-0.12, 0.18]

Notes: 101-103 teachers had a degree of bachelor or below, and 161-166 had a master's or above.

Teachers' years of experience was considered in the analysis (Table 4.19). I calculated the average score for each item on the impact of classroom observations, and I also calculated the average score based on all items to understand the overall impact of classroom observations. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, career teachers who had more than three years of experience in education reported slightly more positive views of the impact of classroom observations than early career teachers who had three or fewer years of experience on three out of the four aspects. To understand whether the differences were significant, I

used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of classroom observations was set to be .05. On average, career teachers and early career teachers did not have significantly different views of the impact of classroom observations.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.28 points) of the perceptions of classroom observations between early career teachers and career teachers on each of the items, as well as the average of all items.

Table 4.19 *Teachers' Views of Classroom Observations Based on Experience*

Impact of Classroom Observations	Early Career (0-3)	Career (3+)	P-value	95% CI
Evaluating teacher performance effectively	2.78	2.81	.771	[-0.20, 0.26]
Improving teachers' instructional practice	2.85	2.83	.865	[-0.25, 0.21]
Promoting student learning	2.80	2.86	.629	[-0.17, 0.28]
Informing teacher PD	2.85	2.87	.856	[-0.20, 0.24]
Overall	2.82	2.84	.831	[-0.18, 0.23]

Notes: 40-41 teachers had 0-3 years of experience, 221-227 had 4 or more years of experience.

Training was considered in the analysis of teachers' views of classroom observations (Table 4.20). I calculated the average score for each item on the impact of classroom observations, and I also calculated the average score based on all items to understand the overall impact of classroom observations. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score

ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, teachers who received classroom observation training reported slightly more positive views of the impact of classroom observations than those who did not on all four aspects. To understand whether the differences were significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of classroom observations was set to be .05. On average, teachers who received training and those who did not receive the training did not have significantly different views of the impact of classroom observations.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.27 points) of the perceptions of classroom observations between teachers who received training and those who did not receive training on each of the items, as well as the average of all items.

Table 4.20 *Teachers' Views of Classroom Observations Based on Training*

Impact of Classroom Observations	No Training	Training	P-value	95% CI
Evaluating teacher performance effectively	2.79	2.83	.627	[-0.13, 0.22]
Improving teachers' instructional practice	2.78	2.88	.246	[-0.07, 0.27]
Promoting student learning	2.81	2.88	.477	[-0.11, 0.24]
Informing teacher PD	2.84	2.89	.496	[-0.11, 0.22]
Overall	2.81	2.87	.422	[-0.09, 0.22]

Notes: 94-98 teachers did not receive trainings and 151-155 received trainings.

School TAP participation was analyzed (Table 4.21). I calculated the average score for each item on the impact of classroom observations, and I also calculated the

average score based on all items to understand the overall impact of classroom observations. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, teachers whose schools participated in TAP reported more positive views than the teachers whose schools did not participate in TAP on three out of the four aspects of the impact of classroom observations. To understand whether the differences were significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of classroom observations was set to be .05. On average, teachers' views of the classroom observations were similar regardless of their schools' TAP participation.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.52 points) of teachers' perceptions of classroom observations due to their schools' TAP participation.

Table 4.21 *Teachers' Views of Classroom Observations Based on TAP*

Impact of Classroom Observations	Non-TAP	TAP	P-value	95% CI
Evaluating teacher performance effectively	2.57	2.85	.026	[0.03, 0.52]
Improving teachers' instructional practice	2.84	2.84	.961	[-0.24, 0.23]
Promoting student learning	2.81	2.86	.664	[-0.19, 0.29]
Informing teacher PD	2.89	2.87	.803	[-0.25, 0.19]
Overall	2.80	2.85	.625	[-0.16, 0.26]

Notes: 35-39 teachers did not participate in TAP and 225-229 participated in TAP.

4.1.5 Summaries of Educator' Views of SLOs and Classroom Observations

To better understand educators' views of the SLOs and classroom observations in the evaluation of teacher effectiveness, I compared their views (Table 4.22). Educators appeared to agree that both SLOs and classroom observations had positive impact on all four aspects with means larger than 2.5. Educators reported slightly more positive views of classroom observation (M = 2.88) in comparison with their views of SLOs (M = 2.83). Educators reported more positive views of the impact of classroom observations on three out of the four aspects in comparison with their views of the impact of SLOs. In particular, educators demonstrated much higher agreement that using classroom observation evaluates teacher performance effectively. In comparison with teachers who had an SLOs mean of 2.82 and classroom observations mean of 2.84, administrators reported more positive views of the impact of both SLOs (M = 3.19) and classroom observations (M = 3.42). Administrators demonstrated more positive views of the impact of classroom observations than the impact of SLOs. Teachers demonstrated slightly more positive views of the impact of classroom observations than the impact of SLOs.

Table 4.22 *Comparing Educators' Views of SLOs and Classroom Observations*

Aspects of Impact	SLOs		Classroom Observations	
	Teachers	Administrator	Teachers	Administrator
Evaluating teacher performance effectively	2.66	3.05	2.81	3.44
Improving teachers' instructional practice	2.84	3.25	2.84	3.37
Promoting student learning	2.92	3.35	2.85	3.44
Informing teacher PD	2.85	3.10	2.87	3.44
Overall	2.82	3.19	2.84	3.42

Teachers' perceptions of the impact of SLOs and classroom observations were summarized and compared based on their academic degrees, years of experience, training, and TAP participation (Table 4.23). The alpha was set to be .05. In comparison with teachers who had a bachelor's degree or below, those with a master's degree or above reported slightly more positive views of the impact of SLOs and the impact of classroom observations, more knowledge about SLOs, and slightly less support needed to implement SLOs.

In comparison with early career teachers, career teachers reported statistically significantly more positive views of the impact of SLOs with a small to medium effect size (p = .001, Cohen's d = 0.45), more positive views of classroom observations, significantly more knowledge about SLOs with a small to medium effect size (p = .006, Cohen's d = 0.38), and slightly more support needed to implement SLOs. In comparison with teachers who did not receive training, those who received training reported statistically significantly more positive views of SLOs with a small to medium effect size (p = .009, Cohen's d = 0.34), more positive views of classroom observations, significantly more knowledge about SLOs with a small to medium effect size (p = .001, Cohen's d = 0.47), and significantly less support needed to implement SLOs with a small to medium effect size (p = .019, Cohen's d = 0.33).

Teachers from the TAP schools reported less positive views of SLOs, more positive views of classroom observations, significantly more knowledge about SLOs with a small to medium effect size (p = .008, Cohen's d = 0.37), and significantly less support needed to implement SLOs with a small to medium effect size (p = .034, Cohen's d = 0.30) (Cohen, 1988).

Table 4.23 *Teachers' Views Based on Education Background*

Variables		SLOs			Classroom Observations
		Impact	Knowledge	Support	Impact
Degree	Bachelors or Below	2.79	3.16	1.64	2.83
	Masters or Above	2.83	3.27	1.61	2.85
Experience in Education	0-3 Years (Early Career)	2.53*	3.02*	1.61	2.82
	3+ Years (Career)	2.86*	3.26*	1.62	2.84
Trainings	No Trainings	2.61*	2.99*	1.75*	2.81
	Trainings	2.86*	3.27*	1.59*	2.87
TAP Participation	Not Participated	2.94	3.04*	1.74*	2.80
	Participated	2.79	3.27*	1.60*	2.85
Overall		2.82	3.23	1.62	2.84

*Note: * indicates statistically significant differences.*

To investigate whether teachers' perceptions of SLOs and their perceptions of classroom observations have some relationships, I conducted a correlation analyses. Considering the ordinal feature of the data, I used Spearman's rho correlation analysis. According to Table 4.24, teachers' perceptions of the four aspects of SLOs and the classroom observations had statistically significant positive associations, and the relationships were moderate.

I also explored the relationship of teachers' overall perceptions of SLOs and their overall perceptions of classroom observations. I calculated the means of the four aspects of SLOs and the four aspects of classroom observations. Considering the data are continuous, I used Pearson's correlation analysis. The correlation coefficient was calculated to be .45 which was statistically significant. It suggests that teachers who had more positive perceptions of classroom observations tended to have more positive perceptions of SLOs.

Table 4.24 *Correlations of Teachers' Views of SLOs and Classroom Observations*

Aspects of Impact	Correlations	p-value
Evaluating teacher performance effectively	.33	<.001
Improving teachers' instructional practice	.41	<.001
Promoting student learning	.39	<.001
Informing teacher PD	.42	<.001
Overall	.45	<.001

To further explore the factors that could predict teachers' perceptions of SLOs, I used multiple regression analysis. Teachers' overall perception of SLOs was the dependent variable, and teachers' perceptions of classroom observations, knowledge about SLOs, years of experience in education, and training of SLOs were the independent variables. Results from the multiple regression analysis suggest a statistically significant association between teachers' overall perceptions of SLOs and the independent variables, F (36.95, 4) = 22.91, $p < .01$.

Teachers' perceptions of classroom observations, training of SLOs, and years of experience in education were identified as the significant predictors of their perceptions of SLOs. Overall, teachers who had more positive perceptions of classroom observations tended to have more positive perceptions of SLOs (B = 0.52, $p < .001$). Teachers who received training of SLOs tended to have more positive perceptions of SLOs (B = 0.30, $p = .006$). Career teachers who had more than three years of experience in education tended to have more positive perceptions of SLOs (B = 0.33, $p = .003$). Based on the analysis results, teachers' knowledge about SLOs was not a significant predictor of their perceptions of SLOs. The effect size adjusted R^2 of .252 indicates that 25.2% of the variation in teachers' perceptions of SLOs could be explained by the predictors.

Table 4.25 *Multiple Regression Analysis Results*

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	.300	.316		0.947	.345
OR Impact	.517	.066	.429	7.832	.000
SLOs Knowledge	.071	.067	.059	1.054	.293
SLOs trainings	.295	.107	.151	2.767	.006
Years of Experience	.329	.111	.162	2.962	.003

4.1.6 The Highlights of the Study 1 Findings

Based on the results from various analysis in Study 1, I present the following highlights of the findings:

- Most educators agreed on the impact of SLOs and classroom observations on evaluating teacher performance effectively. They reported some to substantial knowledge about SLOs, and they reported needing some support in implementing SLOs.
- In comparison with teachers, administrators had more positive views of the overall impact of SLOs and impact of SLOs on evaluating teacher performance effectively. However, teachers and administrators were similar in their knowledge about SLOs and support needed to implement SLOs.
- Teachers and administrators within the same school often differed on their views of the impact of SLOs, their levels of knowledge about SLOs, and their levels of support needed in implementing SLOs.
- Differences in teachers' academic degrees were not associated with differences in their views of SLOs, knowledge about SLOs, or support needed to implement SLOs.

- In comparison with early career teachers, career teachers reported more positive views of the overall impact of SLOs, SLOs' evaluating teacher performance effectively, improving teachers' instructional practices, and promoting student learning. Career teachers reported more overall knowledge about SLOs, the purposes of SLOs, developing high quality of SLOs, and setting growth targets for SLOs. Career teachers and early career teachers were similar in the levels of support needed to implement SLOs.
- In comparison with teachers who did not receive training of SLOs, those who received training had more positive views of the overall impact of SLOs and the impact of SLOs on teachers' instructional practices. Teachers who received SLOs training reported more overall knowledge about SLOs, the purpose of SLOs, content to be include in SLOs, implementation of SLOs in the district, developing high quality of SLOs, and setting growth targets for SLOs. They reported less overall support, and less support in analyzing assessment data in SLOs.
- Teachers' views of the impact of SLOs were similar regardless of the TAP participation status of their schools. In comparison with the non-TAP school teachers, the TAP school teachers reported more overall knowledge about SLOs and knowledge about implementing SLOs in the district. They reported less support needed to implement SLOs.
- In comparison with teachers, administrators had more positive views of the overall impact of classroom observations, and the impact of classroom observations on evaluating teacher performance effectively, improving teachers' instructional

practice, promoting student learning, and informing teachers' professional development.

- Differences in teachers' degrees, years of experience, training, and school TAP participation were not associated with differences in their views of the impact of classroom observations.
- Teachers' perceptions of classroom observations, training of SLOs, and years of experience in education predicted their perceptions of SLOs, and 25.2% of the variation in teachers' perceptions of SLOs could be explained by the predictors.

4.2 STUDY 2 RESULTS

Based on the survey responses of 289 educators from South Carolina, I present the results. First, I present teachers' and administrators' views of SLOs, their knowledge about SLOs, and the support needed to implement SLOs. I compared teachers' and administrators' views of the impact of SLOs, knowledge about SLOs, and support needed in implementing SLOs. I used both percentages and means in reporting teachers' and administrators' views. Second, I present teachers' perceptions of SLOs based on their academic degrees, years of experience in education, the SLOs experience, and grade levels taught. I used descriptive statistics and inferential statistics in analyzing teachers views based on their academic degrees, years of experience in education, the SLOs experience, and grade levels taught. Third, I present educators' views of classroom observations, compare administrators' and teachers' views of classroom observations. I also present teachers' views based on their academic degrees, years of experience in education, SLOs experience, and grade levels taught. In addition, I summarize educators' views of the impact of SLOs and classroom observations. Finally, I present a multiple

regression analysis to examine the prediction of teachers' views of SLOs. In the model, teachers' view of SLOs is the dependent variable, and teachers' views of classroom observations, academic degree, years of experience in education, SLOs experience, and grade levels taught are predictors.

4.2.1 Educators' Perceptions of SLOs

I analyzed educators' perceptions of SLOs using both percentages and means (Table 4.26). About two-thirds of the educators agreed or strongly agreed that using SLOs evaluates teacher performance effectively. About 38% of the educators agreed or strongly agreed that using SLOs improves teachers' instructional practice, promotes student learning, and informs teacher professional development. About 28% of the educators agreed or strongly agreed that using SLOs evaluates teacher performance effectively. Comparatively, administrators reported more positive views of the impact of SLOs in comparison with teachers. Between 43.8% and 56.3% of the administrators and between 25.1% and 36.8% of the teachers agreed or strongly agreed on the four aspects of the impact of SLOs.

I calculated the average score for each item on the SLO impact, and I also calculated the average score based on all items to understand the overall impact of SLOs. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement.

On average, educators disagreed on the impact of SLOs (M = 2.19), administrators reported more positive views (M = 2.35) than teachers (M = 2.16). In addition, administrators were slightly more positive than teachers on all four aspects of the impact of SLOs. Independent t-tests were conducted to help understand whether the views of teachers and administrator are significantly different. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of SLOs was set to be .05. On average, there was no statistically significant difference between teachers' views and administrators' views of the overall impact of SLOs or any of the four aspects of SLOs.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.68 points) of the average scores on the items about the perceptions of SLOs between administrators and teachers on each of the items, as well as the average of all items.

To check the analysis, I also used Mann-Whitney Tests considering the sample size of administrators is small. The results are consistent with those using independent t-tests, and there is no statistically significant difference between teachers' views and administrators' views of the overall impact of SLOs ($Z = -1.29$, $p = .198$) or any of the four aspects. The consistent findings further suggested that teachers and administrators held very similar views of the impact of SLOs, although administrators seemed to be slightly more positive about the impact of SLOs in comparison with teachers.

Table 4.26 *Educators' Perceptions of SLOs*

Impact of SLOs	Percentage			Mean				
	All	Teacher	Administrator	All	Teacher	Administrator	*P*-value	95% CI
Evaluating teacher performance effectively	27.6	25.1	43.8	2.07	2.03	2.25	.296	[-0.20, 0.65]
Improving teachers' instructional practice	38.9	36.4	56.3	2.24	2.21	2.44	.306	[-0.21, 0.68]
Promoting student learning	38.4	36.8	46.7	2.26	2.23	2.33	.651	[-0.36, 0.57]
Informing teacher PD	38.7	36.7	50.0	2.21	2.17	2.38	.354	[-0.23, 0.63]
Overall	35.9	33.8	49.2	2.19	2.16	2.35	.323	[-0.19, 0.58]

Notes: Respondents include 239-242 teachers and 15-16 administrators.

I analyzed educators' knowledge about SLOs using both percentages and means (Table 4.27). Overall, large percentages of teachers (81.4%-91.7%) and administrators (93.8%-100.0%) reported having some or substantial knowledge about different aspects of SLOs. In comparison with teachers, larger percentages of administrators reported having some or substantial knowledge on all nine aspects of SLOs.

I calculated the average score for each item on the SLO knowledge, and I also calculated the average score based on all items to understand the overall knowledge about SLOs. The questions are on a 4-point scale (1-No Knowledge; 2- Limited Knowledge; 3-Some Knowledge; 4- Substantial Knowledge). A higher score point indicates more knowledge. An average score ranging from 3.5 to 4 points indicates substantial knowledge, a score ranging from 2.5 to 3.49 points indicates some knowledge, a score ranging from 1.5 to 2.49 points indicates limited knowledge, and a score ranging from 1 to 1.49 points indicates no knowledge.

On average, educators reported having some knowledge about SLOs (M = 3.31), and administrators reported more knowledge (M = 3.65) than teachers (M = 3.30). To understand whether teachers and administrators reported statistically significantly different levels of knowledge about SLOs, I used independent t-tests. With 9 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .006 (i.e., .05/9). The alpha for the overall knowledge about SLOs was set to be .05. On average, administrators had significantly more knowledge than teachers regarding their overall knowledge with a medium to large effect size (p = .019, Cohen's d = 0.64), student groups to be included in SLOs with a medium to large effect size (p = .003, Cohen's d = 0.69), content to be included in SLOs with a medium to

large effect size (p = .003, Cohen's d = 0.66), and analyzing student assessment data with medium to large effect size (p = .001, Cohen's d = 0.77) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the knowledge about the student groups to be included in SLOs is likely between 0.16 and 0.66 points higher for administrators than for teachers. The average score on the knowledge about the content to be included in SLOs is likely between 0.15 and 0.61 points higher for administrators than for teachers. The average score on the knowledge about analyzing student assessment data is likely between 0.22 and 0.67 points higher for administrators than for teachers. The average score based on the nine items on the knowledge about SLOs is likely between 0.15 and 0.61 points higher for administrators than for teachers. The average score on the knowledge about analyzing student assessment data is likely between 0.06 and 0.64 points higher for administrators than for teachers. These results indicate that administrators have more knowledge than teachers regarding the student groups to be included in SLOs, the content to be included in SLOs, analyzing student assessment data, and average score of the knowledge about SLOs.

To check the analysis, I also used Mann-Whitney Tests considering the sample size of administrators is small. The results were consistent with those by using the independent t-tests, administrators reported statistically significantly more overall knowledge than teachers (Z = -2.58, p = .010). The consistent findings further suggested that administrators seemed to be more knowledge than teachers regarding the Implementation of SLOs.

Table 4.27 *Educators' Knowledge about SLOs*

Knowledge about SLOs	Percentage			Mean				
	All	Teacher	Administrator	All	Teacher	Administrator	*P*-value	95% CI
Purpose of SLOs	90.9	91.7	93.8	3.37	3.35	3.75	.017	[0.08, 0.71]
Student groups to be included in SLOs	89.3	89.7	100.0	3.35	3.34	3.75	.003	[0.16, 0.66]
Content to be included in SLOs	92.1	90.9	100.0	3.46	3.44	3.81	.003	[0.15, 0.61]
Implementation of SLOs in the district	84.8	84.7	93.8	3.25	3.23	3.69	.018	[0.08, 0.84]
Developing high quality SLOs	81.6	82.3	93.8	3.16	3.16	3.50	.081	[-0.04, 0.71]
Selecting appropriate assessments	85.8	86.5	93.8	3.26	3.25	3.50	.201	[-0.13, 0.62]
Setting growth targets for SLOs	80.9	81.4	93.8	3.20	3.20	3.50	.140	[-0.10, 0.69]
Instructional strategies to meet SLOs targets	88.5	89.3	93.8	3.33	3.33	3.50	.353	[-0.19, 0.52]
Analyzing student assessment data	87.9	89.0	100.0	3.38	3.37	3.81	.001	[0.22, 0.67]
Overall	86.9	87.3	95.9	3.31	3.30	3.65	.019	[0.06, 0.64]

Notes: Respondents include 242-244 teachers and 16 administrators.

I analyzed educators' support needed to implement SLOs using both percentages and means (Table 4.28). Overall, some teachers (33.7%-45.1%) and administrators (25.0%-50.0%) reported needing some or a lot of support in implementing different aspects of SLOs, and very similar levels of the overall support needed was reported for teachers and administrators. Larger percentages of administrators reported needing some or a lot of support in four out of the six items in comparison with teachers, and these aspects included implementing standards, in SLOs, understanding the cognitive levels of standards in SLOs, developing assessment for SLOs, and setting growth targets for SLOs. Larger percentages of teachers reported needing some or a lot of support in understanding standards for SLOs and analyzing assessment data in comparison with administrators. Comparatively, larger percentages of teachers and administrators reported needing support in setting growth targets and understanding cognitive levels of standards. Small percentages of teachers and administrators reported needing support in understanding standards and analyzing assessment data.

I calculated the average score for each item on the support needed, and I also calculated the average score based on all items to understand the overall support needed. The questions are on a 3-point scale (1-Need No Support; 2-Need Some Support; 3-Need A Lot of Support). A higher score point indicates more support needed. An average score ranging from 2.5 to 3 points indicates need for a lot of support, a score ranging from 1.5 to 2.49 points indicates need for some support, a score ranging from 1 to 1.49 points indicates need for no support.

On average, educators reported needing little support in implementing SLOs (M = 1.46), and administrators and teachers reported similar levels of the overall support

needed to implement SLOs. Teachers and administrators also reported similar levels of support needed in implementing standards and understanding the cognitive levels of standards. Administrators seemed to need more support in setting growth targets and developing assessment for SLOs in comparison with teachers. Teachers seemed to need more support in understanding standards and analyzing assessment data in comparison with administrators. To understand whether teachers and administrators reported statistically significantly different support needed in implementing SLOs, I used independent t-tests. With 6 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .008 (i.e., .05/6). The alpha for the overall support needed to implement SLOs was set to be .05. On average, there were no statistically significant differences between teachers' and administrators' overall support needed (p = .992) or any of the six aspects of SLOs.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.45 points) of the average support needed between administrators and teachers on each of the items, as well as the average of all items.

To check the analysis, I also used Mann-Whitney Tests considering the sample size of administrators is small. The results are consistent with those by the independent t-tests, and there were no statistically significant differences between teachers' and administrators' support needed to implement SLOs. The consistent findings further suggested that teachers and administrators needed similar levels of support in implementing SLOs.

Table 4.28 *Educators' Need for Support in Implementing SLOs*

Need for Support in Implementing SLOs	Percentage			Mean				
	All	Teacher	Administrator	All	Teacher	Administrator	*P*-value	95% CI
Understanding standards in SLOs	34.7	33.7	25.0	1.39	1.38	1.25	.274	[-0.38, 0.11]
Implementing standards in SLOs	35.5	35.0	37.5	1.39	1.39	1.38	.934	[-0.29, 0.27]
Understanding cognitive levels of standards	46.7	43.8	50.0	1.53	1.50	1.50	.979	[-0.32, 0.31]
Developing assessments for SLOs	38.5	37.0	43.8	1.45	1.42	1.56	.373	[-0.17, 0.45]
Setting growth targets for SLOs	47.1	45.1	50.0	1.55	1.52	1.63	.521	[-0.22, 0.43]
Analyzing assessment data in SLOs	37.9	36.0	25.1	1.44	1.41	1.31	.508	[-0.40, 0.20]
Overall	40.1	38.4	38.6	1.46	1.44	1.44	.992	[-0.26, 0.25]

Notes: Respondents include 242-244 teachers and 16 administrators.

4.2.2 Teachers' Perceptions of SLOs Based on Variables

This section focuses on teachers' views of the impact of SLOs, their knowledge about SLOs, and the support needed to implement SLOs. Specifically, I examined whether there was an association between teachers' perceptions of SLOs and their academic degrees, years of experience in education, SLOs experience, and grade levels taught.

I calculated the average score for each item on the SLO impact, and I also calculated the average score based on all items to understand the overall impact of SLOs (Table 4.29). The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, teachers reported negative perceptions of the impact of SLOs on all four aspects with means being smaller than 2.5. Teachers who had a master's degree or above reported less positive views of the impact of SLOs on all four aspects than those who had a bachelor's degree or below. To understand whether the differences were significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of SLOs was set to be .05. On average, there were no statistically significant differences between teachers who had a master's degree or above and those who had a bachelor's degree or below.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is little difference (at most, 0.47 points) of the perceptions of SLOs between teachers with a bachelor's degree or below and those with a master's degree or above on each of the items, as well as the average of all items.

Table 4.29 *Teachers' Perceptions of SLOs Based on Degree*

Impact of SLOs	Bachelor or Below	Master or Above	P-value	95% CI
Evaluating teacher performance effectively	2.13	1.95	.136	[-0.05, 0.40]
Improving teachers' instructional practice	2.35	2.11	.055	[-0.01, 0.47]
Promoting student learning	2.33	2.16	.156	[-0.07, 0.41]
Informing teacher PD	2.21	2.14	.544	[-0.16, 0.30]
Overall	2.25	2.09	.115	[-0.04, 0.34]

Notes: 80-82 teachers had a degree of bachelor or below, and 151-153 had a master's or above.

Table 4.30 describes teachers' reported knowledge about SLOs based on their degrees. I calculated the average score for each item on the SLO knowledge, and I also calculated the average score based on all items to understand the overall knowledge about SLOs. The questions are on a 4-point scale (1-No Knowledge; 2- Limited Knowledge; 3-Some Knowledge; 4- Substantial Knowledge). A higher score point indicates more knowledge. An average score ranging from 3.5 to 4 points indicates substantial knowledge, a score ranging from 2.5 to 3.49 points indicates some knowledge, a score ranging from 1.5 to 2.49 points indicates limited knowledge, and a score ranging from 1 to 1.49 points indicates no knowledge. On average, teachers reported having some knowledge about SLOs, with means ranging from 3.0 to 3.49. Teachers who had a master's degree or above reported more knowledge about SLOs on all nine aspects than those who had a bachelor's degree or below. To understand whether these differences

were statistically significant based on teachers' degrees, I used independent t-tests. With 9 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .006 (i.e., .05/9). The alpha for the overall knowledge about SLOs was set to be .05. On average, teachers who had a master's degree or above reported statistically significantly more knowledge regarding their overall knowledge about SLOs with a small to medium effect size (p = .011, Cohen's d = 0.35), implementation of SLOs in the district with a small to medium effect size (p = .001, Cohen's d = 0.44), and instructional strategies to meet SLOs targets with a medium effect size (p < .001, Cohen's d = 0.49) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the knowledge about implementation of SLOs in the district is likely between 0.13 and 0.53 points higher for teachers who had a master's degree or above than those who had a bachelor's degree or below. The average score on the knowledge about instructional strategies to meet SLOs targets is likely between 0.15 and 0.51 points higher for teachers who had a master's degree or above than those who had a bachelor's degree or below. The average score on the knowledge about SLOs is likely between 0.05 and 0.35 points higher for teachers who had a master's degree or above than those who had a bachelor's degree or below. These results indicate that teachers who had a master's degree have more knowledge than teachers who had a bachelor's degree or below regarding the implementation of SLOs in the district, instructional strategies to meet SLOs targets, and average score of the knowledge about SLOs.

Table 4.30 *Teachers' Knowledge about SLOs Based on Degree*

Knowledge about SLOs	Bachelor or Below	Master or Above	P-value	95% CI
Purpose of SLOs	3.21	3.45	.006	[0.07, 0.41]
Student groups to be included in SLOs	3.21	3.43	.023	[0.03, 0.40]
Content to be included in SLOs	3.40	3.49	.350	[-0.09, 0.26]
Implementation of SLOs in the district	3.02	3.36	.001	[0.13, 0.53]
Developing high quality SLOs	3.07	3.23	.110	[-0.04, 0.36]
Selecting appropriate assessments	3.16	3.30	.174	[-0.06, 0.34]
Setting growth targets for SLOs	3.10	3.26	.126	[-0.05, 0.38]
Instructional strategies to meet SLOs targets	3.13	3.46	.000	[0.15, 0.51]
Analyzing student assessment data	3.27	3.43	.086	[-0.02, 0.35]
Overall	3.18	3.38	.011	[0.05, 0.35]

Notes: 80-82 teachers had a degree of bachelor or below, and 154-155 had a master's or above.

Table 4.31 describes teachers' need for support in implementing SLOs based on their degrees. I calculated the average score for each item on the support needed, and I also calculated the average score based on all items to understand the overall support needed. The questions are on a 3-point scale (1-Need No Support; 2-Need Some Support; 3-Need A Lot of Support). A higher score point indicates more support needed. An average score ranging from 2.5 to 3 points indicates need for a lot of support, a score ranging from 1.5 to 2.49 points indicates need for some support, a score ranging from 1 to 1.49 points indicates need for no support. On average, teachers reported needing some or no support in implementing SLOs, with means ranging from 1.31 to 1.65. Teachers with a bachelor's degree or below reported needing more support in implementing SLOs than those with a master's degree or above. To understand whether the differences of support needed were significant, I used independent t-tests. With 6 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .008 (i.e., .05/6). The alpha for the overall support needed to implement SLOs

was set to be .05. On average, teachers with a bachelor's degree or below reported statistically significantly more overall support needed in implementing SLOs than those with a master's degree or above with a small to medium effect size (p = .028, Cohen's d = 0.29) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score based on the six items on the support needed is likely between 0.02 and 0.28 points lower for teachers who had a master's degree or above than those who had a bachelor's degree or below, indicating teachers who had a master's degree or above need less support in implementing SLOs than teachers who had a bachelor's degree or below. The confidence intervals show that I am confident that there is very little difference (at most, 0.41 points) of the support needed between teachers with a bachelor's degree or below and those with a master's degree or above on each of the items.

Table 4.31 *Teachers' Need for Support Based on Degree*

Need for Support in Implementing SLOs	Bachelor or Below	Master or Above	P-value	95% CI
Understanding standards in SLOs	1.49	1.31	.032	[0.02, 0.34]
Implementing standards in SLOs	1.50	1.32	.026	[0.02, 0.34]
Understanding cognitive levels of standards	1.65	1.41	.008	[0.06, 0.41]
Developing assessments for SLOs	1.50	1.39	.197	[-0.06, 0.28]
Setting growth targets for SLOs	1.57	1.48	.260	[-0.07, 0.26]
Analyzing assessment data in SLOs	1.48	1.37	.232	[-0.07, 0.27]
Overall	1.53	1.38	.028	[0.02, 0.28]

Notes: 82 teachers had a degree of bachelor or below, and 153-155 had a master's or above.

In the analysis, teachers' years of experience in education was considered. Table 4.32 describes teachers' views of the impact of SLOs based on their years of experience in education. I calculated the average score for each item on the SLO impact, and I also calculated the average score based on all items to understand the overall impact of SLOs.

The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, career teachers who had more than three years of experience in education reported notably less positive views of the impact of SLOs on all four aspects than early career teachers who had three or fewer years of experience. To understand whether the differences were significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of SLOs was set to be .05. On average, career teachers reported statistically significantly less positive views than early career teachers regarding the overall impact of SLOs with a medium to large effect size ($p < .001$, Cohen's $d = 0.68$), impact of SLOs on evaluating teacher performance effectively with a medium to large effect size ($p = .001$, Cohen's $d = 0.67$), improving teachers' instructional practice with a medium to large effect size ($p < .001$, Cohen's $d = 0.68$), and informing teacher professional development with a medium to large effect size ($p = .001$, Cohen's $d = 0.70$) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the item indicating belief in whether SLOs can be used to evaluate teacher performance effectively is likely between 0.25 and 0.90 points higher for early career teachers than for career teachers. The average score on the item indicating belief in whether SLOs can be used to improve teachers' instructional practices

is likely between 0.27 and 0.94 points higher for early career teachers than for career teachers. The average score on the item indicating belief in whether SLOs can be used to inform teacher professional development is likely between 0.24 and 0.88 points higher for early career teachers than for career teachers. The average score based on the four items indicating belief in the impact of SLOs is likely between 0.24 and 0.82 points higher for early career teachers than for career teachers. These results indicate that early career teachers hold higher belief in these statements than career teachers.

Table 4.32 *Teachers' Perceptions of SLOs Based on Experience*

Impact of SLOs	Early Career (0-3)	Career (3+)	P-value	95% CI
Evaluating teacher performance effectively	2.54	1.96	.001	[0.25, 0.90]
Improving teachers' instructional practice	2.73	2.13	.000	[0.27, 0.94]
Promoting student learning	2.59	2.18	.021	[0.06, 0.75]
Informing teacher PD	2.67	2.10	.001	[0.24, 0.88]
Overall	2.63	2.09	.000	[0.24, 0.82]

Notes: 28-30 teachers had 0-3 years of experience, 207-211 had more than 3 years of experience.

Table 4.33 describes teachers' reported knowledge about SLOs based on their years of experience in education. I calculated the average score for each item on the SLO knowledge, and I also calculated the average score based on all items to understand the overall knowledge about SLOs. The questions are on a 4-point scale (1-No Knowledge; 2- Limited Knowledge; 3- Some Knowledge; 4- Substantial Knowledge). A higher score point indicates more knowledge. An average score ranging from 3.5 to 4 points indicates substantial knowledge, a score ranging from 2.5 to 3.49 points indicates some knowledge, a score ranging from 1.5 to 2.49 points indicates limited knowledge, and a score ranging from 1 to 1.49 points indicates no knowledge. On average, career teachers who had more than three years of experience in education reported more knowledge than early career

teachers who had three or fewer years of experience. To understand whether the differences were significant, I used independent t-tests. With 9 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .006 (i.e., .05/9). The alpha for the overall knowledge about SLOs was set to be .05. On average, career teachers reported statistically significantly more knowledge than early career teachers regarding their overall knowledge about SLOs with a medium to large effect size ($p < .001$, Cohen's $d = 0.70$), purposes of SLOs with a medium to large effect size ($p = .003$, Cohen's $d = 0.57$), student groups to be included in SLO with a medium to large effect size ($p= .002$, Cohen's $d = 0.56$), implementation of SLOs in the district with a medium to large effect size ($p < .001$, Cohen's $d = 0.75$), developing high quality SLOs with a medium to large effect size ($p < .001$, Cohen's $d = 0.72$), selecting appropriate assessments with a medium to large effect size ($p = .002$, Cohen's $d = 0.62$), and instructional strategies to meet SLOs targets with a medium to large effect size ($p = .001$, Cohen's $d = 0.58$) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the knowledge about the purpose SLOs is likely between 0.13 and 0.61 points higher for career teachers than for early career teachers. The average score on the knowledge about student groups to be included is likely between 0.16 and 0.69 points higher for career teachers than for early career teachers. The average score on the knowledge about implementation of SLOs in the district is likely between 0.26 and 0.82 points higher for career teachers than for early career teachers. The average score on the knowledge about developing high quality SLOs is likely between 0.23 and 0.78 points higher for career teachers than for early career

teachers. The average score on the knowledge about selecting appropriate assessment is likely between 0.17 and 0.73 points higher for career teachers than for early career teachers. The average score on the knowledge about instructional strategies to meet SLOs targets is likely between 0.17 and 0.69 points higher for career teachers than for early career teachers. The average score based on the nine items on the knowledge about SLOs is likely between 0.20 and 0.63 points higher for career teachers than for early career teachers. These results indicate that career teachers have more knowledge than early career teachers regarding these aspects of the knowledge about SLOs.

Table 4.33 *Teachers' Knowledge about SLOs Based on Experience*

Knowledge about SLOs	Early Career (0-3)	Career (3+)	P-value	95% CI
Purpose of SLOs	3.03	3.40	.003	[0.13, 0.61]
Student groups to be included in SLOs	2.97	3.39	.002	[0.16, 0.69]
Content to be included in SLOs	3.13	3.48	.008	[0.09, 0.61]
Implementation of SLOs in the district	2.77	3.30	.000	[0.26, 0.82]
Developing high quality SLOs	2.73	3.24	.000	[0.23, 0.78]
Selecting appropriate assessments	2.87	3.32	.002	[0.17, 0.73]
Setting growth targets for SLOs	2.90	3.25	.020	[0.06, 0.65]
Instructional strategies to meet SLOs targets	2.97	3.40	.001	[0.17, 0.69]
Analyzing student assessment data	3.07	3.41	.012	[0.08, 0.61]
Overall	2.94	3.35	.000	[0.20, 0.63]

Notes: 29-30 teachers had 0-3 years of experience, 210-212 had more than 3 years of experience.

Table 4.34 describes teachers' need for support in implementing SLOs based on their years of experience in education. I calculated the average score for each item on the support needed, and I also calculated the average score based on all items to understand the overall support needed. The questions are on a 3-point scale (1-Need No Support; 2-Need Some Support; 3-Need A Lot of Support). A higher score point indicates more

support needed. An average score ranging from 2.5 to 3 points indicates need for a lot of support, a score ranging from 1.5 to 2.49 points indicates need for some support, a score ranging from 1 to 1.49 points indicates need for no support. On average, early career teachers with three or fewer years of experience reported more support needed to implement SLOs then career teachers. To understand whether the differences were significant, I used independent t-tests. With 6 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .008 (i.e., .05/6). The alpha for the overall support needed was set to be .05. On average, early career teachers with three or fewer years of experience reported statistically significantly more support needed to implement SLOs than career teachers regarding the overall support needed with a large effect size (p = .001, Cohen's d = 0.84), understanding standards in SLOs with a large effect size (p < .001, Cohen's d = 0.85), implementing standards in SLOs with a large effect size (p < .001, Cohen's d = 1.11), understanding cognitive levels of standards with a medium to large effect size (p < .001, Cohen's d = 0.79), developing assessments for SLOs with a medium to large effect size (p = .006, Cohen's d = 0.64), setting growth targets for SLOs with a medium effect size (p = .006, Cohen's d = 0.50) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the support needed in understanding standards in SLOs is likely between 0.26 and 0.79 points lower for career teachers than for early career teachers. The average score on the support needed in implementing standards in SLOs is likely between 0.44 and 0.83 points lower for career teachers than for early career teachers. The average score on the support needed in understanding cognitive

levels of standards is likely between 0.28 and 0.73 points lower for career teachers than for early career teachers. The average score on the support needed in developing assessment for SLOs is likely between 0.13 and 0.69 points lower for career teachers than for early career teachers. The average score on the support needed in setting growth targets for SLOs is likely between 0.09 and 0.56 points lower for career teachers than for early career teachers. The average score based on the six items on the support needed in implementing SLOs is likely between 0.22 and 0.71 points lower for career teachers than for early career teachers. The results indicated that career teachers need less support than early career teachers on these aspects.

Table 4.34 *Teachers' Need for Support Based on Experience*

Need for Support in Implementing SLOs	Early Career (0-3)	Career (3+)	P-value	95% CI
Understanding standards in SLOs	1.83	1.31	.000	[0.26, 0.79]
Implementing standards in SLOs	1.93	1.30	.000	[0.44, 0.83]
Understanding cognitive levels of standards	1.93	1.43	.000	[0.28, 0.73]
Developing assessments for SLOs	1.77	1.36	.006	[0.13, 0.69]
Setting growth targets for SLOs	1.80	1.47	.006	[0.09, 0.56]
Analyzing assessment data in SLOs	1.73	1.36	.011	[0.09, 0.66]
Overall	1.83	1.37	.001	[0.22, 0.71]

Notes: 30 teachers had 0-3 years of experience, 211-212 had more than 3 years of experience.

In the analysis, teachers' experience of using SLOs was also considered. Table 4.35 describes teachers' views of the impact of SLOs based on their years of experience in using SLOs. I calculated the average score for each item on the SLO impact, and I also calculated the average score based on all items to understand the overall impact of SLOs. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement a score ranging from 2.5 to 3.49

points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, teachers who had three or fewer years of experience of using SLOs reported notably more positive views of the impact of SLOs than those who had more than three years of experience. To understand whether the differences were significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of SLOs was set to be .05. On average, teachers who had three or fewer years of experience of using SLOs reported significantly more positive views than those who had more than three years of experience regarding the overall impact of SLOs with a small to medium effect size (p = .007, Cohen's d = 0.39), the impact of SLOs on teachers' instructional practice with a small to medium effect size (p = .010, Cohen's d = 0.38), and promoting student learning with a small to medium effect size (p = .012, Cohen's d = 0.38) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the item indicating belief in whether SLOs can be used to improve teachers' instructional practices is likely between 0.08 and 0.58 points higher for teachers who had three or fewer years of experience in using SLOs than those who had more than three years of experience in using SLOs. The average score on the item indicating belief in whether SLOs can be used to promote student learning is likely between 0.07 and 0.58 points higher for teachers who had three or fewer years of experience in using SLOs than those who had more than three years of experience in using SLOs. The average score on the four items indicating belief in the impact of SLOs

is likely between 0.08 and 0.52 points higher for teachers who had three or fewer years of experience in using SLOs than those who had more than three years of experience in using SLOs. These results indicate that teachers who had three or fewer years of experience in using SLOs hold higher belief in these statements than those who had more than three years of experience in using SLOs.

Table 4.35 *Teachers' Perceptions of SLOs Based on SLOs Experience*

Impact of SLOs	0-3 Year	3+ Years	P-value	95% CI
Evaluating teacher performance effectively	2.22	1.96	.031	[0.03, 0.50]
Improving teachers' instructional practice	2.45	2.12	.010	[0.08, 0.58]
Promoting student learning	2.47	2.14	.012	[0.07, 0.58]
Informing teacher PD	2.40	2.11	.020	[0.05, 0.53]
Overall	2.38	2.08	.007	[0.08, 0.52]

Notes: 63-65 teachers had 0-3 years of experience, 165-169 had more than 3 years of experience.

Table 4.36 describes teachers' reported knowledge about SLOs based on their years of experience of using SLOs. I calculated the average score for each item on the SLO knowledge, and I also calculated the average score based on all items to understand the overall knowledge about SLOs. The questions are on a 4-point scale (1-No Knowledge; 2- Limited Knowledge; 3- Some Knowledge; 4- Substantial Knowledge). A higher score point indicates more knowledge. An average score ranging from 3.5 to 4 points indicates substantial knowledge, a score ranging from 2.5 to 3.49 points indicates some knowledge, a score ranging from 1.5 to 2.49 points indicates limited knowledge, and a score ranging from 1 to 1.49 points indicates no knowledge. On average, teachers who had more than three years of experience of using SLOs reported more knowledge about SLOs than those who had three or fewer years of experience. To understand whether the differences were significant, I used independent t-tests. With 9 items for

comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .006 (i.e., .05/9). The alpha for the overall knowledge about SLOs was set to be .05. On average, teachers who had more than three years of experience of using SLOs reported statistically significantly more knowledge than those who had three or fewer years of experience regarding their overall knowledge about SLOs with a small to medium effect size (p = .001, Cohen's d = 0.48), the purpose of SLOs with a medium to large effect size (p < .001, Cohen's d = 0.53), developing high quality of SLOs with a small to medium effect size (p = .002, Cohen's d = 0.46), setting growth targets for SLOs with a small to medium effect size (p = .005, Cohen's d = 0.41), instructional strategies to meet SLOs targets with a small to medium effect size (p = .004, Cohen's d = 0.42), and analyzing student assessment data with a small to medium effect size (p = .002, Cohen's d = 0.45) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the knowledge about the purpose SLOs is likely between 0.16 and 0.52 points higher for teachers who had more than three years of experience in using SLOs than those who had three or fewer years of experience in using SLOs. The average score on the knowledge about developing high quality SLOs is likely between 0.12 and 0.53 points higher for teachers who had more than three years of experience in using SLOs than those who had three or fewer years of experience in using SLOs. The average score on the knowledge about setting growth targets for SLOs is likely between 0.10 and 0.53 points higher for teachers who had more than three years of experience in using SLOs than those who had three or fewer years of experience in using SLOs. The average score on the knowledge about instructional strategies to meet SLOs

targets is likely between 0.10 and 0.48 points higher for teachers who had more than three years of experience in using SLOs than those who had three or fewer years of experience in using SLOs. The average score on the knowledge about analyzing student assessment data is likely between 0.11 and 0.50 points higher for teachers who had more than three years of experience in using SLOs than those who had three or fewer years of experience in using SLOs. The average score based on the nine items on the knowledge about SLOs is likely between 0.11 and 0.44 points higher for teachers who had more than three years of experience in using SLOs than those who had three or fewer years of experience in using SLOs. These results indicate that teachers who had more than three years of experience in using SLOs have more knowledge about SLOs regarding these aspects than those who had three or fewer years of experience in using SLOs.

Table 4.36 *Teachers' Knowledge about SLOs Based on SLOs Experience*

Knowledge about SLOs	0-3 Year	3+ Years	P-value	95% CI
Purpose of SLOs	3.11	3.45	.000	[0.16, 0.52]
Student groups to be included in SLOs	3.17	3.43	.011	[0.06, 0.46]
Content to be included in SLOs	3.34	3.49	.140	[-0.05, 0.34]
Implementation of SLOs in the district	3.03	3.32	.008	[0.08, 0.50]
Developing high quality SLOs	2.94	3.27	.002	[0.12, 0.53]
Selecting appropriate assessments	3.09	3.32	.028	[0.03, 0.44]
Setting growth targets for SLOs	2.98	3.29	.005	[0.10, 0.53]
Instructional strategies to meet SLOs targets	3.14	3.43	.004	[0.10, 0.48]
Analyzing student assessment data	3.15	3.46	.002	[0.11, 0.50]
Overall	3.11	3.38	.001	[0.11, 0.44]

Notes: 64-65 teachers had 0-3 years of experience, 169-170 had more than 3 years of experience.

Table 4.37 describes teachers' need for support in implementing SLOs based on their years of experience of using SLOs. I calculated the average score for each item on the support needed, and I also calculated the average score based on all items to understand the overall support needed. The questions are on a 3-point scale (1-Need No Support; 2-Need Some Support; 3-Need A Lot of Support). A higher score point indicates

more support needed. An average score ranging from 2.5 to 3 points indicates need for a lot of support, a score ranging from 1.5 to 2.49 points indicates need for some support, a score ranging from 1 to 1.49 points indicates need for no support. On average, teachers who had three or fewer years of experience in using SLOs reported more support needed to implement SLOs than those who had more than three years of experience. To understand whether the differences of support needed were significant, I used independent t-tests. With 6 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .008 (i.e., .05/6). The alpha for the overall support needed was set to be .05. On average, teachers who had three or fewer years of experience in using SLOs reported statistically significantly more support needed to implement SLOs regarding the overall support needed with a medium to large effect size ($p < .001$, Cohen's $d = 0.60$), understanding standards in SLOs with a medium to large effect size ($p= .002$, Cohen's $d = 0.51$), implementing standards in SLOs with a medium to large effect size ($p < .001$, Cohen's $d = 0.61$), understanding cognitive levels of standards with a medium to large effect size ($p < .001$, Cohen's $d = 0.58$), setting growth targets for SLOs with a small to medium effect size ($p = .005$, Cohen's $d = 0.44$), and analyzing assessment data in SLOs with a medium to large effect size ($p < .001$, Cohen's $d = 0.57$) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the support in understanding standards in SLOs is likely between 0.11 and 0.46 points lower for teachers who had more than three years of experience in using SLOs than those who had three or fewer years of experience in using SLOs. The average score on the support in implementing standards in SLOs is likely

between 0.17 and 0.51 points lower for teachers who had more than three years of experience in using SLOs than those who had three or fewer years of experience in using SLOs. The average score on the support in understanding cognitive levels of standards is likely between 0.18 and 0.52 points lower for teachers who had more than three years of experience in using SLOs than those who had three or fewer years of experience in using SLOs. The average score on the support in setting growth targets for SLOs is likely between 0.09 and 0.47 points lower for teachers who had more than three years of experience in using SLOs than those who had three or fewer years of experience in using SLOs. The average score on the support in analyzing assessment data in SLOs is likely between 0.16 and 0.52 points lower for teachers who had more than three years of experience in using SLOs than those who had three or fewer years of experience in using SLOs. The average score based on the six items about the support needed is likely between 0.15 and 0.46 points lower for teachers who had more than three years of experience in using SLOs than those who had three or fewer years of experience in using SLOs. These results indicate that teachers who had more than three years of experience in using SLOs need less support in these aspects than those who had three or fewer years of experience in using SLOs.

Table 4.37 *Teachers' Need for Support Based on SLOs Experience*

Need for Support in Implementing SLOs	0-3 Year	3+ Years	P-value	95% CI
Understanding standards in SLOs	1.57	1.28	.002	[0.11, 0.46]
Implementing standards in SLOs	1.62	1.28	.000	[0.17, 0.51]
Understanding cognitive levels of standards	1.74	1.39	.000	[0.18, 0.52]
Developing assessments for SLOs	1.57	1.35	.021	[0.03, 0.41]
Setting growth targets for SLOs	1.71	1.43	.005	[0.09, 0.47]
Analyzing assessment data in SLOs	1.65	1.31	.000	[0.16, 0.52]
Overall	1.64	1.34	.000	[0.15, 0.46]

Notes: 65 teachers had 0-3 years of experience, 168-170 had more than 3 years of experience.

The grade levels taught were taken into consideration in the analysis of teachers' views of SLOs (Table 4.38). I calculated the average score for each item on the SLO impact, and I also calculated the average score based on all items to understand the overall impact of SLOs. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, teachers who taught PK-5 reported slightly more positive views of the impact of SLOs than those teachers who taught middle school grade levels (Grades 6-8) and high school grade levels (Grades 9-12). To understand whether these differences were statistically significant, I used ANOVA. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of SLOs was set to be .05. On average, teachers did not have significantly different views of the impact of SLOs based on the grade levels taught.

Table 4.38 *Teachers' Perceptions of SLOs Based on Grade Levels*

Impact of SLOs	PK-5	Grades 6-8	Grades 9-12	P-value
Evaluating teacher performance effectively	2.06	1.97	1.93	.613
Improving teachers' instructional practice	2.22	2.17	2.14	.844
Promoting student learning	2.27	2.03	2.19	.333
Informing teacher PD	2.17	2.14	2.16	.981
Overall	2.18	2.09	2.12	.763

Notes: 147-151 teachers taught PK-5, 35-36 taught Grades 6-8, 42-43 taught Grades 9-12.

Table 4.39 describes teachers' reported knowledge about SLOs based on the grade levels they taught. I calculated the average score for each item on the SLO

knowledge, and I also calculated the average score based on all items to understand the overall knowledge about SLOs. The questions are on a 4-point scale (1-No Knowledge; 2- Limited Knowledge; 3- Some Knowledge; 4- Substantial Knowledge). A higher score point indicates more knowledge. An average score ranging from 3.5 to 4 points indicates substantial knowledge, a score ranging from 2.5 to 3.49 points indicates some knowledge, a score ranging from 1.5 to 2.49 points indicates limited knowledge, and a score ranging from 1 to 1.49 points indicates no knowledge. On average, teachers reported similar knowledge about SLOs regardless of the grade levels they taught. To understand whether the differences were significant, I used ANOVA. With 9 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .006 (i.e., .05/9). The alpha for the overall knowledge about SLOs was set to be .05. On average, teachers did not report significantly different knowledge levels based on the grade levels taught.

Table 4.39 *Teachers' Knowledge about SLOs Based on Grade Levels*

Knowledge about SLOs	PK-5	Grades 6-8	Grades 9-12	P-value
Purpose of SLOs	3.34	3.39	3.42	.754
Student groups to be included in SLOs	3.34	3.43	3.30	.729
Content to be included in SLOs	3.43	3.53	3.45	.742
Implementation of SLOs in the district	3.23	3.42	3.14	.257
Developing high quality SLOs	3.19	3.19	3.14	.937
Selecting appropriate assessments	3.22	3.31	3.37	.449
Setting growth targets for SLOs	3.26	3.25	3.10	.464
Instructional strategies to meet SLOs targets	3.37	3.40	3.28	.676
Analyzing student assessment data in SLOs	3.42	3.47	3.19	.098
Overall	3.31	3.38	3.26	.682

Notes: 150-151 teachers taught PK-5, 35-36 taught Grades 6-8, 42-43 taught Grades 9-12.

Table 4.40 describes teachers' need for support in implementing SLOs based on the grade levels they taught. I calculated the average score for each item on the support needed, and I also calculated the average score based on all items to understand the

overall support needed. The questions are on a 3-point scale (1-Need No Support; 2-Need Some Support; 3-Need A Lot of Support). A higher score point indicates more support needed. An average score ranging from 2.5 to 3 points indicates need for a lot of support, a score ranging from 1.5 to 2.49 points indicates need for some support, a score ranging from 1 to 1.49 points indicates need for no support. On average, teachers reported similar support needed to implement SLOs regardless of the grade levels they taught. To understand whether the differences of support needed were statistically significant, I used ANOVA. With 6 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .008 (i.e., .05/6). The alpha for the overall support needed was set to be .05. On average, there were no statistically significant differences of support needed to implement SLOs among the teachers who taught different grade levels.

Table 4.40 *Teachers' Need for Support Based on Grade Levels*

Need for Support in Implementing SLOs	PK-5	Grades 6-8	Grades 9-12	P-value
Understanding standards in SLOs	1.33	1.42	1.33	.685
Implementing standards in SLOs	1.33	1.42	1.40	.562
Understanding cognitive levels of standards	1.45	1.56	1.48	.668
Developing assessments for SLOs	1.40	1.53	1.33	.301
Setting growth targets for SLOs	1.47	1.58	1.49	.602
Analyzing assessment data in SLOs	1.33	1.56	1.44	.077
Overall	1.39	1.51	1.41	.377

Notes: 150-151 teachers taught PK-5, 36 taught Grades 6-8, 42-43 taught Grades 9-12.

4.2.3 Educators' Perceptions of Classroom Observations

Educators' views of classroom observations were analyzed using both percentages and means (Table 4.41). About 57% or more of the educators agreed or strongly agreed that using classroom observations evaluates teacher performance effectively, improves teachers' instructional practice, promotes student learning, and

informs teachers' professional development. Comparatively, administrators held much more positive views of the impact of SLOs, with more than 93% of the administrators agreeing or strongly agreeing with the four aspects. Between 57.6% and 71.3% of the teachers agreed or strongly agreed on the impact of classroom observations.

I calculated the average score for each item on the impact of classroom observations, and I also calculated the average score based on all items to understand the overall impact of classroom observations. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, educators agreed on the impact of classroom observations (M = 2.75), administrators had more positive views (M = 3.33) than teachers (M = 2.71). Independent t-tests were conducted to help understand whether the views of teachers and administrator are significantly different. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of classroom observations was set to be .05. On average, administrators had significantly more positive views than teachers regarding the overall impact of classroom observations with a large effect size (p = .002, Cohen's d = 0.94), and the impact of classroom observations on evaluating teacher performance effectively with a medium to large effect size (p = .011, Cohen's d = 0.79), improving teachers' instructional practice with a medium to large effect size (p < .001, Cohen's d = 0.77), promoting student learning with a large effect

size (p = .003, Cohen's d = 0.83), and informing teachers' professional development with a large effect size (p = .003, Cohen's d = 0.95) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the item indicating belief in whether classroom observations can be used to evaluate teacher performance effectively is likely between 0.12 and 0.96 points higher for administrators than for teachers. The average score on the item indicating belief in whether classroom observations can be used to improve teachers' instructional practice is likely between 0.28 and 0.83 points higher for administrators than for teachers. The average score on the item indicating belief in whether classroom observations can be used to promote student learning is likely between 0.24 and 1.16 points higher for administrators than for teachers. The average score on the item indicating belief in whether classroom observations can be used to inform teacher professional development is likely between 0.25 and 1.15 points higher for administrators than for teachers. The average score based on the four items on the impact of classroom observations is likely between 0.22 and 1.02 points higher for administrators than for teachers. These results suggest that administrators hold much higher beliefs in these statements about the impact of classroom observations than teachers.

To check the analysis, I also used Mann-Whitney Tests considering the sample size of administrators is small. The results are consistent, and there was a statistically significant difference between teachers' and administrators' views of the overall impact of classroom observations (Z = -3.38, p = .001) and all four aspects of the classroom observations.

Table 4.41 *Educators' Views of Classroom Observations*

Impact of Classroom Observations	Percentage			Mean				
	All	Teacher	Administrator	All	Teacher	Administrator	*P*-value	95% CI
Evaluating teacher performance effectively	71.3	69.3	100.0	2.81	2.77	3.31	.011	[0.12, 0.96]
Improving teachers' instructional practice	70.1	68.8	100.0	2.79	2.76	3.31	.000	[0.28, 0.83]
Promoting student learning	57.6	55.4	93.8	2.59	2.55	3.25	.003	[0.24, 1.16]
Informing teacher PD	70.0	68.5	100.0	2.79	2.74	3.44	.003	[0.25, 1.15]
Overall	67.3	65.5	98.5	2.75	2.71	3.33	.002	[0.22, 1.02]

Notes: Respondents include 235-242 teachers and 16 administrators.

4.2.4 Teachers' Perceptions of Classroom Observations Based on Variables

Classroom observations are a major mode of teacher evaluation. This section focused on the associations of teachers' perceptions of the impact of classroom observations and their educational background. Teachers' academic degrees, years of experience in education, years of experience of using SLOs, and grade levels taught.

Teachers' highest academic degree was analyzed (Table 4.42). I calculated the average score for each item on the impact of classroom observations, and I also calculated the average score based on all items to understand the overall impact of classroom observations. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement.

On average, teachers who had a bachelor's degree or below reported more positive views of the impact of classroom observations than those who had a master's degree or above. To understand whether the differences were significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of classroom observations was set to be .05. On average, teachers who had a bachelor's degree or below reported statistically significantly more positive views of the impact of classroom observations than those who had a master's degree or above regarding the overall impact of classroom observations with a small to

medium effect size (p = .002, Cohen's d = 0.41), impact on evaluating teacher performance effectively with a small to medium effect size (p = .005, Cohen's d = 0.39), improving teachers' instructional practice with a small to medium effect size (p = .001, Cohen's d = 0.45), promoting student learning with a small to medium effect size (p = .012, Cohen's d = 0.35) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the item indicating belief in whether classroom observations can be used to evaluate teacher performance effectively is likely between 0.09 and 0.51 points higher for teachers with a bachelor's degree or below than for teachers with a master's degree or above. The average score on the item indicating belief in whether classroom observations can be used to improve teachers' instructional practice is likely between 0.16 and 0.61 points higher for teachers with a bachelor's degree or below than for teachers with a master's degree or above.

The average score on the item indicating belief in whether classroom observations can be used to promote student learning is likely between 0.07 and 0.55 points higher for teachers with a bachelor's degree or below than for teachers with a master's degree or above. The average score based on the four items on the impact of classroom observations is likely between 0.11 and 0.52 points higher for teachers with a bachelor's degree or below than for teachers with a master's degree or above. These results suggest that teachers with a bachelor's degree or below hold higher beliefs in these statements about the impact of classroom observations than for teachers with a master's degree or above.

Table 4.42 *Teachers' Views of Classroom Observations Based on Degree*

Impact of Classroom Observations	Bachelor or Below	Master or Above	P-value	95% CI
Evaluating teacher performance effectively	2.98	2.67	.005	[0.09, 0.51]
Improving teachers' instructional practice	3.01	2.63	.001	[0.16, 0.61]
Promoting student learning	2.75	2.44	.012	[0.07, 0.55]
Informing teacher PD	2.92	2.66	.027	[0.03, 0.50]
Overall	2.92	2.60	.002	[0.11, 0.52]

Notes: 79-82 teachers had a degree of bachelor or below, and 149-153 had a master's or above.

Teachers' years of experience was considered in the analysis (Table 4.43). I calculated the average score for each item on the impact of classroom observations, and I also calculated the average score based on all items to understand the overall impact of classroom observations. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement. On average, early career teachers who had three or fewer years of experience in education reported more positive views of the impact of classroom observations than career teachers who had more than three years of experience in education. To understand whether the differences were significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of classroom observations was set to be .05. On average, early career teachers who had three or fewer years of experience in education reported statistically significantly more positive views of the impact of classroom observations regarding the

overall impact with a medium to large effect size (p = .001, Cohen's d = 0.58), impact on student learning with a medium to large effect size (p = .004, Cohen's d = 0.53), and impact on informing teacher professional development with a medium to large effect size (p < .001, Cohen's d = 0.67) (Cohen, 1988).

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The average score on the item indicating belief in whether classroom observations can be used to promote student learning is likely between 0.15 and 0.72 points higher for early career teachers than for career teachers. The average score on the item indicating belief in whether classroom observations can be used to inform teacher professional development is likely between 0.26 and 0.80 points higher for early career teachers than for career teachers. The average score based on the four items on the impact of classroom observations is likely between 0.17 and 0.65 points higher for early career teachers than for career teachers. These results suggest that early career teachers hold higher beliefs in these statements about the impact of classroom observations than career teachers.

Table 4.43 *Teachers' Views of Classroom Observations Based on Experience*

Impact of Classroom Observations	Early Career (0-3)	Career (3+)	P-value	95% CI
Evaluating teacher performance effectively	3.00	2.74	.060	[-0.01, 0.54]
Improving teachers' instructional practice	3.14	2.71	.015	[0.08, 0.78]
Promoting student learning	2.93	2.50	.004	[0.15, 0.72]
Informing teacher PD	3.21	2.68	.000	[0.26, 0.80]
Overall	3.07	2.66	.001	[0.17, 0.65]

Notes: 28-30 teachers had 0-3 years of experience, 205-210 had more than 3 years of experience.

Teachers' experience of using SLOs was taken into consideration in the analysis (Table 4.44). I calculated the average score for each item on the impact of classroom

observations, and I also calculated the average score based on all items to understand the overall impact of classroom observations. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement.

On average, teachers who had three or fewer years of experience of using SLOs reported more positive views of the impact of classroom observations than those who had more than three years of experience. To understand whether the differences were significant, I used independent t-tests. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of classroom observations was set to be .05. On average, teachers reported similar views of the impact of classroom observations regardless of their experience of using SLOs.

Confidence intervals were constructed to understand the ranges of the differences for each aspect. The confidence intervals show that I am confident that there is very little difference (at most, 0.41 points) of the perceptions of classroom observations between teachers who had three or fewer years of experience in using SLOs and those who had more than years of experience in using SLOs on each of the items, as well as the average of all items.

Table 4.44 *Teachers' Views of Classroom Observations Based on SLOs Experience*

Impact of Classroom Observations	0-3 Years	3+ Years	P-value	95% CI
Evaluating teacher performance effectively	2.84	2.77	.534	[-0.16, 0.32]
Improving teachers' instructional practice	2.89	2.74	.218	[-0.09, 0.40]
Promoting student learning	2.67	2.53	.311	[-0.13, 0.41]
Informing teacher PD	2.89	2.74	.276	[-0.12, 0.40]
Overall	2.83	2.70	.240	[-0.09, 0.36]

Notes: 61-65 teachers had 0-3 years of experience, 164-168 had more than 3 years of experience.

Grade levels taught were considered in the analysis of teachers' views of classroom observations (Table 4.45). I calculated the average score for each item on the impact of classroom observations, and I also calculated the average score based on all items to understand the overall impact of classroom observations. The questions are on a 4-point scale (1-Strongly Disagree; 2-Disagree; 3-Agree; 4-Strongly Agree). A higher score point indicates more positive views. An average score ranging from 3.5 to 4 points indicates strong agreement, a score ranging from 2.5 to 3.49 points indicates agreement, a score ranging from 1.5 to 2.49 points indicates disagreement, and a score ranging from 1 to 1.49 points indicates strong disagreement.

On average, teachers who taught PK-5 reported slightly more positive views of the overall impact of classroom observations than those who taught middle school grade levels (Grades 6-8) and high school grade levels (Grades 9-12). To understand whether the differences were significant, I used ANOVA. With 4 items for comparison, I applied Bonferroni correction to adjust the familywise alpha of .05 to reduce the Type I error to .013 (i.e., .05/4). The alpha for the overall impact of classroom observations was set to be .05. On average, teachers did not have significantly different views of the impact of classroom observations based on the grade levels they taught.

Table 4.45 *Teachers' Views of Classroom Observations Based on Grade Levels*

Impact of Classroom Observations	PK-5	Grades 6-8	Grades 9-12	P-value
Evaluating teacher performance effectively	2.82	2.89	2.56	.138
Improving teachers' instructional practice	2.82	2.72	2.62	.395
Promoting student learning	2.61	2.47	2.38	.348
Informing teacher PD	2.82	2.79	2.50	.130
Overall	2.77	2.72	2.53	.211

Notes: 146-148 teachers taught PK-5, 33-36 taught Grades 6-8, 39-43 taught Grades 9-12.

4.2.5 Summaries of Educator' Views of SLOs and Classroom Observations

To better understand educators' views of the SLOs and classroom observations in the evaluation of teacher effectiveness, I compared their views (Table 4.46). Overall, they seemed to agree that classroom observations had positive impact on all four aspects with means larger than 2.5. However, they appeared to disagree that SLOs had positive impact on all four aspects with means smaller than 2.5. Administrators reported more positive views of the impact of classroom observations (M = 3.33) than the impact of SLOs (M = 2.35). Teachers reported much more positive views of classroom observations (M = 2.71) than SLOs (M = 2.16). Administrators reported more positive views of the impact of SLOs and classroom observations than teachers.

Table 4.46 *Comparing Educators' Views of SLOs and Classroom Observations*

Aspects of Impact	SLOs		Classroom Observations	
	Teachers	Administrators	Teachers	Administrators
Evaluating teacher performance effectively	2.03	2.25	2.77	3.31
Improving teachers' instructional practice	2.21	2.44	2.76	3.31
Promoting student learning	2.23	2.33	2.55	3.25
Informing teacher PD	2.17	2.38	2.74	3.44
Overall	2.16	2.35	2.71	3.33

Teachers' perceptions of the impact of SLOs and classroom observations were summarized and compared (Table 4.47). Teachers' overall views of the impact of SLOs and classroom observations were calculated based on the four aspects. Teachers' overall knowledge about SLOs was calculated based on the nine aspects. Teachers' overall support needed to implement SLOs was also calculated based on the six aspects. In the analysis, teachers' academic degree, years of experience in education, years of experience in using SLOs, and the grade levels taught were considered.

In comparison with teachers who had a bachelor's degree, those with an EdS, masters, or Ph.D. degree reported statistically significantly less positive views of the impact of classroom observations with a small to medium effect size (p = .002, Cohen's d = 0.41), significantly more knowledge about SLOs with a small to medium effect size (p = .011, Cohen's d = 0.35), and significantly less support needed to implement SLOs with a small to medium effect size (p = .028, Cohen's d = 0.29). In comparison with early career teachers, career teachers reported statistically significantly less positive views of the impact of SLOs with a medium to large effect size (p < .001, Cohen's d = 0.68), significantly less positive views of the impact of classroom observations with a medium to large effect size (p = .001, Cohen's d = 0.58), significantly more knowledge about SLOs with a medium to large effect size (p< .001, Cohen's d = 0.70), and significantly less support needed to implement SLOs with a large effect size (p = .001, Cohen's d = 0.84). In comparison with teachers with three or fewer years of experience of using SLOs, those with more than three years of experience reported statistically significantly less positive views of the impact of SLOs with a small to medium effect size (p = .007, Cohen's d = 0.39), significantly more knowledge about SLOs with a small to medium

effect size (p = .001, Cohen's d = 0.48), and significantly less support needed to implement SLOs with a medium to large effect size (p < .001, Cohen's d = 0.60). Teachers teaching different grade levels reported similar views of the impact of SLOs and classroom observations, similar levels of more knowledge about SLOs and support needed to implement SLOs.

Table 4.47 *Teachers' Views Based on Educational Background*

Variables		SLOs			Classroom Observations
		Impact	Knowledge	Support	Impact
Degree	Bachelor	2.24	3.18*	1.53*	2.92*
	EdS, Master, Ph.D.	2.09	3.38*	1.38*	2.60*
Experience in Education	0-3 Years (Early Career)	2.63*	2.94*	1.83*	3.07*
	3+ Years (Career)	2.09*	3.35*	1.37*	2.66*
Experience in SLOs	0-3 Years	2.38*	3.11*	1.64*	2.83
	3+ Years	2.08*	3.38*	1.34*	2.70
Grade	PK-5	2.18	3.31	1.39	2.77
	Grades 6-8	2.09	3.38	1.51	2.72
	Grades 9-12	2.12	3.26	1.41	2.53
Overall		2.16	3.30	1.44	2.71

*Note: * indicates statistically significant differences.*

To investigate whether teachers' perceptions of SLOs and their perceptions of classroom observations had some relationships, I conducted correlation analyses (Table 4.48). Considering the ordinal feature of the data, I used Spearman's rho correlation analysis. Teachers' perceptions of the four aspects of the impact of SLOs and the classroom observations had statistically significant positive associations, and the relationships were moderate. I also explored the relationship of teachers' overall perceptions of SLOs and their overall perceptions of classroom observations. Considering the data are continuous, I used Pearson's correlation analysis. The correlation coefficient

was calculated to be .395 which is statistically significant. Teachers who had more positive views of classroom observations tended to have more positive views of SLOs.

Table 4.48 *Correlations of Teachers' Views of SLOs and Classroom Observations*

Aspects of Impact	Correlations	p-value
Evaluating teacher performance effectively	.317	<.001
Improving teachers' instructional practice	.294	<.001
Promoting student learning	.343	<.001
Informing teacher professional development	.364	<.001
Overall	.395	<.001

To further understand the factors that might predict teachers' perceptions of SLOs, I used multiple regression analysis. In the study, teachers' overall perception of SLOs was the dependent variable, and variables including teachers' perceptions of classroom observations, knowledge about SLOs, support needed to implement SLOs, years of experience in education, and years of experience of using SLOs were the independent variables.

Results shown in Table 4.49 indicate a statistically significant association between teachers' overall perception of SLOs and the independent variables, $F(223, 5) = 9.32$, $p < .01$. Teachers' perceptions of the classroom observations was identified as a significant predictor of their perceptions of SLOs. Overall, teachers who have more positive perceptions of the classroom observations tended to have more positive perceptions of SLOs ($B = 0.34$, $p < .01$). Teachers' knowledge about SLOs, support needed to implement SLOs, years of experience in education, and years of experience of using SLOs were not significant predictors of their perceptions of SLOs. The effect size adjusted R^2 was .173, indicating that 17.3% of the variation in teachers' perceptions of SLOs could be explained by the predictors.

Table 4.49 *Multiple Regression Analysis Results*

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	0.528	0.500		1.056	.292
Classroom Observation	0.344	0.061	0.352	5.640	.000
SLOs Knowledge	0.206	0.108	0.154	1.913	.057
SLOs Support	0.182	0.127	0.113	1.432	.154
Experience of SLOs	-0.033	0.023	-0.096	-1.434	.153
Experience in Education	-0.007	0.006	-0.079	-1.161	.247

4.2.6 Teachers' Additional Thoughts about SLOs and Teacher Evaluation

At the end of the survey, participants were asked to share their thoughts about using SLOs in teacher evaluation or teacher evaluation in general. Among 289 participants, 114 shared additional thoughts and most of them were teachers. Their responses were coded using R for Qualitative Data Analysis (RQDA). Teachers expressed various concerns about using SLOs in the evaluation of teacher effectiveness. These concerns included issues of using students' test performance to evaluate teacher effectiveness, issues of the assessment methods (e.g., standardized tests) in measuring students' growth, time and timeline in implementing SLOs, paperwork, applicability to special education teachers, arts teachers, ESL teachers, and media specialists, lack of supervision and subjectivity in goal-setting and assessment, teaching standards, teach to the test, and lack of feedback based on evaluation results.

Teachers indicated that SLOs were neither effective nor accurate in the evaluation of teacher effectiveness, and they considered that SLOs results did not accurately reflect teachers' performance or effectiveness. They shared that students' test performance/academic growth were affected by various factors, and it was not fair,

reliable, or valid to judge teacher performance or effectiveness based on student test performance/results. Teachers explained that they should be evaluated based on teacher actions rather than student performance. Some supported the use of classroom observations in the evaluation of teacher effectiveness and considered observations as a better way to capture teacher performance. For example, one teacher indicated, "SLOs are not a fair measure. They can only be fair if either all assessments are standardized by the state or all assessments are teacher made. Teacher observations should always be to assess the teacher, not the students, and help the teacher GROW." However, a few teachers shared that the SCTS 4.0 observational instrument was too cumbersome, and limited numbers of observations might not capture some indicators. In addition, a few teachers shared that some evaluators might not have the required qualifications and might be biased in evaluating teacher effectiveness fairly and accurately. For example, one teacher shared, "Different evaluators have the same guidelines but if their teaching styles differ from yours how can they possibly, fairly do your evaluation with your teaching style?"

Setting growth targets and assessment method were a major concern. Teachers shared that they did not receive guidance in setting reasonable learning goals for students. There were mainly two types of assessment methods: standardized tests and teacher developed assessment. Teachers shared issues of using standardized tests to measure student growth. Some standardized tests (e.g., MAP) did not align with the grade-level standards. The pre/post assessment was invalid for some subject areas (e.g., science). One teacher indicated, "We receive professional development that encourages alternate ways of assessing real growth, but the state department of education still uses standardized tests

to measure growth." Regarding the teacher developed assessment methods, there were flaws as well. A few teachers indicated that some of their colleagues purposefully set low learning objectives/goals/targets and manipulated/altered their students' growth data to meet the goals. As one teacher indicated, "Overall, I like SLOs because there is slightly more accountability than the former GBE system, BUT I am not a fan of assessments which are subjective being used or teachers setting targets to be safe. Manipulating your own results should be through teaching and not by setting low targets or not encouraging students to do their best on pre-assessments." In addition, some teachers set growth targets based on one standard, which led to the neglect of other teaching standards. Some tests change from year to year and the assessments results were not fair or consistent. Most importantly, teachers indicated that some teachers taught to the test to help students achieve the goals. One teacher shared, "We as teachers are offered all kinds of really good progressive professional development opportunities, but they don't align with the reality of how we and our students are measured by the Department of Education. It all results in the same old teaching to the test to survive in this field."

Time, timeline, and paperwork was a major concern. Teachers considered the SLOs process was time-consuming, required a lot of paperwork, added a lot of unnecessary extra work to their schedule, and became a burden or stress. One teacher shared, "I feel that SLO's in general become more of a hassle than an asset. The time spent implementing and assessing SLO processes could be better utilized in creating more meaningful student activities." In particular, many teachers shared that timeline was inappropriate. They were required to submit their SLOs student assessment results in March when they have not completed teaching or assessment of student learning. The

assessment data would not accurately capture students' growth for the year. As one teacher shared, "I believe the SLO's should last longer. We should at least let the evaluation go until May. Having the evaluation end in March does not allow us to monitor student growth effectively. We do not stop teaching in March so we should not stop our evaluation."

Some teachers indicated that SLOs were not applicable or less effective for special education teachers, arts teachers, gifted teachers, ESL teachers, librarians, speech pathologists, media specialists, and guidance counselors. As one teacher indicated, "Some positions SLO cannot be used such as: gifted teachers, ESL teachers, and librarians/speech pathologists/guidance counselors, but administrators implement them anyway." Teachers believed that SLOs should not be one size fits all. They indicated that it was challenging for special education students to show growth required by SLOs, and some special education teachers were already using IEP goals and objectives and using SLOs was redundant to them. One teacher shared, "As a self-contained special education teacher, SLO should not be one size fits all. My students are on a modified curriculum and need to have a modified SLO. It is not fair for my SLO to be exactly the same as a regular education teacher."

Although teachers indicated various concerns about using SLOs in the evaluation of teacher effectiveness, some did indicate that they supported the idea of using SLOs to measure student growth. Teachers understood the importance of teacher evaluation and they believed that teachers should be accountable for student learning. A few teachers indicated that SLOs might be of importance in theory, measure students' growth, and help teachers reflect on student progress. Overall, it appears that early career teachers

need some training or instruction in implementing SLOs. A few early career teachers shared that the SLOs information was overwhelming, and they did not have enough knowledge about SLOs, and they needed better orientation and trainings about the implementation of SLOs. One teacher indicated, "It would be very beneficial if you were to speak with a panel of teachers to get valuable input about SLOs and teacher evaluations."

Regarding the purposes of evaluation, teachers expected the evaluation to help them grow and improve rather than judging their teaching. Teachers hoped that the evaluators should provide them with constructive feedback and help them grow/improve instead of assigning them with a number/rating or checking the box. One teacher shared, "If you want to know how a teacher is doing, you have to actually watch them teach for an extended period of time. If you want them to improve, then you have to have actual mentorships, not evaluations." Another teacher shared, "Often, these evaluations are seen and felt as punitive. If we do things wrong, they go on our formal evaluations, rather than being used to provide teachers with areas of growth." It appeared that teachers expected that the evaluation results should be used to inform their instruction and help them improve rather than being judgmental and punitive.

4.2.7 The Highlights of the Study 2 Findings

Based on the results from various analyses in Study 2, I present the following highlights of the findings:

- Most educators disagreed on the impact of SLOs on evaluating teacher effectiveness effectively, and they agreed on the impact of classroom observations. They reported

some to substantial knowledge about SLOs and some support needed in implementing SLOs.

- Teachers and administrators had similar views of the impact of SLOs, and they reported similar levels of support needed to implement SLOs. In comparison with teachers, administrators reported more overall knowledge, as well as knowledge about student groups to be included, content to be included, and analyzing student assessment data.
- Differences in teachers' academic degrees were not associated with differences in their views of SLOs. Teachers who had a master's degree or above reported more overall knowledge about SLOs, knowledge about implementation of SLOs in the district, and instructional strategies to meet SLOs targets. Teachers with a bachelor's degree or below reported more overall support needed in implementing SLOs.
- In comparison with early career teachers, career teachers reported less positive views of the overall impact of SLOs, impact on evaluating teacher performance effectively, improving teachers' instructional practice, and informing teacher professional development. Career teachers reported more overall knowledge about SLOs and knowledge of six out of nine aspects of SLOs. Early career teachers reported more overall support needed to implement SLOs and five out of six aspects of SLOs support.
- Teachers who had three or fewer years of experience of using SLOs reported more positive views of the overall impact of SLOs, the impact of SLOs on teachers' instructional practice and promoting student learning. Teachers who had more than three years of experience of using SLOs reported more overall knowledge about

SLOs, and five out of nine aspects of their knowledge about SLOs. Teachers who had three or fewer years of experience in using SLOs reported more overall support needed and five out of six aspects of SLOs support.

- Teachers teaching different grade levels had similar views of the impact of SLOs, knowledge about SLOs, and support needed to implement SLOs.
- Teachers and administrators differed in their views of the overall impact of classroom observations, and all four aspects of SLOs impact. Teachers who had a bachelor's degree or below reported more positive views of the overall impact of classroom observations and three aspects of SLOs impact. Early career teachers reported more positive views of the overall impact of classroom observations and two aspects of SLOs impact. Teachers reported similar views of the impact of classroom observations regardless of their experience of using SLOs or the grade levels taught.
- Teachers' perceptions of the classroom observations predicted their perceptions of SLOs, and 17.3% of the variation in teachers' perceptions of SLOs could be explained by the predictors.
- Teachers expressed various concerns about using SLOs. Their concerns are related to students' test performance, assessment methods (e.g., standardized tests), time and timeline in implementing SLOs, paperwork, applicability to special education teachers, arts teachers, ESL teachers, and media specialists, lack of supervision and subjectivity in goal-setting and assessment, teaching standards, teach to the test, and lack of feedback.

4.3 STUDY 3 RESULTS

Study 3 investigated teachers' views of the impact and implementation of SLOs in the evaluation of teacher effectiveness. Through interviews with 18 schoolteachers in South Carolina, this study reported the results. The results mainly include the impact of SLOs on teaching and learning, successes and challenges in implementing SLOs, the assessment methods used to measure student growth, and whether SLOs was an additional reliable method in the evaluation of teacher effectiveness. In addition to the four key results, the study also sought to understand teachers' experience of using SLOs, the trainings of SLOs, and whether teachers felt confident in implementing SLOs.

Regarding the experience of using SLOs, teachers reported between three and nine years of experience of using SLOs. Fourteen teachers indicated that they had received various types of training, professional development, and mentoring. The training was mainly from school administrators, coaches, and mentors, and many were through professional development sessions, faculty and staff meetings, workshops, and mentoring from colleagues. Regarding the effectiveness of the training, nine teachers indicated the training was effective and helpful. For example, one teacher said, "The training is effective in that they do a good job explaining to me what the SLO means and what it requires." Another teacher shared, "The training has been comprehensive, such that I walked away with a full understanding of what was expected of me and why it was important." However, a few teachers indicated that the training was not effective. For example, one teacher shared that the training was not necessarily effective, and she was just taught how to write them, and nothing beyond that.

Teachers were asked about whether they felt confident and needed any support in implementing SLOs. Twelve out of 18 teachers explicitly expressed that they felt very confident, comfortable, and did not need any support in implementing SLOs. Five teachers explained that the SLOs procedures were simple and easy, and not hard or difficult at all. One teacher indicated, "It's not hard, it's just tedious." Three teachers shared that they had a very good and supportive school administration. For example, one teacher shared, "My school, my faculty, and my administration has done a really good job at recognizing we are not getting a lot of support from the state, so they are working hard to give us some support." Three teachers expressed that they had years of experience of using SLOs and felt confident and comfortable to use them. As one teacher shared "I feel like I'm more comfortable with the SLO after using it for several years even if it changed for the past two years." A few teachers indicated that they felt confident because they had received training of SLOs and they and their colleagues had been working together and supporting one another in the implementation of SLOs. Therefore, they did not need any support. It was important to notice that four teachers suggested that the beginning teachers might need some support in implementing SLOs. For example, one teacher shared that the first year when she was using SLOs, she was not confident and needed help especially in setting appropriate growth targets for students.

Regarding teachers' confidence and support in using SLOs, five teachers indicated that it would be helpful to have some support. They did not explicitly express whether they felt confident or not. They did emphasize the importance of having some support or professional development in the implementation of SLOs. For example, one teacher shared, "I think I'm definitely a champion for any types of professional

development and learning. I always think anyone of us need some development, hearing and seeing how other people do things and how to make it valid and reliable." Similarly, another teacher said, "I think everyone always need support. We are always learning." One teacher indicated that she and her colleagues needed some outside support in implementing SLOs, hoping that the state could help them with step-by-step videos demonstrating the SLOs procedures. Another teacher said she needed some support due to slight changes of the SLOs requirements.

Table 4.50 *Teachers' Confidence and Support Needed*

Category	Theme	Selected Coding
Teachers feeling confident and no support needed (12 teachers)	Simple SLOs process	They are very simple and easy. It's not hard, it's just tedious. I don't think it's real hard. It wasn't difficult with the SLO process. That's about it, you don't need to do anything.
	Supportive school administration	We have a very good administration at our school. Our principal and curriculum coaches are very responsive. School administration and faculty are working hard to give us some support.
	Years of experience in using SLOs	This is my third year, and I'm using the same one. I'm more comfortable after using SLOs for several years. I've done it for three years.
Teachers' need for some support (5 teachers)	General learning and PD	Everyone always needs support. I'm a champion for any type of professional development and learning.
	Specific support	I need a little help this year (due to the change of SLOs requirement). I don't feel I've been adequately prepared (due to a new kit). It would be really good for us to have some outside support (from the state).
Additional Findings (4 teachers)	Beginning teachers' need for support	I'm sure probably for beginning teachers. My mentee is currently stressed. The first year, I was not confident. I had difficulties in my first year.

4.3.1 SLOs Impact on Teaching and Learning

In the interviews, two questions were about the impact of SLOs on teachers' instructional practice and students' learning outcomes. Regarding the impact of SLOs on instructional practices, six teachers (three elementary and three middle school teachers) showed that SLOs had impact on their instructional practices, indicating that they became more aware of the learning goals and specific skills for students, they were more intentional about instructional content, and they had a better understanding of the guidelines. For example, one teacher said, "I don't think it changes everything about my teaching. But it definitely changes my instructional practices. They are more intentional, more planned and thoughtful." Teachers also suggested an association between assessment and teaching. They indicated that SLOs help them gather information in a more organized way, and they made plans and adjusted their teaching based on assessment results. One teacher shared, "When we do SLOs, we use pre assessment and post assessment. So, based on our pre assessment, we teach them and make sure they understand the information."

Twelve teachers (six high school teachers, three elementary school teachers, and three middle school teachers) indicated that SLOs did not impact/change their instructional practice. They shared that they already had the knowledge about standards, learning goals, and assessment, and they taught in the same way with or without SLOs. For example, one teacher said, "I feel like it just makes you write down the goals, I already have goals, I already know what they need to know, so it won't change." Some pointed out the issues of the assessment used to measure student growth including repeating the tests, and the consistency and validity of the assessment. For example, one

teacher said, “I wouldn’t say that it necessarily changes the way I teach my students. Honestly, I gave them a pretest, a mid-year test, and an end-of-the-year test, they are the same.” Similarly, another teacher shared “One way it does change, but this is a negative change because I have to give additional test of that sort. So, it actually ended up taking a little away from my instruction.” In addition, a few teachers considered SLOs as “just another way to complete a long-range plan,” or “another hoop to jump through” in order to stay certified.

Table 4.51 *Impact of SLOs on Teaching*

Category	Theme	Selected Coding
SLOs having Impact on Teaching (6 teachers)	Awareness and understanding	It makes me more aware of what my goals are. It makes me more aware of what specific skills students need to improve on. I’m much more intentional about instructional content. It gives me a better understanding of guidelines.
	Associations between assessment and teaching	I make adjustment of teaching my students. I have to teach the material because I have a test again. It makes me cover the learning targets effectively.
SLOs having No Impact on Teaching (12 teachers)	Having knowledge	I know what the standards are. I pretty much know what I’m doing. I already know what they need to know. I know the class.
	Same way of teaching with or without SLOs	With or without it, you want your students to achieve. I was already teaching, so I don’t have to alter it. It’s the same thing that I would have done with or without SLO.
	Assessment issues	It is a teacher generated exam, there is no consistency, there is no validity. I gave then a pretest, mid-year test, and a posttest, they are the same. It’s an assessment I generally give anyway. The outcome has always maintained the same.

Regarding the impact of SLOs on student learning, eight teachers (four elementary and four middle school teachers) expressed that SLOs had impact on student learning outcomes. Three teachers indicated without hesitation that SLOs definitely impacted their students' learning outcomes, indicating that SLOs helped better outcomes for students, students had met or exceeded expectations, and there was improvement in their students' writing skills. For example, one teacher said, "Since we use it, their writing has become more detailed, they have improved in their writing conventions, and their paragraph writings also improved." Another teacher shared, "They have met and exceeded the expectation that I placed on them. So, I learned that I get expect more and more, they can get more and more. So, I think it's a very good thing."

However, five teachers expressed that SLOs might have some impact on certain skills of students, and SLOs might not cause substantial change on student learning outcomes. As one teacher shared, "It does support their mastery in a particular skill. But there is a trade off because SLOs only require one standard, obviously we have many standards. If we spend plenty of time on that one particular skill. In order for students to master that particular skill, you may have to sacrifice another skill along the way. It's a tradeoff." Similarly, another teacher said, "I think, for that assessment, yes. I teach science, if it's something focusing on something other than science, it may not change their outcomes. But for that assessment, yes." Other teachers were a little hesitant about the impact of SLOs on student learning outcomes. One teacher said, "I don't know if it changes their outcomes, … I always looked back to see their growth, it's definitely there, I feel like SLOs cause you to pause and really look at the growth."

Table 4.52 *Impact of SLOs on Learning*

Category	Theme	Selected Coding
SLOs having Impact on Learning (8 teachers)	Definite impact	It definitely helps better outcome for them. They have met and exceeded the expectations. I have noticed the improvement in their writing skills.
	Some impact on certain skills	It does support their mastery in a particular skill. For the assessment, yes. I have always seen growth in my students. It does not change learning outcomes any more than other learning standards I teach. Probably not as much as they are intending SLOs to change student outcome.
SLOs having No Impact on Learning (10 teachers)	Impact of effective teaching on student learning	The only thing that changes students learning outcomes is effective teaching. I already do as a teacher who utilizes good practices. I don't really credit SLO with their achievement as much as what I am doing in the classroom. Any good teachers want student to perform well and be successful. I teach them all six units regardless of SLOs. I teach my students the content and skill, and they need to know anyway.
	SLOs focus on one standard	I try to be effective in everything I do, not just this one task. I teach all subjects, but SLO is simply geared on one. We pick a standard.

Ten teachers (all six high school teachers, two elementary school teachers, and two middle school teachers) indicated that SLOs did not impact/change student learning outcomes. They believed that effective teaching had an impact on student learning. For example, one teacher explicitly stated, "In my experience, the only thing that changes student learning outcomes is effective teaching." Another teacher said, "My kids have done really really well. I don't really credit the SLO with their achievement as much as what I am doing in the classroom." Some teachers considered that SLOs did not impact student learning because SLOs focused on only one or a few standards. They shared that focusing on one standard did not improve students' overall learning outcomes. For

example, one teacher said, “They are growing in one area, but are actually causing a deficiency in another area. So, this is something you have to think about.” In addition, a few teachers shared that the assessment was too subjective, students made the decision whether they would meet the targets or not, and the SLOs was just a “formality.”

4.3.2 Successes and Challenges in Implementing SLOs

In the interview, two questions were about the successes/benefits and challenges/obstacles in the implementation of SLOs. Regarding the successes or benefits of using SLO, eleven teachers (all six elementary school teachers, three middle school teachers, and two high school teachers) showed that SLOs have some benefits. SLOs help track students’ growth or progress. Through using SLOs, teachers can see where the students start with, and how much progress they can make. One teacher said, “SLO shows us how to track students’ progress. I just think seeing the kids making progress, that’s the benefit, make sure that teachers are doing what they need to do.” Teachers indicated that using SLOs influenced their teaching and made them be more reflective of their teaching. One teacher said, “The SLOs make teachers reflect on what they teach, the reflection aspect of SLOs is powerful.” Some teachers suggested that using SLOs helps hold them accountable for student learning and making progress. One teacher shared, “I think the benefit from using SLOs is holding each teacher accountable for their students’ progress.” In addition, teachers showed that using SLOs provided them opportunities to study and use student data to inform their instruction. Other successes or benefits of using SLOs included helping teachers set learning targets/goals for students, teachers’ making decisions about standards, and partially measuring teacher effectiveness.

Table 4.53 *Successes/Benefits of Using SLOs*

Category	Theme	Selected Coding
Some Benefits (11 teachers)	Tracking student growth or progress	SLO shows us how to track students' progress. You could see where the students start with and where are they to end to show growth. I am tracking student growth. I can test what they know, so I can push them to learn.
	Teaching and teacher reflection	The reflection aspect of SLOs is powerful. It does provide opportunities for teachers to reflect on what they've taught. It definitely made me thoughtful in my instructional practices. It influences my teaching. Students are guaranteed to get exceptional instruction in that area.
	Accountability	SLO is holding each teacher accountable for their students' progress. It holds me more accountable as a teacher. Every teacher should be working hard toward the success of all students.
	Using data	Teachers look more about their data. It's good for teachers to review data to learn where their students are. Teachers are forced to study their data. Prove to my administration that I/m tracking data.
No Benefits (7 teachers)	Various explanation	I don't think it determines your effectiveness. Once you are a veteran teacher, it's just the distraction because you know what you need to do anyway. Administrators do not read the SLOs. Teachers just want to check marks. Teacher has the autonomy to create the growth targets. I don't feel like it really helps me to be better prepared.

Although many teachers considered that SLOs had some benefits, seven teachers (three middle school teachers, and four high school teachers) stated that there were no benefits of using SLOs in the evaluation of teacher effectiveness. They provided various explanations, indicating that SLOs cannot determine teacher effectiveness, administrators did not read the SLOs due to time constraints, it's a distraction, and teachers just wanted to "check marks." One teacher summarized, "So when you are evaluating a teacher

based on student success, it's sort of an old conundrum. ... If you teach to the test, sure students are going to have that information. Ultimately, they have to use their knowledge to score and show that growth. The other conundrum is that SLOs are written by the teacher. So, a teacher who is very smart will create SLOs in a way they know every student will have some kind of growth. In addition to that, the teacher has the autonomy to create the growth targets and adjust the growth targets as the year progresses. So, I feel that for teacher effectiveness, I don't feel like SLO itself can monitor or evaluate how effective a teacher is."

Teachers were also asked about the challenges or obstacles in using SLOs. All 18 teachers indicated some types of challenges or obstacles in using SLOs. It appeared that time had been considered as a major challenge or obstacle for some teachers. Eight teachers indicated that time is a big challenge/obstacle, sharing that it took a long time to write out the SLOs plans and complete all the paperwork. Some administrators required their teachers to turn in the SLOs within a short period of time. For example, one teacher said, "The time when it's due for all the paperwork has been really really hard on the teachers." Similarly, another teacher shared, "Sometimes, I don't feel like I know my students long enough to really make those judgements, or assess, or plan on the SLO that I really want to accomplish because I don't know everything about where they are, developmentally, academically, and emotionally." In addition to the time spent in completing the SLOs, two music and visual arts teachers indicated that inadequate instructional time may create some challenges. For example, one music teacher said, "I teach for 36 days. SLO is supposed to cover 180 days. But, because I only see my students only one day a week, a year's growth for me is actually 36 days. That does

create a bit of problem to translate something that is supposed to be over 180 days to accomplish in 36 days." One visual arts teacher said, "There's a lot of instructional time lost, maybe there is a school assembly or other events that create instructional time loss. …That's one reason why the instruction needs to be objective. That is definitely a challenge."

Table 4.54 *Challenges/Obstacles of Using SLOs*

Theme	Selected Coding
Time	It takes a long time to write out your SLOs plan. The timeline is challenging. Time is the obstacle to creating and using SLOs. Having a time to give the pre and post assessment can be difficult. Just the timelines of the SLOs is very frustrating. That's hard to really judge your students early on. The only challenge is it's time consuming. There is a lot of instructional time loss.
Paperwork	The major challenge is the mass of paperwork that it involves. There are a lot of papers that we have to do. They have paperwork due this weekend. I feel like it is more like a hoop, some extra paperwork to do.
Teachers' autonomy (lack of supervision)	Because they are teacher created, I think it's the biggest obstacle. No one even check your data, you can immediately just change your data to meet your SLOs expectations. I think the only obstacle is that every teacher does it differently.
Missing other standards	Maybe they might miss another standard. If you are so focused on the goals of your SLOs,…are you going to change the others?

Another challenge that teachers indicated was about the paperwork involved to complete SLOs. For example, one teacher said, "The major challenge is the mass of paperwork that it involves. I don't know if you see it, but it's a lot of paperwork, a lot of time involved." Teachers shared that teachers had a lot of autonomy, and it could be an obstacle due to the lack of supervision. Teachers could set goals that are easy for their students to achieve so this will be in favor of their evaluation scores. There was no

supervision for the assessment data. A few teachers indicated that some of their colleagues modified student data to meet the expectations. Also, teachers were implementing SLOs differently, and there was no common standards to evaluate teacher effectiveness. In addition, missing other standards was also a challenge. Teachers might be so focused on the one standard that they use for their SLOs that they may neglect other standards that should be taught as well. As one teacher said, "I think the biggest obstacle is that you can definitely lose other standards that need to be taught because you spent extra time on the SLO material."

Teachers also expressed other challenges and obstacles in using SLOs. Teachers did not get any feedback about SLOs and student growth. There were some changes of SLOs requirements, and teachers had to learn and meet the requirements every year. For example, one teacher said, "They change every year about how to turn in the paperwork, which means we have to learn a new different format. … On top of doing the paperwork, I have to learn how to do it. That's another stack. That's very frustrating and that's very time consuming. It often happens the last minute. It's very challenging." In addition, teachers thought the group of students they teach affected their SLO scores, and students looked at SLOs as unimportant, which would affect their evaluation results. Furthermore, schools implemented SLOs differently, and it created some confusions for the teachers who switched schools. Another challenge was the assessment methods teachers used to measure student growth, which might not be objective and reliable to accurately measure student growth.

4.3.3 SLOs Assessment Methods

Teachers described the assessment methods they used in measuring student growth in teacher evaluation. Table 4.55 presents the assessment methods that teachers used to measure student growth in the evaluation of teacher effectiveness. Overall, there were two major methods: standardized assessment and teacher designed assessment. Among the 18 teachers, four teachers (three elementary school teachers and one middle school teacher) reported using standardized assessment, and 14 teachers (three elementary school teachers, five middle school teachers, and all six high school teachers) used teacher designed assessment. For the teacher designed assessment methods, some were designed by individual teachers, some were designed by a team of teachers who taught the same subject area, and some were designed by administrators and specialists at the school. Schools had variations in decision making regarding the assessment methods used in implementing SLOs. Some teachers shared that they could decide the assessment methods, while other teachers indicated that it was the school administration that made decisions on the assessment methods used in implementing SLOs.

Some teachers reported having been using assessment methods designed by individual teachers, a team of teachers, or school administrators. There were various assessment methods including tests, performance tasks, multiple-choice questions, reading comprehension, writing prompts, MasteryConnect (a platform for teachers to develop assessment items), and other assessment methods. Teachers had overall positive views of the assessment methods. For example, one teacher indicated that she was using writing prompts, sharing “I think it is a good measure, it’s like a snapshot, it’s one piece of the puzzle I use to track their growth, and I do think it’s effective.” Another teacher

also had positive views of the performance tasks she was using, indicating "I think this is a reliable method to capture my students' growth over time." A few teachers reported negative views of the assessment methods showing that the methods were not reliable, cannot judge student growth, and did not see accountability. Teachers had major concerns about the objectivity of assessment in measuring student growth, especially when they used the teacher-designed assessment methods. In addition to one specific assessment method, teachers shared that they were actually using multiple methods in measuring student growth. For example, one elementary teacher said, "We use MAP, and we also use the math curriculum to create our assessment," Another teacher said, "I use performance tasks, I use conferences and interviews with my students, I also take anecdotal data of the questions they ask."

Four teachers indicated that they used standardized test results to measure their students' growth in teacher evaluation. Three used MAP and one used a new formative assessment method named FastBridge. For the use of MAP test results, two teachers said it's either a district decision or school decision, and one said it's decided by the team of teachers. They also had mixed views of this method. Two believed that it was a good idea to use MAP and it was an accurate method to track student growth. One teacher disagreed and considered the method inaccurate. The teacher who was using FastBridge showed some hesitation in using this new method, indicating that it posed some difficulty in using the data for teacher evaluation due to the score reporting scale. It appears that the major concern of using standardized assessment is that students' test scores might not reflect the effectiveness of teachers.

Table 4.55 *Assessment Methods in Measuring Student Growth*

Category	Methods	Decision	Teacher View
Teacher designed assessment (14 teachers)	Performance tasks	Teacher designed	I think this is a reliable method to capture my students' growth over time. I do think this method is reliable. I think it's a good indicator of what's happening in the classroom.
	Writing prompts	Assistant principal and reading specialist	I think it is reliable, I think it is just one piece of information used to measure their growth.
	Multiple-choice questions	Teacher or teacher team designed	I do not think the assessment method is reliable to evaluate student growth. The method is reliable in a sense that I just measure the same thing every time I use it.
	Reading comprehension	Teacher designed	I don't know
	MasteryConnect	Teacher or teacher team designed	The data I receive from the platform does help me understand who is below mastery, approaching mastery, or has attained mastery. I guess you can see their growth. It's not reliable because teachers set up their own thing.
	Test	Teacher team designed	Not reliable. I haven't really seen any accountability.
	Multiple assessment methods	Teacher and teacher team designed	I think they are reliable. I don't think I could effectively judge growth. For this unit of study, I would say it's a reliable method. I think the method is reliable.
Standardized tests (4 teachers)	MAP	Teachers, school, or district decision	We decided this is a good idea. I like MAP,…what we can see is their growth. It isn't really an accurate method to show growth.
	FastBridge (First time)	School or district decision	They are letters and levels for the reading,… It's harder for us to use that as data in SLOs.

4.3.4 SLOs as an Additional Evaluation Method

Teachers were asked whether SLOs was an additional reliable method in the evaluation of teacher effectiveness. Half of the 18 teachers (five elementary, three middle, and one high school teachers) agreed that SLOs was an additional reliable method used in the evaluation of teacher effectiveness. Half of the teachers (one elementary, three middle, and five high school teachers) did not consider SLOs an additional reliable method in teacher evaluation.

The teachers who thought of SLOs as an additional reliable method in teacher evaluation indicated that SLOs was a good method to track student growth, and it provided a snapshot of students' growth over time. For example, one teacher said, "I think if it's used correctly, SLO is a reliable method. Because it helps show how students grow in the classroom." Some teachers also considered SLOs as a good indicator of teaching, and an effective way to measure teacher performance. As one teacher shared, "I think it's a good indicator of what I do in my classroom." Some teachers viewed SLOs as an additional reliable method when used in combination with other methods. They thought the classroom observations were not adequate in teacher evaluation because classroom observation was just "a small glimpse into what happens in the classroom year-round." One teacher said, "It is impossible for people who do observations to see everything that goes on every day. They see a point at a time." Teachers seemed to like the idea of employing multiple methods in evaluating teacher effectiveness. One teacher shared, "In combination with classroom observations, and student feedback, and peer feedback, I think it's not a bad idea." Another teacher indicated, "I do feel like there are

several pieces that should be considered when evaluating a teacher's effectiveness. But I do feel like SLO is a good piece of that."

Table 4.56 *SLOs as an Additional Reliable Method*

Category	Theme	Selected Coding
Agreement (9 teachers)	Tracking student growth	I think it's good to show student progress. It gives you a snapshot of the growth of your students. It helps show how students grow in the classroom.
	Good indicator of teaching	It's a good indicator of what I do in my classroom. It is an effective way to measure teacher performance. It's a good thing to have as part of an understanding about whether a teacher is teaching effectively.
	Combination of multiple methods and inadequacy of classroom observation	In combination with classroom observation, and student feedback, and peer feedback, I think it's not a bad idea. There are several pieces that should be considered when evaluating a teacher or his/her effectiveness. It is impossible for people who do observations to see everything that goes on every day.
Disagreement (9 teachers)	Student performance	They (students) don't really care. They are not willing to learn, they are not prepared. They are trying to compare this year's class with last year's class, with completely different students.
	Assessment issues	I don't think it is very reliable because it's too subjective. They are not taking the same test at the beginning and in the end, and they are not tracked. I don't think SLOs are reliable unless we have an objective test. It's a small snapshot of what students are actually being exposed to learning in the class.
	Observation	For the classroom observation, I think it's adequate to judge a teacher. When they come to observe me, they are fabulous with giving me feedback.
	Teacher autonomy (lack of supervision)	You can craft that SLO to make you look like the most effective teacher. If a teacher is fudging the numbers, creating an easily achievable growth targets. Honestly, I have heard teachers just making the numbers up.

Although teachers appeared to be in favor of using SLOs as an additional reliable method in teacher evaluation, they pointed out that SLOs should carry small weight, and count as a small piece in the evaluation of teacher effectiveness. They believed that the student group played an important role in their academic growth. There were various factors that could influence student learning. One teacher shared that one of her students had a learning disorder and five students had ADHD (not medicated), and the assessment results of these students were counted as the evaluation of her performance. Another teacher showed that students' home life, attendance, and home support come into play in determining students' achievement and growth. It was not fair to evaluate teacher performance based on student performance. In addition, a few teachers emphasized that SLOs should be implemented correctly, and teachers should be honest and set appropriate learning targets for their students. For example, one teacher said, "Some teachers do not take it seriously. They pick an easy standard to assess, and then teach to the test such that all kids show growth. For these types, SLO is not effective."

Half of the teachers indicated that SLOs was not an additional reliable method in the evaluation of teacher effectiveness. They considered that student performance in the test were affected by many reasons. There were various issues in the assessment used to measure student growth. Some teachers thought SLOs was not a reliable method because the assessment method was not objective or reliable and cannot reflect student learning. One teacher described her experience of the SLOs conference. She was given an evaluation score that was not based on her students' assessment results. She said, "That assessment was very unreliable, not based on real data." A few teachers considered the classroom observation was adequate in the evaluation of teacher effectiveness, indicating

"With observational rubrics, the administration, when they come to observe me, they are fabulous with giving me feedback." Teachers did not have any feedback on assessment results. For example, one teacher said, "We don't get any feedback of the test anyway. They just give us the score. At the end, I have no clue what they are good at and what they missed." Another teacher also thought the teachers should have some feedback and support from the state. In addition, evaluators' expertise seemed to be a concern for some teachers. One world language teacher indicated that evaluator's expertise was her concern. She said, "It's a foreign language, how do you really know what's going on there. If the teacher asks some higher level thinking questions, how do you know? If you don't know the language, you don't know."

In addition to all the key questions discussed above, teachers were also asked to share additional thoughts regarding SLOs or teacher evaluation in general, fourteen teachers provided their thoughts. Teachers had mixed feelings about using SLOs in teacher evaluation. Some teachers seemed to be in favor of using SLOs in teacher evaluation. They considered that teachers should be held accountable for student learning, and they appreciated coaching from their administrators in the process of teacher evaluation. They seemed to like the idea of measuring student growth rather than using student achievement in teacher evaluation. They believed that SLOs gave another snapshot of what teachers were doing, and they considered SLOs as an effective way to evaluate teachers. For example, one teacher said, "I've taught for 26 years, I've been evaluated in a lot of different ways, to me, this has been one of the most effective ways."

At the same time, teachers reported a lot of concerns about using SLOs in teacher evaluation. There were a lot of variability of linking student performance to teacher

effectiveness. The group of students (e.g., gifted and talented students, ELL) affected their academic growth, and thus affected teachers' evaluation scores. There were many factors (e.g., students having a bad day, or lack of support from family) that could impact student test performance, and it was not a fair judgement of teaching based on student test performance. For example, one teacher shared, "I think sometimes that the teacher evaluation part is frustrating for us because we don't want our job performance to be reflected by a student's inability to focus on a test." Teachers considered using student growth measure in evaluating teacher effectiveness as stressful, intimidating, scary, paranoid, and frustrating. One teacher said, "When the SLOs were approached to everybody, there was a lot of panic." Some teachers also considered SLOs as "time-consuming," "all about paperwork," and "just another fed evaluation and will fade away in five to 10 years."

In addition, teachers had concerns about the disconnect/inconsistency between classroom observation results and SLOs results, the validity and reliability of the assessment used in measuring student growth, issues of teacher retention/teacher shortage, and the fidelity of implementing SLOs. The lack of feedback from the state regarding teacher performance based on SLOs was also a concern for teachers. For example, one teacher shared, "We need some feedback that is supportive, that could make us feel supported, feel valued. When we feel valued, we are more likely to improve." Overall, a few teachers indicated the improvement of teacher evaluation. One veteran teacher with more than 30 years of teaching experience shared, "I have seen many evaluations, I do think we are improving."

4.3.5 Summaries of the Interview Findings

Based on the qualitative analysis of the interviews with the 18 teachers, I summarized the findings. Overall, fewer than half of the teachers thought that SLOs had some positive impact on teaching and learning. They shared that using SLOs made them be more aware of their teaching, and the assessment results could inform their teaching. Some teachers considered that using SLOs contributed to student learning outcomes and students made notable growth. However, more than half of the teachers did not think that SLOs had any positive impact on teaching and learning. They shared that they had the knowledge about teaching and would teach in the same way with or without SLOs. They also had concerns about the assessment methods used in SLOs. Some teachers thought that SLOs focused on only one standard, and effective teaching contributed to student learning.

In the implementation of SLOs, teachers shared some benefits of SLOs. They said SLOs could be used to track student growth, made teachers reflect on their teaching, held teachers accountable for student learning, and involved data using to better understand student learning. However, teachers indicated that there were some challenges/obstacles in implementing SLOs. The process was very time-consuming, the timeline was not appropriate, and it required a lot of paperwork. SLOs only required the assessment of one standard, and many standards were missing while some teachers were so focused on the one standard.

In addition, some teachers had the concern that teachers had too much autonomy, there was a lack of supervision, and the assessment results of students' growth were subjective. Regarding the assessment methods used in measuring student growth, many

teachers reported that they used teacher-designed assessment methods including multiple-choice tests, reading comprehension, essay writing, performance assessment, and multiple assessment methods. A few teachers also reported using standardized test (e.g., MAP) in measuring their students' growth.

Overall, most teachers shared that they felt confident in using SLOs, indicating that the SLOs process was simple and easy, they had a supportive administration at their schools, and they had years of experience of using SLOs. A few teachers reported that they needed some support due to changes in SLOs requirements. Some indicated the importance of professional development. The teachers believed that the new teachers should need more support in understanding and implementing SLOs.

Although most teachers showed confidence in implementing SLOs, only half considered SLOs as an additional reliable method in teacher evaluation. They argued that SLOs could be used to track student growth, was a good indicator of teaching, and was reliable when used with other evaluation methods (e.g., classroom observation). The other half of the teachers did not consider SLOs as a reliable method, and they indicated concerns about student performance in assessment, assessment issues, lack of supervision, and subjectivity. Some teachers thought that classroom observation alone was adequate in evaluating teacher effectiveness. It appears that teachers had different reasons for making decisions about whether SLOs is an additional reliable method in the evaluation of teacher effectiveness. The major issues lie in the implementation of SLOs in teacher evaluation.

Table 4.57 *Summaries of Interview Findings*

Aspect	Findings
SLOs Impact on Teaching	• One-third of the teachers reported positive impact of SLOs on teaching. They were more aware of their teaching and used assessment results to inform their instruction. • Two-thirds of the teachers reported no impact of SLOs on teaching. They already had the knowledge and taught in the same way with or without SLOs, and they also had concerns about the assessment methods.
SLOs Impact on Learning	• Fewer than half of the teachers reported positive impact of SLOs on learning. Students had better learning outcomes and made progress. • More than half of the teachers reported no impact of SLOs on learning. They believe effective teaching can impact student learning. SLOs only focused on one standard.
Successes in SLOs Implementation	• Tracking students' growth/progress • Providing opportunities for teacher reflection • Holding teachers accountable for student learning • Using data to understand student learning
Challenges in SLOs Implementation	• Time consuming and timeline • Paperwork • Teachers having too much autonomy (lack of supervision) • Focusing on one standard but missing other standards
Assessment Methods	• Many teachers reported using teacher-designed assessment methods: multiple-choice test, reading comprehension, essay writing, test, performance assessment, and multiple assessment methods. • A few teachers reported using standardized test (e.g., MAP)
Teacher Confidence	• Most teachers reported feeling confident in using SLOs. • A few reported needs for some support due to changes of SLOs. • Some considered new teachers need more support.
SLOs in Teacher Evaluation	• Half of the teachers considered SLOs as an additional reliable method in teacher evaluation. SLOs can track student growth, was a good indicator of teaching, and was reliable. • Half of the teachers did not consider SLOs as a reliable method because they had concerns about student performance, teacher autonomy (lack of supervision), subjectivity, and classroom observation.

4.3.6 The Highlights of the Study 3 Findings

Based on the results from various analyses in Study 3, I present the following highlights of the findings:

- Teachers had mixed views about SLOs' impact on teaching and learning. Fewer than half of the teachers considered that SLOs have a positive impact on teaching and learning. Elementary school teachers were more positive of SLOs.
- SLOs benefits included tracking students' progress, encouraging reflective teaching, holding teachers accountable for student learning, and using student assessment data to inform instruction.
- SLOs challenges included time, timeline for SLOs submission, lots of paperwork, too much teacher autonomy and lack of supervision in goal setting, issues in assessment method, and missing some standards.
- Most teachers felt confident in using SLOs. Support should be provided to new teachers or teachers whenever there are some changes of the SLOs implementation policy or requirements.
- Half of the teachers considered SLOs as an additional reliable method in teacher evaluation.

4.4 STUDY 4 RESULTS

Study 4 examined 275 TAP school teachers' evaluation scores based on SLOs and classroom observations. This study focused on examining whether teachers' SLO scores could better differentiate teacher performance in comparison with their classroom observation scores. In addition, this study also examined the relationship between teachers' SLO scores and their classroom observation scores. In the analysis of teacher's

SLOs score and observational scores, districts, school poverty, school type, and teacher type were taken into consideration.

4.4.1 Differentiation of Teacher Performance

To examine the differentiation of teacher performance based on their SLO scores and classroom observation scores, I calculated the frequencies of each score point (Table 4.58). Based on the evaluation using SLOs, teachers' minimum SLOs score was 1 point, and 9.1% of the teachers obtained 1 point. Teachers' maximum SLOs score was 5 points, and 26.9% of the teachers obtained 5 points. Overall, about 36% of the teachers obtained either 1 point or 5 points, and about 64% of the teachers obtained between 2 and 4 points. Based on the evaluation using classroom observations, about 1.1% of the teachers obtained a minimum score of 2 points, and 1.1% of the teachers obtained a maximum score of 4.5 points. A large percentage (97.8%) of teachers obtained between 2.5 and 4 points.

Table 4.58 *Frequencies of SLO Scores and Classroom Observation Scores (N=275)*

SLO Scores			Observational Scores		
Score	N	%	Score	N	%
1.0	25	9.1	2.0	3	1.1
2.0	24	8.7	2.5	29	10.5
3.0	67	24.4	3.0	104	37.8
4.0	85	30.9	3.5	96	34.9
5.0	74	26.9	4.0	40	14.5
			4.5	3	1.1

The frequencies of teachers' SLO scores and their classroom observation scores revealed that there was notable differentiation of teacher performance based on their SLO scores, with certain percentages of teachers on each score point or performance level. However, little differentiation was identified based on their classroom observation scores, with about 98% of teachers obtaining between 2.5 and 4 points. Comparatively, SLO

scores appeared to be able to better differentiate teacher performance in comparison with classroom observation scores.

4.4.2 Relationship of Teachers' SLO Scores and Classroom Observation Scores

To understand teachers' performance based on their SLO scores and classroom observation scores, I calculated means and standard deviations. Overall, teachers had a mean SLO score of 3.58 and a mean classroom observation score of 3.27. A paired sample t-test revealed that there was a statistically significant difference between teachers' SLO scores and their classroom observation scores ($p < .01$), with a small effect size (Cohen's $d = 0.24$). The standard deviation of teachers' SLO scores was 1.23, which was much larger than the standard deviation of their classroom observation scores (.47). It further indicated teachers' SLO scores were more spreading out in comparison with their classroom observation scores. To explore the relationship between teachers' SLO scores and classroom observation scores, I calculated Spearman's rho correlations considering the ordinal data characteristics of the scores. The correlation coefficient was .12, which was statistically significant ($p = .04$). However, the magnitude of the relationship was small. It suggests that there is a small positive association between teachers' SLO scores and their classroom observation scores. Teachers who had a higher SLO score tended to have a higher classroom observation score.

4.4.3 Teachers' SLO Scores and Classroom Observation Scores by Factors

To explore whether teachers from different districts obtained similar or different SLO scores and classroom observation scores, I calculated means and standard deviations by district (Table 4.59). Teachers' SLO score means ranged from 2.17 (Chase) to 3.77 (Mills), and their observational score means ranged from 3.07 (Moon Mountain) to 3.41

(Mills). Further Kruskal-Wallis H test revealed that teachers' SLO scores were significantly different based on their districts, H = 25.00, $p < .01$, with a medium to large effect size ($\eta^2 = .08$). Teachers observational scores were also significantly different based on their districts, H = 26.46, $p < .01$, with a medium to large effect size ($\eta^2 = .09$).

Table 4.59 *SLOs and Classroom Observation Scores based on District (N=275)*

District	N	SLO Scores		Observational Rubric Scores	
		Mean	Std. Dev	Mean	Std. Dev
Chase	23	2.17	1.37	3.13	0.41
Mills	153	3.77	1.03	3.41	0.43
Moon Mountain	22	3.41	1.14	3.07	0.56
Summer View	77	3.66	1.31	3.10	0.46
Total	275	3.58	1.23	3.27	0.47

To understand whether teachers at different school types had different SLO scores and classroom observation scores, I calculated means and standard deviations based on school types (Table 4.60). For all 16 TAP schools involved in this study, five schools with both elementary and middle school types, or both middle and high school types were excluded from the analysis. Eleven schools that were clearly classified as primary/elementary, middle, or high school were included. In this analysis, 67 teachers taught at primary or elementary schools, 19 teachers taught at middle schools, and 103 teachers taught at high schools. Teachers' SLOs score means ranged from 3.53 (high school) to 4.11 (middle school), and their observation score means ranged from 2.97 (middle school) to 3.41 (primary/elementary school). Further Kruskal-Wallis H test revealed that teachers' SLO scores were not significantly different based on school type, H = 3.68, $p = .16$. Teachers' classroom observation scores were significantly different based on school type, H = 18.75, $p < .01$, with a medium to large effect size ($\eta^2 = .09$).

Table 4.60 *SLOs and Classroom Observation Scores based on School Level (N=189)*

School Type	N	SLO Scores		Observational Scores	
		Mean	Std. Dev	Mean	Std. Dev
Primary/Elementary School	67	3.81	1.03	3.41	0.51
Middle School	19	4.11	1.10	2.97	0.49
High School	103	3.53	1.31	3.16	0.38

To explore whether school poverty impacted teachers' SLO scores and classroom observation scores, I calculated means and standard deviations (Table 4.61). Fifty-four teachers were from schools that had a poverty index of 60% or below, 127 teachers were from schools that had a poverty index between 60% and 80%, and 94 teachers were from schools that had a poverty index of 81% or above. Teachers from low poverty schools obtained the lowest SLOs score of 2.81, and teachers from moderate poverty schools obtained the highest SLOs score of 3.87. Teachers from low poverty schools obtained the lowest classroom observation score of 3.07, teachers from high poverty schools obtained the highest classroom observation score of 3.47. It appears that teachers' SLO scores varied substantially based on the school poverty levels. However, their observation scores were slightly different based on school poverty. The one-way ANOVA revealed that teachers' SLO scores were significantly different based on school poverty, $F = 15.69$, $p < .01$, with a medium to large effect size ($\eta^2 = .10$). Teachers observational scores were significantly different based on school poverty, $F = 15.45$, $p < .01$, with a medium to large effect size ($\eta^2 = .09$).

Table 4.61 *SLOs and Classroom Observation Scores based on School Poverty (N=275)*

School Poverty	N	SLO Scores		Observational Scores	
		Mean	Std. Dev	Mean	Std. Dev
Poverty Index (60% or Below)	54	2.81	1.49	3.07	0.39
Poverty Index (60% - 80%)	127	3.87	1.20	3.21	0.41
Poverty Index (80% or Above)	94	3.62	0.87	3.47	0.52

To explore whether different types of teachers obtained similar or different SLO scores and classroom observation scores, I calculated means and standard deviations (Table 4.62). In this analysis, most teachers (83%) were career teachers, and some were master teachers (8%) or mentor teachers (8%). Master teachers obtained the lowest SLOs score of 2.91, and mentor teachers obtained the highest SLOs score of 3.96. Career teachers obtained the lowest observation score of 3.18, and master teachers obtained the highest observation score of 3.83. Further Kruskal-Wallis H test revealed that teachers' SLO scores were significantly different based on teacher type, $H = 6.68$, $p = .04$, with a small to medium effect size ($\eta^2 = .02$). Teachers observation scores were significantly different based on teacher type, $H = 52.26$, $p < .01$, with a large effect size ($\eta^2 = .18$).

Table 4.62 *SLOs and Classroom Observation Scores based on Teacher Type (N=275)*

Teacher Type	N	SLO Scores		Observational Scores	
		Mean	Std. Dev	Mean	Std. Dev
Career Teachers	229	3.61	1.16	3.18	0.42
Master Teachers	23	2.91	1.65	3.83	0.47
Mentor Teachers	23	3.96	1.26	3.61	0.40

4.4.4 The Highlights of the Study 4 Findings

Based on the results from various analysis in Study 4, I present the following highlights of the findings:

- Teachers' SLO scores could better differentiate teacher performance in comparison with their classroom observation scores.
- There was a small positive association between teachers' SLO scores and their classroom observation scores.
- Teachers' SLO scores and their classroom observation scores demonstrated notable variations among districts, schools of different poverty levels, and different types of

teachers. Teachers from low-poverty schools had the lowest SLO scores and classroom observation scores. Mentor teachers had the highest SLO scores, and master teachers had the highest classroom observation scores.

CHAPTER 5

DISCUSSION

Through cross-checking the results from the four studies, I discovered very interesting and informative findings about educators' views of the impact of SLOs, their knowledge about SLOs, the support needed to implement SLOs, and their views of the impact of classroom observations. Table 5.1 compares the findings from Study 1 and those from Study 2. Table 5.2 illustrates the findings from Study 3 and those from Study 4. The findings of the four studies are discussed from four aspects: educators' views of SLOs, educators' views of classroom observations, the SLOs implementation, and teachers' evaluation scores.

5.1 EDUCATORS' VIEWS OF SLOS

Educators reported very different views of the impact of SLOs on evaluating teacher effectiveness effectively before and after the full implementation of the teacher evaluation system in South Carolina. Before the full implementation of the system, between 64.5% and 78.3% of the teachers and between 90% and 95% of the administrators agreed on the four aspects of the impact of SLOs. However, after the full implementation of the system, between 25.1% and 36.8% of the teachers and between 43.8% and 56.3% of the administrators agreed on the four aspects of the impact of SLOs. Comparing educators' views before and after the full implementation of the evaluation system, there are large decreases of the percentages of both teachers and administrators who agreed on the impact of SLOs.

Table 5.1 *Highlights of the Findings from Study 1 and 2*

Aspect	Study 1 Highlights	Study 2 Highlights
Educators' Views of SLOs	Most educators agreed on the impact of SLOs, reported some to substantial knowledge about SLOs, and needed some support to implement SLOs.	Most educators disagreed on the impact of SLOs, reported some to substantial knowledge about SLOs, and needed some support to implement SLOs.
Comparing teachers' and administrators' views of SLOs	In comparison with teachers, administrators reported more positive views of SLOs. Teachers and administrators had similar levels of knowledge about SLOs and support needed.	Teachers and administrators had similar views of SLOs and similar levels of support needed. Administrators had more knowledge about SLOs.
Teachers' views of SLOs based on degrees	Differences in teachers' academic degree were not associated with their views of SLOs, knowledge about SLOs, or support needed to implement SLOs.	Differences in teachers' academic degree were not associated with their views of SLOs. Teachers who had a master's degree or above reported more knowledge about SLOs and less support needed to implement SLOs.
Teachers' views of SLOs based on experience in education	Career teachers reported more positive views of SLOs and more knowledge about SLOs. Career teachers and early career teachers needed similar levels of support to implement SLOs.	Career teachers reported less positive views of SLOs, more knowledge about SLOs, and less support needed to implement SLOs.
Educators' views of classroom observations	Most educators agreed on the impact of classroom observations. In comparison with teachers, administrators had more positive views.	Most educators agreed on the impact of classroom observations. In comparison with teachers, administrators had more positive views.
Teachers' views of classroom observations based on degrees and experience	Differences in teachers' degrees or years of experience were not associated with their views of classroom observations.	Teachers with a bachelor's degree or below and the early career teachers reported more positive views of classroom observations.
Prediction of teachers' views of SLOs	Teachers' views of classroom observations, SLOs training, and experience in education predicted their views of SLOs.	Teachers' views of the classroom observations predicted their views of SLOs.

Table 5.2 *Highlights of the Findings from Study 3 and 4*

Study 3 Highlights	Study 4 Highlights
Teachers had mixed views about SLOs' impact on teaching and learning. Fewer than half of the teachers considered that SLOs have positive impact on teaching and learning. Elementary school teachers were more positive of SLOs.	Teachers' SLO scores could better differentiate teacher performance in comparison with their classroom observation scores.
SLOs benefits included tracking students' progress, teacher reflection, accountability, and using data.	There was a small positive association between teachers' SLO scores and their classroom observation scores.
SLOs challenges included time/timeline, paperwork, teacher autonomy (lack of supervision), assessment, and standards.	Teachers' SLO scores and their classroom observation scores demonstrated notable variations among districts, schools of different poverty levels, and different types of teachers.
Most teachers felt confident in using SLOs. Support should be provided to new teachers or teachers whenever there are some changes of the SLOs implementation policy or requirements.	
Half of the teachers considered SLOs as an additional reliable method in teacher evaluation.	

This might be due to the issues in SLOs implementation. Educators' views of SLOs in Study 2 were collected about one and half years after the full implementation of the evaluation system. Educators expressed various concerns about the implementation of SLOs through responding to the open-ended question in Study 2. Their concerns were related to students' test performance, issues in assessment methods, paperwork, limited time, inappropriate timeline for SLOs submission, teacher autonomy and lack of supervision in goal setting, subjectivity in assessment, teaching standards, and lack of feedback. In addition, the evaluation using SLOs might not apply to the special education teachers, arts teachers, ESL teachers, and media specialists. Similarly, findings from the

interviews with the teachers in Study 3 confirmed their overall negative views of the impact of SLOs. Fewer than half of the teachers interviewed shared that SLOs had positive impact on their teaching and student learning. They indicated the challenges of using SLOs in the evaluation of teacher effectiveness, which included time/timeline, paperwork, too much teacher autonomy and lack of supervision, fairness of assessment, and missing some standards. Therefore, these issues in the SLOs implementation might be the major reason for their negative views of the impact of SLOs after the implementation of the teacher evaluation system.

Despite educators' very different views of the impact of SLOs before and after the full implementation of the teacher evaluation system, they did report very similar levels of knowledge about SLOs in both occasions. Most educators (88% before the implementation of the system and 87% after the implementation of the system) reported having some to substantial knowledge about SLOs. Educators' knowledge about SLOs might be related to their familiarity with learning goals/objectives in teaching standards. Educators are generally exposed to learning goals in teacher preparation programs and their teaching practices at schools. As teachers interviewed in Study 3 expressed that they had the knowledge about SLOs, they knew the teaching standards and learning goals/objectives, and they felt very confident in using SLOs in teacher evaluation. Therefore, I conclude that most educators have some to substantial knowledge about SLOs either before or after the full implementation of the teacher evaluation system.

Regarding the support needed to implement SLOs, slightly more than half of the educators reported needing some or a lot of support before the full implementation of the evaluation system. After the full implementation of the system, about slightly fewer than

half of the educators reported needing some or a lot of support. Overall, not many educators reported needing some or a lot of support in using SLOs. The percentages of educators who reported needing some or a lot of support decreased after the full implementation of the evaluation system. Findings from Study 3 also revealed that teachers were confident about using SLOs and they did not need much support in implementing SLOs. It appears that after the full implementation of the system, many educators had a full experience of using SLOs and thus reported less support needed in implementing SLOs. Therefore, I believe that when educators have more experience of using SLOs, they probably need less support to implement SLOs.

In this study, teachers' and administrators' views of the impact of SLOs, knowledge about SLOs, and support needed to implement SLOs were compared before and after the full implementation of the teacher evaluation system. Before the full implementation of the system, administrators reported statistically significantly more positive views of the impact of SLOs than teachers. However, teachers and administrators did not report significantly different levels of knowledge about SLOs or support needed to implement SLOs. After the full implementation of the system, teachers and administrators did not report significantly different views of the impact of SLOs or support needed to implement SLOs. However, administrators had significantly more knowledge of SLOs. The differences of the views between teachers and administrators might be due to the different roles that they play in the process of teacher evaluation. Administrators are actively involved in decision-making about teacher recruitment, preparation, evaluation, and employment. Teachers are classroom instructors who work closely with students, and are often evaluated through classroom observations, SLOs, and

other professional conducts. Therefore, it is understandable that teachers and administrators hold different views of SLOs in the evaluation of teacher effectiveness.

In particular, teachers' views of the impact of SLOs, their knowledge about SLOs, and the support needed to implement SLOs were reported based on their academic degrees. Before the full implementation of the system, differences in teachers' academic degrees were not associated with differences in their views of SLOs, knowledge about SLOs, or support needed to implement SLOs. After the full implementation of the system, differences in teachers' academic degrees were not associated with the differences in their views of SLOs. However, teachers who had a master's degree or above reported statistically significantly more knowledge about SLOs and significantly less support needed to implement SLOs. Though teachers reported similar views of the impact of SLOs regardless of their highest academic degree both before and after the full implementation of the evaluation system, they did report significantly more knowledge and less support needed after the full implementation of the system. This could probably be attributable to the impact of their higher levels of coursework in the degree study and the implementation process.

Teachers' views of the impact of SLOs, their knowledge about SLOs, and support needed to implement SLOs were reported based on their years of experience in education. Before the full implementation of the evaluation system, career teachers reported statistically significantly more positive views of SLOs and significantly more knowledge about SLOs in comparison with early career teachers. There were no statistically significant differences of support needed between career teachers and early career teachers. After the full implementation of the system, career teachers reported statistically

significantly less positive views of SLOs. They reported statistically significantly more knowledge about SLOs and significantly less support needed to implement SLOs. The findings were consistent with those by Delvaux et al. (2013) who indicated that teachers with fewer than five years of teaching experience reported greater impact of the evaluation system on their professional development. It further suggests that teachers' experience in education is an important factor shaping their views of the impact of SLOs, knowledge about SLOs, and support needed to implement SLOs. Teachers who have longer years of experience in education might obtain more knowledge about teaching standards and learning goals, which further contributes to their higher levels of knowledge about SLOs.

5.2 EDUCATORS' VIEWS OF CLASSROOM OBSERVATIONS

Classroom observations have been used as a major mode in evaluating teacher effectiveness. The study also explored educators' views of the impact of classroom observations before and after the full implementation of the teacher evaluation system in South Carolina. Before the full implementation of the system, between 76.0% and 79.9% of the teachers and 100% of the administrators agreed or strongly agreed on the four aspects of the impact of classroom observations. After the full implementation of the system, between 55.4% and 69.3% of the teachers and between 93.8% and 100% of the administrators agreed or strongly agreed on the four aspects of the impact of classroom observations. There is a notable decrease in the percentages of teachers who agreed on the impact of classroom observations after the full implementation of the system. However, administrators' views of the impact of classroom observations are very similar before and after the implementation of the system.

Although educators had slightly less positive views of classroom observations after the full implementation of the system, the changes are small compared with the change of educators' views of SLOs. Overall, educators agreed on the impact of classroom observations and are positive about the impact of classroom observations both before and after the full implementation of the system. Administrators had significantly more positive views of classroom observations than teachers both before and after the full implementation of the system. The findings are consistent from those by Study 3. Some teachers interviewed in Study 3 indicated that classroom observations can reflect their teaching and are a reliable method in teacher evaluation. These finding echo those by Garrett and Steinberg (2015) who indicated that teachers' evaluation scores based on classroom observations are more straightforward and transparent and should be used for teacher performance evaluation.

In particular, teachers' views of the impact of classroom observations were reported based on their academic degrees and years of experience in education. Before the full implementation of the evaluation system, differences in teachers' degrees or years of experience were not associated with the differences in their views of the impact of classroom observations. After the full implementation of the evaluation system, teachers who had a bachelor's degree or below and the early career teachers reported statistically significantly more positive views of the impact of classroom observations. It suggests that the implementation process might be beneficial for those teachers who have a bachelor's degree or below or those early career teachers who have three or fewer years of experience in education.

In addition, the studies found that teachers' views of the impact of SLOs and their views of the impact of classroom observations are associated. Before the full implementation of the teacher evaluation system, teachers' views of classroom observations, SLOs training, and experience in education were significant predictors of their views of SLOs. After the full implementation of the system, teachers' views of the classroom observations were a significant predictor of their views of SLOs. It suggests that teachers who have positive views of classroom observations tend to have positive views of SLOs. Very interestingly, Study 4 revealed that teachers' SLO scores and their classroom observation scores have a significantly positive relationship. It suggests that teachers who have higher classroom observation scores tend to have higher SLO scores. It further indicates that there is some consistency of teachers' views of SLOs and classroom observations, and teachers' evaluation scores based on SLOs and classroom observations in the evaluation of teacher effectiveness.

5.3 SLOS IMPLEMENTATION

Based on the findings from the interview study (Study 3), teachers had mixed views about SLOs' impact on teaching and learning. Some teachers indicated that SLOs had positive impact on teaching and learning, sharing that SLOs could hold teachers accountable for student learning, make them be more reflective in their teaching, help track student growth, and use student data to inform teaching. These findings are consistent with the findings by previous studies. For example, Donaldson (2012) found that teachers had mixed views regarding the impact and implementation of SLOs. Makkonen et al. (2015) indicated that teachers in Utah considered the SLOs process was beneficial to students and teachers' professional development. Donaldson et al. (2014)

indicated that SLOs provided teachers with opportunities to use data and analyzing student data was valuable.

Despite some teachers' positive views of SLOs, others indicated that SLOs had little positive impact on teaching or learning, sharing that they already had the knowledge about teaching and would teach in the same way with or without SLOs. The findings are consistent with those by other researchers. For example, Makkonen et al. (2015) indicated that teachers in Utah did not consider the implementation of SLOs to have positive impact on instruction. However, some studies did find that teachers reported more and more positive views of SLOs along with longer implementation of SLOs. A series of studies of SLOs in Austin Independent School District (Texas) revealed that about 48% of teachers in 2008-2009 and 68% of teacher in 2009-2010 agreed or strongly agreed that using SLOs had improved their teaching (Courtemanche et al., 2014; Schmitt, et al., 2013). It appeared to show that more teachers might recognize the positive impact of using SLOs on their teaching along with the longer time of SLOs implementation.

According to Study 2 and 3, some teachers appeared to have concerns about the assessment methods used in measuring student growth and the standards assessed in SLOs. Teachers use teacher-designed assessment methods, and they have too much autonomy and there was a lack of supervision, and their decisions on students' growth were subjective. It appeared that the reliability and validity of the assessment was a key issue in SLOs. The findings were consistent with those by other studies. For example, Crouse et al. (2016) indicated that a limitation of SLOs was the validity, reliability, and accuracy of teachers' SLO scores due to the quality of the assessment designed by teachers and the quality of evaluators. Evidently, the assessment method is a key issue to

be addressed to ensure the fair assessment of student growth in the evaluation of teacher effectiveness.

Another major obstacle in the implementation of SLOs was related to time and paperwork. According to the teachers interviewed, the SLOs process was very time-consuming, the timeline was not appropriate, and it required a lot of paperwork. These obstacles resulted in frustration and stress for teachers. Similarly, Schmitt et al. (2008) reported that teachers considered the SLOs process as frustrating and time-consuming when SLOs were first piloted in 2007-2008. However, Courtemanche et al. (2014) and Schmitt et al., 2013) reported that by the academic year of 2013-2014, majority of teachers indicated that they often considered SLOs when planning and conducting their daily work, and the student achievement results of using SLOs were worth the extra work. It appeared that teachers' views of the SLOs changed in a more positive way along with the SLOs implementation. The obstacle of time, timeline, and paperwork seemed to be less a problem when teachers have more experience of using SLOs in teacher evaluation (Courtemanche et al., 2014; Schmitt et al., 2013).

In addition, this interview study found that elementary school teachers tended to have more positive views of using SLOs in teacher evaluation, and high school teachers tended to have less positive views of using SLOs in teacher evaluation. However, the results from Study 2 revealed that differences in the grade levels teachers taught were not associated with the differences in their views of the impact of SLOs or the impact of classroom observations. Considering that the interview study involved 18 teachers, and we should be cautious about making decisions that elementary school teachers are more positive. However, the differences in their views might be related to the curriculum in

different grade levels. For example, elementary schools provide ELA, math, science, social studies, PE, music, and visual arts classes, and many have state standardized assessment. Some elementary school teachers did report using standardized assessment to measure student growth in SLOs. However, high schools provide a variety of courses and most do not have standardized assessments and teachers design their own assessment to measure student growth. This posed an issue of validity of the assessment. As some teachers mentioned in the interview, teachers have too much autonomy, there was a lack of supervision, and the decisions of student growth were very subjective. This could probably explain why the elementary school teachers held more positive views of SLOs while high school teachers held less positive views of SLOs.

Furthermore, this interview study did not find an association of teachers' views of SLOs and such factors as subject taught, teaching experience, school poverty levels, and school location. This was not consistent with a study by Delvaux et al. (2013) who indicated that teachers with fewer than five years of teaching experience reported greater impact of the evaluation system on their professional development. Jiang et al. (2015) found that teachers' perceptions of the educator evaluation system were dependent on teacher characteristics, and teachers' perceptions about the school leadership and the professional community were positively associated with their perceptions of the evaluation system. The inconsistent findings of this study and the previous studies might be due to the number of teachers involved in the interview. Only eighteen teachers were interviewed, and it may not capture a complete picture of the impact of teacher and school characteristics on their views of the evaluation system.

Despite the concerns and obstacles of implementing SLOs, most of the teachers indicated that they felt confident in using SLOs. They shared that the SLOs process was simple, they had supportive school administration, and they had years of experience of using SLOs. At the same time, a few teachers reported that they needed some support due to changes of SLOs requirements. Some teachers believed that the early career teachers need more support in understanding and implementing SLOs. This is consistent with the findings from Study 2 that early career teachers reported needing statistically significantly more support to implement SLOs in comparison with career teachers. Similarly, Slotnik et al. (2014) surveyed educators (teachers and principals) in Maryland and found that about 50% of the educators reported needing support to have access to and analyze student data. Therefore, providing support to teachers, especially to early career teachers is important in implementing SLOs.

Finally, should SLOs be used as an additional reliable method in the evaluation of teacher effectiveness? Teachers reported mixed views. Half of the teachers considered SLOs as an additional reliable method in teacher evaluation, showing that SLOs could be used to track student growth, was a good indicator of teaching, and was reliable when used with other evaluation methods (e.g., classroom observation). However, half of the teachers showed that SLOs were not a reliable method in evaluating teacher effectiveness, and they had concerns about student performance in assessment, assessment issues, and teacher autonomy and lack of supervision. A few teachers considered classroom observation alone as an adequate method in the evaluation of teacher effectiveness. Therefore, policy makers should consider carefully the impact of SLOs, both benefits and obstacles of using SLOs, and assessment methods in decision

making regarding the evaluation of teacher effectiveness. In addition, the state department of education should consider teachers' views and make improvements of the teacher evaluation system.

5.4 TEACHERS' EVALUATION SCORES

This study revealed that about 99% of the teachers obtained a score of 2.5 or above, which suggests that teachers' classroom observation scores could not differentiate teacher performance. The findings of this study are consistent with the findings of previous studies. For example, 99.7% of the teachers in Chicago were evaluated as satisfactory to distinguished based on classroom observations (Rich, 2012), and 97.0% of teachers in New Jersey, 97.7% of the teachers in Florida, 95.0% of the teachers in New York, 98.0% of the teachers in Michigan, and 98.0% of the teachers in Tennessee were rated as effective or highly effective based on classroom observations (NCTQ, 2015, p. 12). This study of the evaluation scores of the teachers in South Carolina further confirmed that one major limitation of using classroom observation lies in its inability of differentiating between teacher performance.

The finding of this study revealed that 9.1% of the teachers obtained a minimum score of 1 point, 64.0% obtained a score between 2 and 4 points, and 26.9% obtained a maximum score of 5 points based on SLOs. There are notable percentages of teachers at each SLOs score point, which suggests that the SLO scores of teachers in South Carolina could better differentiate teacher performance. The finding of this study echoed those by Makkonen et al. (2015) who found that the end-of-year SLO scores of teachers in Arizona could differentiate between high- and low-performing teachers.

There was a small positive relationship between teachers' SLO scores and their classroom observation scores. The finding is consistent with those by Makkonen et al. (2015) who indicated that teachers' SLO scores had a statistically significant association with their ratings based on classroom observations. The findings suggest that teachers who obtain a higher SLO score tended to obtain a higher classroom observation score, with a weak relationship. It is very understandable because classroom observations are designed to capture teachers' instructional practices, while SLOs are used to measure students' academic growth. In addition, the small correlation coefficient might be due to the lack of alignment between goals of teaching practices captured by classroom observations and the student learning outcomes measured by standardized tests (Grossman et al., 2014).

This study discovered that school poverty level was significantly associated with teachers' evaluation scores based on both SLOs and classroom observations. Study 3 found that teachers from high poverty schools had significantly higher evaluation scores (both SLO scores and classroom observation scores) than those from low poverty schools. These results are inconsistent with previous studies (e.g., Steinberg & Sartain, 2015). A recent study by Dickenson et al. (2020) revealed that 96% of the teachers from high-poverty schools and 93% of the teachers from low-poverty schools scored Met based on the South Carolina teacher evaluation system Expanded ADEPT. Regarding teachers' SLO ratings, 86.26% of the teachers from high-poverty schools and 86.09% of the teachers from low-poverty schools scored Exemplary or Proficient. About 22% of the teachers from high-poverty schools and 32% of the teachers from low-poverty schools scored Exemplary on SLO. It suggests that higher percentage of teachers

from low-poverty schools were in the Exemplary category (highly effective), though very similar percentages of teachers from high and low-poverty schools were in the Exemplary or Proficient category. In generally, low poverty schools have more effective teachers, which means that the average evaluation scores of the teachers in the low poverty schools should be higher. The findings of the current study are not consistent with other studies. This is probably due to the sample of teachers used in the study. A convenience sampling method was used, and a small sample of 275 teachers from TAP schools were involved in the study. In addition, I believe that the variations of teachers' evaluation scores might be attributed to the evaluators, implementation strategies, school leadership, teacher workforce, and student population at the schools or districts. Further studies are needed to explore the relationship between school poverty and teachers' evaluation scores.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

This section includes the conclusions, limitations, and recommendations based on the studies on the SLOs in the evaluation of teacher effectiveness in South Carolina. The conclusions are drawn based on the four studies, delineating educators' views of SLOs before and after the full implementation of SLOs, comparing teachers' and administrators' views of SLOs, teachers' views of SLOs based on their degrees and experience in education, educators' views of classroom observations, the benefits and obstacles in implementing SLOs, SLOs assessment methods, and teachers' evaluation scores based on SLOs and classroom observations. The limitations are described in the aspects including sampling, self-reported survey responses, single year data about evaluation scores, quantifying qualitative data, and collecting educators' views across years. The recommendations are provided regarding the reform of the teacher evaluation system in South Carolina, policy making about teacher evaluation by considering teacher preparation, teacher recruitment, teacher retention, and using evaluation results to inform teaching, learning, and assessment. The recommendations are provided based on the findings of the four studies, which could inform teachers' professional development, improve teacher evaluation, and enhance teaching and student learning.

6.1 CONCLUSIONS

Through a a mixed-methods sequential explanatory design (Creswell & Creswell, 2018), this research investigated using SLOs as a measure of student growth in the evaluation of teacher effectiveness. The four studies that employed surveys, interviews, and teachers' evaluation scores from a total of 1,020 participants revealed significant findings. Several overarching conclusions could be drawn based on the findings from the four studies.

First, most educators disagreed on the impact of SLOs after the full implementation of the teacher evaluation system though they reported similar levels of knowledge about SLOs. Second, most educators agreed on the impact of classroom observations both before and after the full implementation of the evaluation system. Third, in comparison with teachers, administrators reported more positive views of both SLOs and classroom observations. Fourth, after the full implementation of the evaluation system, early career teachers reported more positive views of both SLOs and classroom observations. In addition, there were various issues in the implementation of SLOs in teacher evaluation. Finally, teachers who had a positive view of classroom observations tended to have more positive views of SLOs, and teachers who had a higher classroom observation score tended to have a higher SLO score.

The four studies revealed both the positive and negative social consequences of using SLOs in the evaluation of teacher effectiveness. Regarding the consequential validity of educational tests, Messick (1989) suggested that tests should be labeled correctly and thoughtfully, and the potential and actual social consequences of applied testing should be identified. Shepard (1993) further argued that social consequences of

educational tests should be investigated. In the context of teacher evaluation, the social consequences should be considered as well. The findings of the current research showed that using SLOs in the evaluation of teacher effectiveness have both positive and negative consequences. SLOs are appropriate to measure teachers' contribution to student learning, especially considering teachers' accountability. SLOs are complementary to the classroom observations because the evaluation of teacher effectiveness should involve multiple methods and the teacher ratings based on classroom observations could not differentiate teacher performance. At the same time, SLOs could possibly cause teachers' frustration and stress and teacher attrition due to the issues in implementation. Therefore, the school administrators and policy makers should acknowledge the social consequences of using SLOs in teacher evaluation and employ implementation strategies that could avoid or reduce the negative consequences.

6.2 LIMITATIONS

Although the findings of this study can inform teacher evaluation, teachers' professional development, and decision making about teacher recruitment, preparation, and retention, there are several limitations. First, one major limitation of this study is related to sampling. Study 1 used a convenience sample, and educators from the schools that were involved in the TAP and the South Carolina Department of Education (SCDE) Partnership Program participated in the survey study. These schools were high-need schools, and they might not fully represent all schools in South Carolina. In Study 2, I used a stratified random sampling and eight school districts were sampled to participate in the survey study. Three out of the eight school districts approved the study. I compared the three school districts and all school districts in South Carolina on multiple indicators

including school poverty level, school location, and school enrollment. The three school districts are not completely representative of all school districts in South Carolina, but I believe data collected from the three school districts are valuable information about teacher evaluation.

Second, Study 1 and 2 used surveys and educators' self-reported data were analyzed. Educators reported their perceptions of the impact of SLOs and classroom observations, their knowledge about SLOs, and the support needed to implement SLOs. For example, I asked educators to report their levels of knowledge about SLOs by selecting from no knowledge, limited knowledge, some knowledge, and substantial knowledge. Educators' reported knowledge in a survey might be different from their knowledge measured by a test or scale. In addition, educators might use different criteria to make judgement about their knowledge level. This is a common limitation for self-reported data.

Third, teachers' evaluation scores from SLOs and classroom observations were based on one-year evaluation results. I found a relationship between teachers' SLO scores and classroom observation scores, and SLO scores can better differentiate teacher performance. However, I did not track the possible changes across years of evaluation, when teachers might teach different classes, different students, and different subjects, and are evaluated by different evaluators. Morgan et al. (2014) investigated teachers' observational ratings across four years and found that teachers' observational ratings were not stable across time. Similarly, Rowan et al. (2013) indicated that teachers' observational scores demonstrated considerable variability based on evaluators and lessons taught. However, I used only a one-year one-time evaluation scores of teachers,

and tracking teachers' evaluation scores for several years might be able to help portray a full picture about teacher effectiveness by multiple evaluation methods.

Fourth, the interview study (Study 3) revealed that elementary school teachers (five out of six) reported positive views and considered SLOs as a reliable method in teacher evaluation. Middle school teachers reported divided views and half of them reported positive views and considered SLOs as a reliable method. High school teachers (five out of six) tended to have negative views and did not consider SLOs as a reliable method in teacher evaluation. I attempted to employ a quantitative method chi-square test to examine whether school level (elementary, middle, and high) and teachers' decisions about whether SLOs were a reliable method (Yes vs No) have any associations. However, due to the small sample size of 18 teachers, the requirement of at least five counts in each cell was not met for conducting a Chi-square test. A large sample size in a qualitative study is preferred to help quantify the qualitative study and discover more important information.

Finally, Study 2 and 3 were conducted about one year and half after the full implementation of the teacher evaluation system in South Carolina. Both studies found that educators did not hold a very positive view of using SLOs in the evaluation of teacher effectiveness. Teachers reported various concerns about using SLOs in teacher evaluation. I believe educators' views are dynamic, and their experience, school leadership, students, subjects taught, and policy might shape their views about SLOs. It is understandable that some teachers reported frustration and stress of using SLOs, especially considering that they just had more than one year of experience after the full implementation of the system. As Courtemanche et al. (2014) and Schmitt et al. (2013)

conducted a series of studies of SLOs in Austin Independent School District (Texas), and they found that more teachers reported positive impact of SLOs with longer time of implementation. I expect to conduct more surveys to understand educators' views of using SLOs in the evaluation of teacher effectiveness in the future studies.

6.3 RECOMMENDATIONS

Based on the findings of the four studies, I provide recommendations for school districts, schools, administrators, policymakers, teachers, and educational researchers. First, we should recognize that teacher evaluation is one of many important elements in education. All elements in education are interrelated. While making decision about teacher evaluation, we should be aware of other elements in the system including teacher preparation, teacher recruitment, teacher induction, teacher retention, teacher employment, instruction, assessment, etc. We should predict and visualize the consequences of policymaking about teacher evaluation, especially considering school poverty and location. As Garcia and Weiss (2019) described, teacher shortage in the United States was real, large, and growing. Sutcher et al. (2019) indicated that the most important driving factor of teacher shortages was high teacher attrition, and the highest overall turnover rates were in the South (Carver-Thomas & Darling-Hammond, 2019). Teacher attrition was also described by one teacher whom I interviewed. As a department chair, she had witnessed several of new teachers leaving their profession due to the stress from teacher evaluation. Therefore, policy making regarding teacher evaluation should take such elements as teacher shortage and teacher attrition into consideration. We should also acknowledge that teacher shortage is especially serious in certain subject areas including mathematics, science, special education, and English as a second language, and

in high poverty schools and schools with a large percentage of minority enrollments (Garcia & Weiss, 2019). Similarly, some teachers whom I interviewed indicated that the SLOs evaluation rules and principles cannot be applied to the special education teachers, music and visual arts teachers, and English as a second language teachers. Therefore, policy makers should take these into consideration while making decisions about teacher evaluation.

Second, we should understand the relationship of teaching, learning, and assessment, especially when we make decisions about teacher evaluation. Green and Johnson (2010) indicated that classroom assessment is an essential facet of education that is used to evaluate student learning, and there are two major purposes of assessment: assessment *of* learning and assessment *for* learning. Assessment is part of the teaching process and is an effective way for teachers to gather useful information about student learning. Therefore, a teacher's professional responsibility in assessment is to use high quality assessment information to make decisions about students' learning (Brookhart & Nitko, 2014). In teacher evaluation, SLOs is used to measure student academic growth, which is considered as a measure of a teacher's contribution to student learning. However, there are many factors that contribute to student learning. As many teachers who were involved in Study 2 and 3 indicated that they were concerned about linking student test performance to their effectiveness. Although teachers believed that they should be accountable for student learning, they did not consider it fair to be solely responsible for student learning outcomes. Therefore, we should be cautious while making decisions about using student learning outcomes in determining teacher

effectiveness, and teachers' SLO scores should not be weighed too much especially in making high-stake decisions for teacher employment and teacher compensation.

Third, the implementation of SLOs should be reformed and improved. Educators did not report positive views of the impact of SLOs after the full implementation of the system. Educators in Study 2 and 3 explicitly expressed various concerns related to SLOs implementations. These issues are related to inappropriate timeline for SLOs submission, teachers' autonomy and lack of supervision, and subjectivity in setting goals and assessment, validity and reliability of assessment methods in measuring student growth, teaching to the test for achieving the goals, missing some standards, lack of feedback, and other issues. Some teachers considered SLOs as "another hoop to jump through," or "just another chore to complete." However, as Schneider and Johnson (2019) indicated, "The SLO process is not a template that teachers complete at the beginning of the year and return to at the end of the year. It is a formative assessment process of understanding where students are in their learning and where they need to go next in regard to the SLO learning goal" (p. 142). Similarly, some teachers indicated that teacher-designed assessment are too subjective and could not fairly assess student learning. Goe and Holdheide (2011) recommended that student growth measures should be designed and monitored at the state level to ensure standardization. Therefore, I recommend that SLOs process should be incorporated in the process of instruction, and SLOs results should be used to individualize learning for students and inform instruction for teachers.

In addition, teachers' professional characteristics should be taken into consideration in making decisions about teacher evaluation. Study 2 and 3 revealed that early career teachers reported more positive views of SLOs and classroom observations,

and they needed more support in using SLOs. Therefore, providing professional development and other types of support is especially important for early career teachers. Teacher evaluation is a very important element in teacher education, as well as teaching and learning. It is used as a method to summarize teacher performance and determine teacher effectiveness. Most importantly, it should be used as a method to evaluate teacher performance, inform teachers' professional development, enhance teacher quality, and improve student learning. As Donaldson (2016) indicated that "the key to getting the most out of teacher evaluation is figuring out how to implement it in a way that challenges, supports, and motivates teachers." (p. 76). Considering the current issue of teacher shortage in South Carolina and the rest of the country, it is important to use evaluation as a mode to prepare teachers, help them grow, and retain effective teachers. At the same time, schools should develop supportive administrative leadership and shared leadership (Kukla-Acevedo, 2009; Podolsky et al., 2019), build welcoming and supportive teaching and learning community, empower teachers, build strong teacher workforce in South Carolina, and ultimately improve student learning.

Therefore, I recommend that school districts, schools, and administrators should acknowledge the current situation of teacher workforce in South Carolina, use the evaluation as a strategy for teacher development, build a strong professional community, and develop mentoring programs for the teachers in need. Most importantly, schools should include some effective teachers in the evaluation team, guide and supervise the evaluation process, ensure fair assessment of student growth, and provide valuable feedback to the teachers. In addition, the SCDE should reform the teacher evaluation system by taking teachers' workload, busy schedule, and stress into consideration.

Instead of evaluating all teachers every year, the SCDE might consider evaluating teachers selectively and the evaluation could be conducted every three years.

The four studies focused on the state of South Carolina that has its unique characteristics as a southern state. About 40% of the students in South Carolina are educated in rural schools, which is much higher than the approximately 24% of the students in rural schools in the nation (Irvin et al., 2020). About 33% of K-12 education students are Black in South Carolina (South Carolina Department of Education, 2020) in comparison with about 15% of Black students nationwide (National Center for Education Statistics [NCES], 2020). Therefore, the findings of the four studies might not be generalized to other states in the nation, especially those in the north. I recommend more studies should be conducted to better understand the perspectives of educators on the teacher evaluation system and the characteristics of teachers' evaluation scores nationwide. In addition, I recommend that longitudinal data about the perspectives of educators on the teacher evaluation system and the characteristics of teachers' evaluation scores should be collected yearly to track the changes over time. Data from multiple years could provide a better picture of the evaluation of teacher effectiveness, which will ultimately better inform policy making regarding teacher evaluation. Finally, I recommend that more studies should focus on the evaluation of principal effectiveness considering the impact of school leadership on teacher preparation, retention, and effectiveness. More studies should be conducted to investigate the impact of school poverty, school location, and school climate on teacher effectiveness.

REFERENCES

Aaronson, D., Barrow, L., & Sanders, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25* (1), 95-135.

AERA. (2015). *AERA statement on use of Value-Added Models (VAM) for the evaluation of educators and educators preparation programs*. Educational Researcher. 1-5.

Amrein-Beardsley, A., & Collins, C. (2012). The SAS Education Value-Added Assessment System (SAS®EVAAS®) in the Houston Independent School District (HISD): Intended and Unintended Consequences. *Education Policy Analysis Archives, 20*(12). http://epaa.asu.edu/ojs/article/view/1096

Balch, R. (2012). *The validation of a student survey on teacher practice*. Vanderbilt University. http://mystudentsurvey.com/wp-content/uploads/2012/06/Balch-Student-Surveys-2012.pdf

Balch, R., & Springer, M. G. (2015). Performance pay, test scores, and student learning objectives. *Economics of Education Review, 44*(2015), 114-125.

Basileo, L. D., & Toth, M. (2019). A state level analysis of the Marzano teacher evaluation model: Predicting teacher value-added measure with observation scores. *Practical Assessment, Research & Evaluation, 24*(6).

Bell, C. A., Gitomer, D. H., McCaffrey, D. F., Hamre, B. K., Pianta, R. C., & Qi, Y. (2012). An argument approach to observation protocol validity. *Educational Assessment, 17*(2/3), 62-87.

Brill, S. (2009). The rubber room. *The New Yorker,* August 31, 2009.

Brookhart, S. M., & Nitko, A. J. (2014). *Educational assessment of students* (7th ed.). Upper Saddle River, NJ: Pearson.

Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Whitrock (Ed.), *Handbook of research on teaching* (pp. 328-375). New York, NY: Macmillan.

Callahan, K. & Sadeghi, L. (2015). Teacher perceptions of the value of teacher evaluations: New Jersey's ACHIEVE NJ. *NCPEA International Journal of Educational Leadership Preparation*, *10* (21). ISSN: 2155-9635, 2014 National Council of Professors of Educational Administration.

Campbell, R. J., Kyriakides, L., Mujis, R. D., & Robinson, W. (2004). Differential teacher 218ffect- tiveness: Towards a model for research and teacher appraisal. *Oxford Review of Education, 29,* 347-362.

Carver-Thomas, D., & Darling-Hammond, L. (2019). The trouble with teacher turnover: How teacher attrition affects students and schools. *Education Policy Analysis Archives, 27*(36). DOI: http://dx.doi.org/10.14507/epaa.27.3699

Center on Great Teachers and Leaders. (2015). *Databases on state teacher and principal evaluation policies* (Website). Washington, DC: Author. http://resource.tqsource.org/stateevaldb/STEPEightKeys.aspx

Cherasaro, T. L., Brodersen, R. M., Reale, M. L., & Yanoski, D. C. (2016). *Teachers' response to feedback from evaluators: What feedback characteristics matter?* (REL 2017-190). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Center. http://ies.ed.gov/ncee/edlabs.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014a). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review 104*(9):2593–2632. doi:10.1257/aer.104.9.2593

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014b). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review 104*(9):2633–2679. doi:10.1257/aer.104.9.2633

Coburn, C. E. (2005). Shaping teacher sensemaking: School leaders and the enactment of reading policy. *Educational Policy*, *19*(3), 476-509.

Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic.

Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social-emotional learning: Teacher stress, job satisfaction and teaching efficacy. *Journal of Educational Psychology*, *104*(4): 1189-1204.

Collins, C. (2014). Houston, we have a problem: Teachers find no value in the SAS education value-added assessment system (EVAAS). *Education Policy Analysis Archives*, *22*(98).

Community Training and Assistance Center [CTAC]. (2008*). Tying earning to learning: The link between teacher compensation and student learning objectives*. Boston, MA: Author.

Community Training and Assistance Center [CTAC]. (2013*). It's more than money: Teacher incentive fund leadership for educators' advanced performance, Charlotte-Mecklenburg schools*. Boston, MA: Author.

Corcoran, S. (2010). Can teachers be evaluated by their students' test scores? Should they be? (Report for the Educational Policy for Action Series). Providence, RI: Annenberg Institute for School Reform at Brown University.

Courtemanche, M., Orr, A., & Schmitt, L. (2014). *AISD REACH program update: 2013–2014 participant feedback* (DRE Report No. 13.39). Austin, TX: Austin Independent School District. http://www.austinisd.org/sites/default/files/dre-reports/DRE_13.39_AISD_Reach_Program_Update_2013_2014_Participant_Feedback.pdf

Creswell, J. W., & Creswell, J.D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (5th ed.)*. Thousand Oaks, CA: Sage.

Crouse, K., Gitomer, D. H., & Joyce, J. (2016). An analysis of the meaning and use of student learning objectives. In A. Amrein-Beardsley & K. Kappler Hewitt (Eds.), *Student growth measures: Where policy meets practice*. New York, NY: Palgrave Macmillan.

Danielson, C. (1996). *Enhancing Professional Practice:* A *framework for Teaching.* Alexandria, VA: Association for Supervision and Curriculum Development.

Danielson, C. (2001). New trends in teacher evaluations. *Educational Leadership, 58*(5), 12-15.

Danielson, C. (2013). *The framework for teaching evaluation instrument*. Princeton, NJ: The Danielson Group.

Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional learning. Princeton*, NJ: Educational Testing Service.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis, 8*(1), 15-17.

Darling-Hammond, L. (2012). *Creating a comprehensive system for evaluating and supporting effective teaching*. Stanford, CA. Stanford Center for Opportunity Policy in Education.

Darling-Hammond, L. (2013). *Getting teacher evaluations right: What really matters for effectiveness and improvement*. New York, NY: Teachers College Press.

Darling-Hammond, L., & Young, P. (2002). Defining "highly qualified teachers": What does "scientifically based research" actually tell us? *Educational Researcher, 31*(9), 13-25.

Datnow, A., & Castellano, M. (2000). Teachers' responses to success for all: how beliefs, experiences, and adaptations shape implementation. *American Educational Research Journal, 37*, 775-799.

Delvaux, E., Vanhoof, J., Tuytens, M., Vekeman, E., Devos, G., & Petegem, P. V. (2013). How may teacher evaluation have an impact on professional development? A multilevel analysis. *Teaching and Teacher Education, 36* (2013), 1-11.

de Winter, J. F.C. & Dodou, D. (2010). Five-Point Likert Items: t test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research, and Evaluation*, 15(11).

Dickenson, T., Fan, X., Pan, F., Kunz, G., & Hodges, T. (2020). *Profile of the South Carolina teacher workforce for 2018-2019*. Columbia, SC: South Carolina Teacher Education Advancement Consortium through Higher Education Research (SC-TEACHER).

Donaldson, M. L. (2009). *So long, Lake Wobegon*? Center for American Progress, Washington, D.C.

Donaldson, M. L. (2012). *Teachers' perspectives on evaluation reform.* Washington, DC: Center for American Progress. http://www.americanprogress.org/wp-content/uploads/2012/12/TeacherPerspectives.pdf

Donaldson, M. L. (2016). Teacher evaluation reform: focus, feedback, and fear. *Educational Leadership*.

Donaldson, M. L., Cobb, C., LeChasseur, K., Gabriel, R., Gonzales, R., Woulfin, S., & Makuch, A. (2014). *An evaluation of the pilot implementation of Connecticut's system for educator evaluation and development.* http://aftct.org/sites/aftct.org/files/neag_seed_report_1_1_14.pdf

English, D., Burniske, J., Meibaum, D., & Lachlan-Haché, L. (2016). *Using student surveys as a measure of teaching effectiveness.* American Institutes for Research. https://education.ohio.gov/getattachment/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System/Alternative-Components/Student-Surveys-in-Teacher-Evaluation-FINAL-for-Ohio-060816.pdf.aspx

Ferguson, R. F. (2010). *Student perceptions of teaching effectiveness* (Discussion Brief). Cambridge, MA: National Center for Teacher Effectiveness and the Achievement Gap Initiative.

Finster, M., & Milanowski, A. (2018). Teacher perception of a new performance evaluation system and their influence on practice: A within- and between-school level analysis. *Education Policy Analysis Archives, 26*(41). Doi: 10.14507/epaa.26.3500

Follman, J. (1992). Secondary school students' ratings of teacher effectiveness. *The High School Journal, 75*(3), 168-178.

Follman, J. (1995). Elementary public school pupil rating of teacher effectiveness. *Child Study Journal, 25*(1), 57-78.

Ford, T. G., Urick, A., & Wilson, A. S. P. (2018). Exploring the effect of supportive teacher evaluation experience on U.S. teachers' job satisfaction. *Education Policy Analysis Archives, 26*(59).

Fullan, M. (2011). *Leading in a culture of change*. San Francisco, CA: Jossey-Bass.

Gallagher, H. A. (2004). Vaughn Elementary's innovative teacher evaluation system: Are teacher evaluation scores related to growth in student achievement? *Peabody Journal of Education, 79*, 79-107.

Garcia, E., & Weiss, E. (2019). The teacher shortage is real, large and growing, and worse than we thought. Economic Policy Institute. https://www.epi.org/publication/the-teacher-shortage-is-real-large-and-growing-and-worse-than-we-thought-the-first-report-in-the-perfect-storm-in-the-teacher-labor-market-series/

Garrett, R., & Steinberg, M. P. (2015). Examining teacher effectiveness using classroom observation scores: Evidence from the randomization of teachers to students. *Educational Evaluation and Policy Analysis, 37*(2), 224-242.

Gitlin, A., & Margonis, F. (1995) The political aspect of Reform: Teacher resistance as good sense. *American Journal of Education, 103*(3), 377-405.

Goe, L., Bell. C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. http://www.gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf

Goe, L., & Holdheide, L. (2011). *Measuring teachers' contributions to student learning growth for nontested grades and subjects*. Washington, DC: National Comprehensive Center for Teacher Quality.

Goldhaber, D. D., & Brewer, D. J. (1997). Evaluating the effect of teacher degree level on educational performance. In W. J. Fowler (Ed.), *Developments in school finance, 1996* (pp. 197-210). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Green, S. K., & Johnson, R. L. (2010). Assessment is essential. Burr Ridge, IL: McGraw-Hill Higher Education.

Grossman, P., Cohen, J., Ronfeldt, M., & Brown, L. (2014). The test matters: The relationship between classroom observation scores and teacher value added on multiple types of assessment. Educational Researcher, 43(6), 293-303.

Grossman, P., Loeb, S., Cohen, J., & Wyckoff, J. (2013). Measure for measure: The relationship between measures of instructional practice in middle school English language arts and teachers' value-added scores. *American Journal of Education, 119*(3), 445-470.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy, 100*, 84-117.

Hanushek, E. A. (2009). Teacher deselection. In D. Goldhaber & J. Hannaway (Eds.), *Creating a new teaching profession* (pp. 165-180). Washington, DC: Urban Institute Press.

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review, 30*(3), 466-479.

Harris, D. N. (2011). *Value-added measures in education: What every educator needs to know*. Cambridge, MA: Harvard Education Press.

Harris, D. N., Ingle, w. K., & Rutledge, S. A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal, 51*(1), 73-112. Doi: 10.3102/0002831213517130

Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

Heck, R. (2009). Teacher effectiveness and student achievement, *Journal of Educational Administration, 47*(2), 227-249
Doi: https://doi.org/10.1108/09578230910941066

Heneman, H. G., Milanowski, A., Kimball, S. M., & Odden, A. (2006). *Standards-based teacher evaluation as a foundation for knowledge and skill-based pay* (CPRE Research Report RB-45). Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education. http://www.cpre.org/images/stories/cpre_pdfs/RB45.pdf

Irvin, M. J., Kunz, G. M., Starrett, A. D., & Harrist, J. N. (2020). *Recruitment and retention of teachers in rural South Carolina*. Columbia, SC: South Carolina Teacher Education Advancement Consortium through Higher Education Research (SC-TEACHER). https://sc-teacher.org/ruralsc/

JCSEE. (2009). *The personnel evaluation standards: How to assess systems for evaluating educators* (2nd ed.). Thousand Oaks, CA: Corwin.

Jiang, J.Y., Sporte, S.E., and Luppescu, S. (2015). Teacher perspectives on evaluation reform: Chicago's REACH students, *Educational Researcher, 44*(2), 105-116.

Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). *Have we identified effective teachers? Validating measures of effective teaching using random assignment.* Seattle, WA: Bill & Melinda Gates Foundation.

Kearney, W. S., & Garfiend, T. (2019). Student readiness to learn and teacher effectiveness: The key factors in middle grades mathematics achievement. *Research in Middle Level Education, 42*(5).

Koedel, C., Li, J., Springer, M. G., & Tan, L. (2017). The impact of performance ratings on job satisfaction for public school teachers. American Educational Research Journal, 54(2), 241-278. Doi: 10.3102/0002831216687531

Kraft, M. A., & Gilmour, A. F. (2017). Revisiting the widget effect: Teacher evaluation reforms and the distribution of teacher effectiveness. *Educational Researcher, 46*(50), 234-249. Doi: 10.3102/0013189X17718797

Kukla-Acevedo, S. (2009). Leavers, movers, and stayers: The role of workplace conditions in teacher mobility decisions. *The Journal of Educational Research, 102*(6), 443-452. DOI: https://doi.org/10.3200/JOER.102.6.443-452

Lachlan-Haché, L. (2015). *The art and science of Student Learning Objectives A research synthesis.* Washington, DC: American Institutes for Research.

Lachlan-Haché, L., Cushing, L. & Bivona, L. (2012). *Student learning objectives as measures of educator effectiveness: The basics*. American Institutes for Research. http://educatortalent.org/inc/docs/SLOs_Measures_of_Educator_Effectiveness.pdf

Lacireno-Paquet, N., Morgan, C., & Mello, D. (2014). *How states use student learning objectives in teacher evaluation systems: A review of state websites* (REL 2014–013). U.S. Department of Education, Institute of Education Sciences. http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=REL2014013

Lamb, L. M., Schmitt, L. N. T., Gross, R., & Cornetto, K. M. (2013). *Austin Independent School District (AISD) pilot teacher appraisal system update: 2012−2013 focus group and survey summary* (DRE Publication No. 12.70). Austin, TX: Austin Independent School District. http://www.austinisd.org/sites/default/files/dre-reports/12.70_Austin_Independent_School_District_AISD_Pilot_Teacher_Appraisal_System_Update.pdf

Liu, E., & Johnson, S. M. (2006). New teachers' experience of hiring: Late, rushed, and information-poor. *Educational Administration Quarterly, 42*, 324-360.

Liu, Y., Visone, J., Mongillo, M. B., & Lisi, P. (2019). What matters to Teachers if evaluation is meant to help them improve? *Studies in Educational Evaluation, 61*(2019), 41-54.

Lockwood, J. R., McCaffrey, D. F., Hamilton, L.S., Stetcher, B., Le, V. N., & Martinez, J. F.(2007). The sensitivity of value-added teacher effect estimates to different mathematics achievement measures. *Journal of Educational Measurement, 44* (1), 47-67.

Makkonen, R., Tejwani, J., & Rodrigues Jr., F. (2015). *A descriptive study of the pilot implementation of student learning objectives in Arizona and Utah* (REL 2016-102). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. http://ies.ed.gov/ncee/edlabs.

Marion, S., DePascale, C., Domaleski, C., Gong, B., & Diaz-Bello, E. (2012). Considering for analyzing educators' contributions to student learning in non-tested subjects and grades with a focus on student learning objectives. Concord, NH: Center for Assessment. http://www.nciea.org/publication_PDFs/Measurement%20Considerations%20for%20NTSG_052212.pdf

Marshall, J. C., Smart, J., & Alston, D. M. (2016). Development and validation of Teacher Intentionality of Practice Scale (TIPS): A measure to evaluate and scaffold teacher effectiveness. *Teaching and Teacher Education. 59*(2016), 159-168.

Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership, 70*(3), 14-19.

Marzano, R. J., & Toth, M. (2013). *Teacher evaluation that makes a difference.* Alexandria, VA: ASCD.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-103). New York: Macmillan.

MET Project. (2012). *Asking students about teaching: Student perception surveys and their implementation* (Policy & Practice Brief). Seattle, WA: Bill and Melinda Gates Foundation. http://k12education.gatesfoundation.org/wp-content/uploads/2015/12/Asking_Students_Summary_Doc.pdf

Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education, 79*(4), 33-53.

Moran, R. M. R. (2017) The impact of a high stakes teacher evaluation system: Educator perspectives on accountability, *Educational Studies, 53*(2), 178-193, DOI: 10.1080/00131946.2017.1283319

Morgan, G. B., Hodge, K. J., Trepinski, T. M., & Anderson, L. W. (2014). The stability of teacher performance and effectiveness: Implications for policies concerning teacher evaluation. Education Policy Analysis Archives, 22(95). Doi: http://dx.doi.org/10.14507/epaa.v22n95.2014

Morganstein, D., & Wasserstein, R. (2014). ASA statement on Value-Added Models, *Statistics and Public Policy, 1*:1, 108-110.

Muncey, D., & McQuillan, P. (1996). *Reform and resistance in schools and classrooms*. New Haven, CT: Yale University Press.

Muñoz, M. A., Prather, J. R., & Stronge, J. H. (2011). Exploring teacher effectiveness using hierarchical linear models: Student- and classroom-level predictors and cross-year stability in elementary school reading. *Planning and Changing, 42*(3/4), 241-273.

National Center for Education Statistics (2020). Racial/Ethnic enrollment in public schools. https://nces.ed.gov/programs/coe/indicator_cge.asp

National Council on Teacher Quality. (2015). State of the States Evaluating Teaching, Leading, and Learning. http://www.nctq.org/dmsView/StateofStates2015

National Council on Teacher Quality. (2017a). NCTQ State Teacher Policy Yearbook 2017: Executive summary. http://www.nctq.org/dmsView/NCTQ_2017-Yearbook_Executive_Summary

National Education Association. (2010). *Teacher assessment and evaluation*. http://www.nea.org/assets/docs/HE/TeachrAssmntWhtPaperTransform10_2.pdf

New Teacher Project. (2012). Summer report: Creating a culture of excellence in Indiana schools. Indianapolis, IN: Indiana Department of Education. http://www.doe.in.gov/sites/default/files/evaluations/tntp-summer-report-june-2012.pdf

Newton, X. A., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Educational Policy Analysis Archives, 18*(23).

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Odden, A., Borman, G., & Fermanich, M. (2004). Assessing teacher, classroom, and school effects, including fiscal effects. *Peabody Journal of Education, 79*(4), 4-32.

Papay, J. P. (2011). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *American Educational Research Journal, 48*, 163-193.

Paufler, N. A. (2018). Declining morale, diminishing autonomy, and decreasing value: Principal reflections on a high-stakes teacher evaluation system. *International Journal of Education Policy & Leadership, 13*(8).

Peterson, K. D., & Peterson, C. (2006*). Effective teacher evaluation. Guide for principals.* California: Corwin Press.

Peterson, K., Wahlquist, C., & Bone, K. (2000). Student surveys for school teacher evaluation. *Journal of Personnel Evaluation in Education, 14*(2), 135-153.

Podolsky, A., Kini, T., Darling-Hammond, L., & Bishop, J. (2019). Strategies for attracting and retaining educators: What does the evidence say? *Education Policy Analysis Archives, 27*(38). DOI: https://doi.org/10.14507/epaa.27.3722

Popham, W. J. (2013). *Evaluating America's teachers: Mission possible?* Thousand Oaks, CA: Corwin.

Rich, M. (2012). National school debate is on in Chicago. *New York Times*, September 11, 2012.

Rivkin, S. G., Hanushek, E. A. & Kain, J. F.(2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement. *American Economic Review, 94*(2):247-252.

Rockoff, J. E., & Speroni, C. (2010). Subjective and Objective Evaluations of Teacher Effectiveness. *American Economic Review, 100*(2), 261-266.

Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York, NY: Free Press.

Rowan, B., Chiang, F. S., & Miller, R. J. (1996). Using research on employee performance to study the effects of teachers on students' achievement. *Sociology of Education, 70,* 256-284.

Rowan, B., Jacob, R., & Correnti, R. (2009). Using instructional logs to identify quality in educational settings. *New Directions for Youth Development,* 13-31.

Rowan, B., Schilling, S. G., Spain, A., Bhandari, P., Berger, D., & Graves, J. (2013). Promoting high quality teacher evaluations in Michigan: Lessons from a pilot of educator effectiveness tools. Institute for Social Research, University of Michigan.

Ruffini, S., Makkonen, R., Tejwani, J., & Diaz, M. (2014). *Principal and teacher perceptions of implementation of multiple-measure teacher evaluation systems in Arizona* (REL 2015-062). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West.

Sanders, W. L. (2000). Value-Added Assessment from Student Achievement Data: Opportunities and Hurdles. *Journal of Personnel Evaluation in Education, 14*(4), 329-339.

Schmitt, L. N. T., Lamb, L. M., Cornetto, K. M., & Courtemanche, M. (2013). *AISD REACH program update, 2012−2013: Student learning objectives* (DRE Publication No. 12.83). Austin, TX: Austin Independent School District. https://www.austinisd.org/sites/default/files/dre-reports/DRE_12.83_AISD_REACH_Program_Update_2012_2013_Student_Learning_Objectives.pdf

Schmidt, M., & Datnow, A. (2005). Teachers' sense-making about comprehensive school reform: The influence of emotions. *Teaching and Teacher Education, 21*(8), 949-965.

Schmitt, L., Malerba, C., Cornetto, K., & Bush-Richards, A. (2008). *Strategic compensation interim report 2: Teacher focus group summary, spring 2008* (DPE Publication No. 07.32). Austin, TX: Austin Independent School District. http://www.austinisd.org/sites/default/files/dre-reports/07.32_Strategic_Compensation_Interim_Report_2_Teacher_Focus_Group_Summary_Spring_2008.pdf

Schneider, M. C., & Johnson, R. L. (2019). *Using formative assessment to support student learning objectives.* New York and London: Routledge.

Shepard, L.A. (1993). Evaluating Test Validity. In L. Darling-Hammon (Ed.), *Review of Research in Education*, *19*. Washington, DC: AERA.

Skaalvik, E. M., & Skaalvik, S. (2011). Teacher job satisfaction and motivation to leave the teaching profession: Relations with school context, feeling of belonging, and emotional exhaustion. *Teaching and Teacher Education, 27*(4):1029-1038.

Slotnik, W. J., Bugler, D., & Liang, G. (2014). *Real progress in Maryland: Student learning objectives and teacher and principal evaluation.* Washington, DC: Mid-Atlantic Comprehensive Center. https://www.wested.org/resources/real-progress-in-maryland-59136/

Slotnik, W. J., Smith, M., Glass, R., & Helms, B. J., (2004). *Catalyst for change: Pay for performance in Denver (Final Report)*. Boston, MA: Community Training and Assistance Center (CTAC). https://www.ctacusa.com/wp-content/uploads/2013/11/CatalystForChange.pdf

Slotnik, W. J., Smith, M. D., Helms, B. J., & Qiao, Z. (2013). *It's more than money: Teacher Incentive Fund-Leadership for educators' advanced performance, Charlotte-Mecklenburg Schools.* Boston, MA: Community Training and Assistance Center (CTAC). http://ctacusa.com/wp-content/uploads/2013/11/MoreThanMoney.pdf

Slotnik, W. J., Smith, M. D., & Liang, G. (2013). *Focus on Rhode Island: Student learning objectives and evaluation.* Boston, MA: Community Training and Assistance Center (CTAC). http://www.ctacusa.com/wp-content/uploads/2013/11/FocusOnRhodeIsland.pdf

South Carolina Department of Education. (2017). Poverty Rating Indicator in PowerSchool. https://ed.sc.gov/newsroom/school-district-memoranda-archive/poverty-rating-indicator-in-powerschool/poverty-rating-indicator-in-powerschool-memo/

South Carolina Department of Education (2020). Active student headcounts. https://ed.sc.gov/data/other/student-counts/active-student-headcounts/

Southern Regional Education Board. (2013). *Toward better teaching - a view of evaluation policies, practices, and lessons in SREB states*.

Spillane, J. P. (1998). State policy and the non-monolithic nature of the local school district: Organizational and professional considerations. *American Educational Research Journal*, *35*(1), 33-63.

Spillane, J. P. (1999). State and local government relations in the era of standards-based reform: Standards, state policy instruments, and local instructional policy making. *Educational Policy*, *13*(4), 546-572.

Steinberg, M. P., & Donaldson, M. L. (2016). The new educational accountability: Understanding the landscape of teacher evaluation in the post-NCLB Era. *Education Finance and Policy*, *11*(3), 340-359.

Steinberg, M. & Sartain, L. (2015). Does teacher evaluation improve school performance? Experimental evidence from Chicago's Excellence in Teaching Project. *Education Finance and Policy, 10*(4), 535-572.

Stronge, J. H., Ward, T. J., Tucker, P. D., & Hindman, J. L. (2008). What is the relationship between teacher quality and student achievement? An exploratory study. Journal of *Personnel Evaluation in Education, 20*(3/4), 165-184.

Strunk, K., Weinstein, T., & Makkonen, R. (2014). Sorting out the signal: Do multiple measures of teachers' effectiveness provide consistent information to teachers and principals? *Education Policy Analysis Archives, 22*(100).

Sutcher, L., Darling-Hammond, L., Carver-Thomas, D. (2019). Understanding teacher shortages: An analysis of teacher supply and demand in the U.S. *Education Policy Analysis Archives, 27*(35).

Tucker, P.D. and Stronge, J. H. (2005). *Linking teacher evaluation and student learning*. Alexandria, VA: ASCD.

Tuytens, M. & Devos, G. (2009). Teachers' perception of the new teacher evaluation policy: A validity study of the Policy Characteristics Scale. *Teaching and Teacher Education, 25*(6), 924-930.

Tuytens, M. & Devos, G. (2010). The influence of school leadership on teachers' perceptions of teacher evaluation policy. *Education Studies, 36*(5), 521-536.

U.S. Department of Education [USDOE]. (2010a). *A blueprint for reform: The reauthorization of the Elementary and Secondary Education Act. Washington, DC: Author*. The White House. http://www.whitehouse.gov/sites/default/files/docs/settingthepacerttreport_3-2414_b.pdf

van den Berg, R., Vandenberghe, R., & Sleegers, P. (1999). Management of innovations from a cultural-individual perspective. *School Effectiveness and School Improvement, 10,* 321-351.

Varlas, L. (2009). Making teacher evaluations meaningful. Education Update, 51:11, ASCD, Alexandria, VA.

Weatherley, R., & Lipsky, M. (1977). Street-level bureaucrats and institutional innovation: Implementing special-education reform. *Harvard Educational Review*, *47*(2), 171-197.

Weisberg, D., Sexton, S., Mulhern, J. and Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Washington, DC: New Teacher Project.

Welsh, M. E. (2011). Measuring teacher effectiveness in gifted education: Some challenges and suggestions. *Journal of Advanced Academics, 22*(5), 750-770.

Wilkerson, D. J., Manatt, R. P., Rogers, M. A., & Maughan, R. (2000). Validation of student, principal and self-ratings in 360º feedback® for teacher evaluation. *Journal of Personnel Evaluation in Education, 14*(2), 179-192

APPENDIX A

SURVEY QUESTIONS (STUDY 1)

Directions: This survey seeks information from teachers and administrators on educator evaluation in the state. Questions are about Student Learning Objectives (SLOs), classroom observations, and your educational background. Your responses will be anonymous, and I will be the only person who has access to the data. Your views are valued. Thank you for your participation!

Part 1: Impact of Student Learning Objectives (SLOs)

To what extent do you agree that using SLOs…	Strongly Disagree	Disagree	Agree	Strongly Disagree	I don't know
evaluates teacher performance effectively?	1	2	3	4	9
improves teachers' instructional practice?	1	2	3	4	9
promotes student learning?	1	2	3	4	9
informs teacher professional development?	1	2	3	4	9

Part 2: Understanding and Knowledge about Student Learning Objectives (SLOs)

How much knowledge do you have about…	No knowledge	Limited knowledge	Some knowledge	Substantial knowledge
the purpose of SLOs?	1	2	3	4
student groups to be included in SLOs?	1	2	3	4
the content to be included in SLOs?	1	2	3	4
implementation of SLOs in the district?	1	2	3	4
developing high quality SLOs?	1	2	3	4
selecting appropriate assessments for SLOs?	1	2	3	4
setting growth targets for SLOs?	1	2	3	4
instructional strategies to meet SLOs targets?	1	2	3	4
analyzing student assessment data in SLOs?	1	2	3	4

Part 3: Support Needed in Implementing Student Learning Objectives (SLOs)

How much support do you need in…	Need no support	Need some Support	Need a lot of Support
understanding standards in SLOs?	1	2	3
implementing standards in SLOs?	1	2	3
understanding cognitive levels of standards in SLOs?	1	2	3
developing assessments for SLOs?	1	2	3
setting growth targets for SLOs?	1	2	3
analyzing assessment data in SLOs?	1	2	3

Part 4: Impact of Observational Rubric

To what extent do you agree that using Observational Rubric…	Strongly Disagree	Disagree	Agree	Strongly Disagree	I don't Know
evaluates teacher performance effectively?	1	2	3	4	9
improves teachers' instructional practice?	1	2	3	4	9
promotes student learning?	1	2	3	4	9
informs teacher professional development?	1	2	3	4	9

Part 5: Other questions

1. In which school do you work?
2. Which of the following best describes your current position?
   Teacher Administrator Other (please specify)
3. What is the highest educational degree you have attained?
   Bachelor Master Doctorate Other (please specify)
4. How many years have you been working in the field of education?
   0-3 4-6 7-9 10+
5. Including the current year, how many years have you been using SLOs?
   0-1 1-2 2-3 3+
6. Have you participated SLOs training(s)?
   Yes No
7. Does your school participate in TAP?
   Yes No I don't know

APPENDIX B

SURVEY QUESTIONS (STUDY 2)

Directions: This survey seeks information from teachers and administrators on educator evaluation in the state. Questions are about Student Learning Objectives (SLOs), classroom observations, and your educational background. Your responses will be anonymous, and I will be the only person who has access to the data. Your views are valued. Thank you for your participation!

Part 1: Impact of Student Learning Objectives (SLOs)

To what extent do you agree that using SLOs…	Strongly Disagree	Disagree	Agree	Strongly Disagree	I don't Know
evaluates teacher performance effectively?	1	2	3	4	9
improves teachers' instructional practice?	1	2	3	4	9
promotes student learning?	1	2	3	4	9
informs teacher professional development?	1	2	3	4	9

Part 2: Understanding and Knowledge about Student Learning Objectives (SLOs)

How much knowledge do you have about…	No knowledge	Limited knowledge	Some knowledge	Substantial knowledge
the purpose of SLOs?	1	2	3	4
student groups to be included in SLOs?	1	2	3	4
the content to be included in SLOs?	1	2	3	4
implementation of SLOs in the district?	1	2	3	4
developing high quality SLOs?	1	2	3	4
selecting appropriate assessments for SLOs?	1	2	3	4
setting growth targets for SLOs?	1	2	3	4
instructional strategies to meet SLOs targets?	1	2	3	4
analyzing student assessment data in SLOs?	1	2	3	4

Part 3: Support Needed in Implementing Student Learning Objectives (SLOs)

How much support do you need in…	Need no support	Need some Support	Need a lot of Support
understanding standards in SLOs?	1	2	3
implementing standards in SLOs?	1	2	3
understanding cognitive levels of standards in SLOs?	1	2	3
developing assessments for SLOs?	1	2	3
setting growth targets for SLOs?	1	2	3
analyzing assessment data in SLOs?	1	2	3

Part 4: Impact of Observational Rubric

To what extent do you agree that using Observational Rubric…	Strongly Disagree	Disagree	Agree	Strongly Disagree	I don't Know
evaluates teacher performance effectively?	1	2	3	4	9
improves teachers' instructional practice?	1	2	3	4	9
promotes student learning?	1	2	3	4	9
informs teacher professional development?	1	2	3	4	9

Part 5: Other questions

1. In which district do you work?
2. Which of the following best describes your current position?
   Teacher Administrator Other (please specify)
3. What is the highest educational degree you have attained?
   Bachelor Master Doctorate Other (please specify)
4. How many years have you been working in the field of education?
5. How many years have you been using SLOs?
6. Do you have additional thoughts about SLOs and teacher evaluation in general?

APPENDIX C

ALIGNMENT OF SURVEY QUESTIONS

Survey Questions	Alignment
Impact of SLOs	
Using SLOs evaluates teacher performance effectively	SLOs impact
Using SLOs improves teachers' instructional practice	SLOs impact
Using SLOs promotes student learning	SLOs impact
Using SLOs informs teacher professional development	SLOs impact
Perceived Knowledge of SLOs	
Knowledge about purpose of SLOs	SLOs understanding
Knowledge about student groups to be included in SLOs	SLOs understanding
Knowledge about content to be included in SLOs	SLOs understanding
Knowledge about implementation of SLOs in the District	SLOs understanding
Knowledge about developing high quality SLOs?	SLOs understanding
Knowledge about selecting appropriate assessments for SLOs	SLOs understanding
Knowledge about setting growth targets for SLOs	SLOs understanding
Knowledge about instructional strategies to meet SLOs targets	SLOs understanding
Knowledge about analyzing student assessment data in SLOs	SLOs understanding
Support Needed for Implementing SLOs	
Support needed in understanding standards in SLOs	SLOs implementation
Support needed in implementing standards in SLOs	SLOs implementation
Support needed in understanding cognitive levels of standards in SLOs	SLOs implementation
Support needed in developing assessments for SLOs	SLOs implementation
Support needed in setting growth targets for SLOs	SLOs implementation
Support needed in analyzing assessment data in SLOs	SLOs implementation
Impact of Classroom Observations	
Using Observation evaluates teacher performance effectively.	Observation impact
Using Observation improves teachers' instructional practice.	Observation impact
Using Observation promotes student learning.	Observation impact
Using Observation informs teacher professional development.	Observation impact

APPENDIX D

SLOS TEACHER INTERVIEW PROTOCOL

Directions:
You are invited to share your views of using Student Learning Objectives (SLOs) in evaluating teacher effectiveness. The sole purpose of this interview is for my dissertation research. Your personal information will be kept confidential. This interview is estimated to last about 30 minutes, and you will be paid $30 as an appreciation of your time. Do you have any questions or concerns?

Questions:
1. Let's start with your experience of using SLOs. How long have you been using SLOs? Have you received any trainings about SLOs? If so, how effective were the trainings?
2. Based on your experience, does using SLOs change your instructional practice? If so, how?
3. Based on your experience, does using SLOs change student learning outcomes? If so, how?
4. Are there any successes or benefits of using SLOs in evaluating teacher effectiveness? If so, what are they?
5. Are there any challenges or obstacles of using SLOs in evaluating teacher effectiveness? If so, what are they?
6. In implementing SLOs, what types of assessment are used in evaluating your students' academic growth? Who choose(s) the assessment methods? What are your opinions on the assessment methods?
7. Do you feel confident to use SLOs? Do you need any support in implementing SLOs? If so, what support do you need?
8. Observational rubrics are commonly used to evaluate teacher effectiveness. Do you think SLOs is an additional reliable method in the evaluation of teacher effectiveness? If so, why?
9. Do you have any additional thoughts that you would like to share with me?

Closing Comments:
Thank you so much for your time! Your views about using SLOs in evaluating teacher effectiveness are truly valued.

APPENDIX E

AN EXAMPLE OF SLOS TEMPLATE

☐ This SLO serves as the Professional Growth and Development Plan (Section I only)
☐ This SLO serves as **one** of multiple goals of the Professional Growth and Development Plan. (Section I and II)

Section I. SLO

Teacher Name: Click here to enter text.	**Teacher School:** Click here to enter text.
SLO Evaluator Name: Click here to enter text. **SLO Evaluator Position/Role:** Click here to enter text.	
Grade Level: Click here to enter text.	**SLO Content Area:** Click here to enter text.
SLO Type: *Choose One* ☐Individual (*written by an individual teacher)* ☐Team *(team of teachers focus on a similar goal but are held accountable for only their students)*	**SLO Approach:** *Choose One* ☐Class *(covers all of the students in one class period i.e., 2nd period Biology, 4th period Beginning Pottery, etc.)* ☐Course (c*overs all of the students enrolled in multiple sections of the course (i.e., all of a teacher's Biology 2 students, all of a teacher's Beginning Pottery students, etc.)*
SLO Interval of Instruction *Choose One* ☐ Year ☐ Semester ☐ Other If *Other*, provide rationale (i.e. quarter long course) and indicate days of instruction. Rationale: Click here to enter text. Days of Instruction: Click here to enter text.	**Assessment Dates** Pre-Assessment Date: Click here to enter text. Post-Assessment Date: Click here to enter text.
I. Student Population	

	Provide a detailed description of the student population. Information should include, but is not limited to, the following: the number of students in the class, a description of students with exceptionalities (e.g., learning disability, gifted and talented, English language learner [ELL] status, etc.), and a description of academic supports provided to students (e.g., extended time, resource time with EC teacher, any classroom supports that students receive to help them access the core curriculum).
II.	**Historical and Trend Data** Describe the applicable past data for the students. In your description included the students' level of knowledge prior to instruction, including the source(s) of data (e.g., formative and summative assessments, anecdotal data gathered from collaboration with other educators) and reflect on the relevance to the overall course objectives.
III.	**Baseline Data** Describe which pre-assessment(s) will be used to measure student learning and why the assessment is appropriate for measuring the objective(s). Provide baseline assessment results for the student population. Attach the assessment and grading scale and/or rubric used to score the assessment(s).
IV.	**Post Assessment** Indicate what assessment will be used as a post assessment and how it is aligned to the baseline assessment.
V.	**Progress Monitoring Plan** How frequently will you progress monitor students' mastery of content? Indicate what ongoing sources of evidence you will collect in order to monitor student progress. (Other evidence of student growth can include student work samples, portfolios, etc.)
VI.	**Learning Goal (Objective)** Provide a description of what students will be able to do at the end of the SLO Interval. The Learning Goal (objective) is based on and aligned with course- or grade-level content standards and curriculum. The goal should be broad enough to capture major content, but focused enough to be measurable.
VII.	**Standard (s)** Identify the content standard(s) and indicators that align to the SLO learning goal (objective).
VIII.	**Growth Targets** *A. Choose One* ☐ Tiered ☐ Individual ☐ Targeted (*Sub population(s) of students are the focus of the SLO goal. Appropriate for course approach as a second SLO when the first includes all students.)* B. Considering all available data, identify the targets the students are expected to reach by the end of the SLO interval. List the growth target information below or on an attached spreadsheet.

C. Provide a rationale for the growth targets. Rationale may reflect typical vs. pretest performance, may include reasoning for using individualized targets for some but not all students, or any other influencing information used to determine anticipated growth.

IX. Instructional Strategies

A. Describe the best instructional practices you will use to teach this content to students. Include how instruction will be differentiated based on data. What interventions will be used if more assistance is needed during the learning progress?

B. Around which SCTS 4.0 Rubric Indicator(s) will you focus your professional learning

Choose an item.

X. Conference Reflection

A. Percentage of Students Who Met Growth Targets

_____ %

B. Reflection on Data

How does the data inform your instructional practice, goal setting, or your professional development for next year?

Conference	**Date**	**Signatures**
SLO Preliminary Conference		
SLO Mid-Course Conference		
SLO Summative Conference		

Section II. Additional Professional Growth and Development Goals

Evidence that the supervisor will consider in determining progress/goal

Area to be addressed: **(optional)** **South Carolina Teaching Standard Indicator(s):** Choose an item.	**Area to be addressed:** **(optional)** **South Carolina Teaching Standard Indicator(s):** Choose an item.
Goal 2:	**Goal 3:**
Strategies:	**Strategies:**
Desired Outcome:	**Desired Outcome:**
Reflect how these goals are related to your Professional Learning: (Teacher and Supervisor)	

accomplishment:

Preliminary performance review *(to be completed by the supervisor on the basis of the evidence)*

___The educator has *met* the above goal.
___The educator is making *satisfactory progress* toward achieving this goal.
___The educator is *not* making satisfactory progress toward achieving this goal.

Comments

The signatures below verify that the teacher has received written and oral explanations of the preliminary performance review.

Teacher _____ Date: _____

Supervisor: _____ Date: _____

Final performance review *(to be completed by the supervisor on the basis of the evidence*)

___ The educator has *met* the above goal.
___ The educator is making *satisfactory progress* toward achieving this goal.
___ The educator is *not* making satisfactory progress toward achieving this goal.

Comments

The signatures below verify that the teacher has received written and oral explanations of the final performance review.

Teacher _____ Date: _____

Supervisor: _____ Date: _____

APPENDIX F

IRB APPROVAL LETTER

UNIVERSITY OF
SOUTH CAROLINA

OFFICE OF RESEARCH COMPLIANCE

INSTITUTIONAL REVIEW BOARD FOR HUMAN RESEARCH
APPROVAL LETTER for EXEMPT REVIEW

Xumei Fan
College of Education
Department of Educational Psychology
820 Main Street, Wardlaw 004
Columbia, SC 29208
Re: **Pro00089844**

Dear Ms. Xumei Fan:

This is to certify that the research study ***A Mixed* Method Study of Student Learning Objectives (SLOs) in Evaluating Teacher Effectiveness** was reviewed in accordance with 45 CFR 46.104(d)(2) and 45 CFR 46.111(a)(7), the study received an exemption from Human Research Subject Regulations on **10/3/2019**. No further action or Institutional Review Board (IRB) oversight is required, as long as the study remains the same.

However, the Principal Investigator must inform the Office of Research Compliance of any changes in procedures involving human subjects. Changes to the current research study could result in a reclassification of the study and further review by the IRB.
Because this study was determined to be exempt from further IRB oversight, consent document(s), if applicable, are not stamped with an expiration date.
All research related records are to be retained for at least three (3) years after termination of the study.
The Office of Research Compliance is an administrative office that supports the University of South Carolina Institutional Review Board (USC IRB). If you have questions, contact Lisa Johnson at lisaj@mailbox.sc.edu or (803) 777-6670.
Sincerely,

Lisa M. Johnson
ORC Assistant Director and IRB Manager