

Abstract

Evaluating the impact of new teachers on student learning is a requirement for accredited teacher preparation programs. This article shares findings from a mixed methods, multiple case study investigating P-12 student learning in the classrooms of six graduates two-three years after completing a teacher preparation program. Data collection included student engagement surveys, de-identified student growth percentile scores, teacher selected pre-post assessment data, and structured phone interviews with each graduate. Results indicated a majority of students from participants' classrooms demonstrated learning growth, and participants viewed their teaching as effective. Implications for programs include pre-service opportunities to identify and respond to authentic student engagement and opportunities to build collective teacher efficacy. Additional insights describe pre-service teacher training to measure student progress and measure impact based on classroom assessments. The investigation provides a replicable case study design for teacher educators to examine relationships between teacher preparation, program graduates, and P-12 student outcomes.

Keywords: teacher preparation, student growth and achievement, efficacy, case study, graduate impact

CASE STUDY EXAMINATION OF COMPLETER IMPACT

Multiple Measures of Student Learning: A Case Study Examination of Completer Impact

In recent years, accountability measures for teacher preparation programs based on student-learning outcomes have been emphasized by the U.S. Department of Education (2011) and accrediting agencies (Tatto, Richmond, & Andrews, 2016) who seek to “judge programs by examining how well P-12 students learn and develop from the completers of the preparation program” (CAEP, 2015, p. 1). As declared by a Task Force of the American Psychological Association (2014), evaluating the impact of new teachers on student learning “is arguably the most critically needed type of data to engage in a cycle of evaluation and continuous improvement” (p. 15). This focus on preparation impact has been associated with emerging literature suggesting that teacher preparation programs (TPPs) vary in the extent to which recent graduates contribute to student achievement gains (Boyd, Grossman, Lankford, Loeb, & Wycoff, 2008; Gansle, Noell, & Burns, 2012; Goldhaber, Liddle & Theobald, 2012).

The research presents conflicting justification in using student learning and value-added measures to identify quality teaching. Effective teaching, and properly evaluating it, presents an array of challenges due to its complex nature (Chung Wei & Pecheone, 2010; Noell, Brownell, Buzick, & Jones, 2014). Critics suggest student achievement data are not always a strong indicator of teaching effectiveness and should not be used in isolation of other measures (Baker et al., 2010; Everson, Feinauer, & Sudweeks, 2013; Toch & Rothman, 2008) or “erroneous conclusions about teachers” could be implied (Marzano, Frontier, & Livingston, 2011, p. 103). As made evident by Cochran-Smith and Zeichner (2005), arguments about using data for evaluation exist due to complex variables involved in linking outcomes measures to preparation programs. Hattie (Visible Learning, 2018) recognized 252 influences related to student learning outcomes categorized in six main areas: the student, the home, the school, the curricula, the teacher, and teaching and learning approaches. Of these, the teacher (i.e., attributes, interactions, and education) was recognized as contributing 14 factors. However, a teacher’s initial training programs was found to have an effect size of 0.10, indicating only a small likelihood of impact on student achievement. Teaching effects related to student learning and instructional strategies, along with implementation methods, contributed 103 factors. Hattie’s results, alongside cautions concerning the use of achievement

CASE STUDY EXAMINATION OF COMPLETER IMPACT

data, should be considered before evaluating TPPs in this manner. As Tschannen-Moran and Johnson (2011) stated, it is not possible to ascertain whether a teacher or a school is effective or not, only to identify perceptions, at a given moment and under specific conditions, of these beliefs.

In a study that surveyed perceptions of all first-year teachers in New York City representing graduates from over 30 TPPs, Boyd et al. (2008) found initial indicators that pre-service preparation can influence the effectiveness of teachers, particularly those in their first year. The study estimated the effects of preparation program features on teachers' value-added to student test-score performance. Findings linked the amount of practice teaching during preparation as a benefit to first-year teachers. However, uncertainty remained regarding the extent to which value-added measures of student achievement are actually good measures of student learning or of teachers' impact on learning (p. 6).

An extensive literature summary report prepared by the Regional Educational Laboratory (REL) Central provided a review of research that associated educator preparation to both teacher and student outcomes (REL Central, 2014). The thorough review of 1,891 publications resulted in only 56 studies where research connected teacher preparation to any type of outcomes; 24 focused on student outcomes, generally based on state standardized achievement tests of English/language arts or math. The examined aspects included characteristics of a preparation program (e.g., number of credits, a co-teaching model), type of program (traditional or alternative), performance of multiple or individual programs based on value-added results, perception surveys of pre-service candidates, in-service teachers and supervisors, and also types of licenses or degrees. Only three of these studies implemented case study design. Since the research described in this manuscript followed case study design, these three studies were of particular interest. The first included evaluation of performance assessments of two pre-service elementary teachers through surveys and focus groups (Chung, 2008). The second case study involved 16 pre-service teachers' use of universal design evaluated through lesson plans, unit assessments, and reflection (Frey, Andres, McKeeman & Lane, 2012). The final study involved an examination of characteristics of seven effective alternative certification programs (Humphrey, Wechsler & Hough, 2008). The review did not

CASE STUDY EXAMINATION OF COMPLETER IMPACT

include any case studies of in-service teachers, linking teachers or programs to student outcomes, nor did any studies examine student learning beyond value-added or achievement measures.

Acknowledging the issues of connecting teacher preparation to student outcomes, accrediting bodies encourage TPPs to build evidence around their choice of measures (CAEP, 2015), utilize case study methodology, and make arguments for decisions to better understand graduate performance. Research that examines P-12 student learning of in-service teachers using classroom-based, student-centered, and multiple measures is needed to further understandings of how student data can be used to inform and improve teacher preparation. A number of challenges have been noted by the American Psychological Association (APA, 2014) for TPPs to attain evidence of new teachers' impact on student learning: (a) completers are dispersed in many different schools, districts and locations; (b) appropriate measures are difficult given the range of content and levels taught; (c) the financial demand of gathering student learning measures is prohibitive; (d) teachers are responsible for much more than state standards that are tested; (e) heterogeneity of student starting points are challenging for interpretation; and (f) methodological challenges are posed by pre-test and post-test data. Additionally, trustworthiness depends on how value-added modeling is both defined and calculated. As David (2010) noted, "...to provide trustworthy estimates of teacher effectiveness, value-added measures require complicated formulas that take into account as many influences on student achievement as possible" (p. 81). Furthermore, value-added measures and teacher-linked student achievement data are not readily available as many states do not have systems that connect student scores to individual teachers.

Drawing on descriptive case studies of six practicing teachers who graduated from one teacher preparation program two to three years prior, this study describes student learning via engagement, achievement and growth measures, and participant perceptions of learning factors as outcome indicators of program quality. This study is part of a larger case study project which investigated knowledge, skills, and dispositions of TPP completers and preparation program evaluation [Author, 2019a]. Since the prior study found that completers successfully met professional expectations for knowledge, skills, and

CASE STUDY EXAMINATION OF COMPLETER IMPACT

dispositions, researchers used the current study to investigate learning outcomes of P-12 students in those teachers' classrooms. The study gathered data related to the following questions:

- Do P-12 students in the classrooms of TPP completers demonstrate expected levels of learning and development?
- What factors do graduates perceive impact students' learning and development?
- How can a TPP improve program quality based on results of P-12 student learning in classrooms in which graduates are teaching?

Reflective Experiential Framework

This study was grounded in the *Reflective Experiential Teacher* preparation framework as a guide to both research design and case study analysis. The framework encourages learning through reflection and contemplation of beliefs and experiences as knowledge (Dewey, 1938). These experiences occur through observation of learning as cooperating teachers in the field model effective pedagogy, as well as practice teaching paired with targeted feedback, the essence of social cognitive theory (Bandura, 1989). Experiences must be thought about critically within social, cultural, and environmental contexts as an educator continues to refine skills that help students learn. These experiences and skills affect teachers' prior knowledge and how they reflect on theory and evidence-based practices in this constructivist frame.

Inherent in the framework are the 10 core teaching standards of the Interstate Teacher Assessment and Support Consortium (InTASC) and associated skill progressions as they relate to the complexity of teaching practices educators employ, becoming more refined with practice over time (CCSSO, 2013). The framework includes the knowledge, skills, and dispositions of the standards that impact students aligned to four general categories. Learner and Learning, Content, Instructional Practice, and Professional Responsibility. These categories provide constructs for content validity of data collection measures, a valid system for evaluation of teacher competencies and overall performance (ND DPI, 2015), a foundation on which teaching practice and its effect on student growth and achievement can be based, and a system of concepts, assumptions, expectations, and beliefs that inform inquiry (Maxwell, 2005).

CASE STUDY EXAMINATION OF COMPLETER IMPACT

Related to this study, the framework emphasizes how professional growth takes place through teachers' own reflections on experiences and those of their students, as well as the cyclical informal and formal teacher evaluation and professional growth processes found in schools.

Methods

A descriptive, multiple case approach was utilized to describe outcomes of students in the classrooms of six TPP graduates. The four elementary teachers and two secondary teachers had completed a traditional teacher education program at a regional, four-year institution located in the Midwest. The graduates are interchangeably referred to in the multiple roles they represent: program completers, study participants, graduates, and teachers. Given the challenges to attaining teacher impact data identified by the APA (2014) and the recommendation of accrediting agencies to utilize case study for program evaluation (CAEP, 2015), a case study methodology was selected. Case study allowed data collection from several sources to provide a multi-dimensional examination of the interrelated attributes of teaching and learning (Kennedy, 2016). The case study design rigor of Yin (2014) provided recognized research methods and an explicit set of procedures. The constructivist-education epistemology of Merriam (1998) was also used to inform methodological decisions. This orientation matched well the constructivist conceptual framework underpinning a study on TPP completer outcomes and teacher influence of student learning (Yazan, 2015). As a descriptive case study, there is not an attempt to control variables, infer direct causality of teacher preparation, effectiveness of teaching on the results of student learning, or generalization to all graduates, but instead to inform data-based decisions regarding TPP efforts for continuous improvement.

Participants

Participants were selected through purposeful sampling. The goal was not to randomly select completers to make generalizations, but instead to select participant cases that could be illustrative and best enable answers to the research questions. The six graduates from the TPP demonstrate a perspective within a defined context and with enough information to portray a detailed picture (Merriam, 1998, p. 62) of student learning in their classrooms. Recruitment criteria included program completers who were

CASE STUDY EXAMINATION OF COMPLETER IMPACT

actively teaching full-time within an 80-mile radius of the TPP and with a program completion date from one to three years prior. Initially, all graduates were contacted via email and invited to participate. Two responded and consented. A second recruitment script was sent out the following semester including additional graduates meeting the graduation date criteria. Two elementary teachers and two secondary education teachers responded and consented to participate resulting in six total participants. Table 1 provides a description of each participant referenced using pseudonyms: Jamie, Terry, Reese, Stacy, Taylor, and Alex. Graduate participation was dependent on administrative agreement for research to be conducted; all principals agreed.

[Table 1 here]

Instruments

The existing data set from the prior study [Author, 2019a] was expanded to include four additional participants and reanalyzed for this manuscript. The additional data were collected over three months using multiple sources: (1) student achievement and growth scores; (2) student pre and post assessment data; (3) student engagement surveys; and (4) structured interview questions about teaching impact on student learning. Use of multiple measures provided a more complete description of student growth and completers' success in the classroom and followed recommendations of best-practices to investigate teaching effectiveness [Author, 2019b].

Participants were asked to submit digital copies of de-identified, teacher-linked P-12 student achievement data which resulted in varied metrics. Five participants submitted data from the Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP); one participant (Reese) submitted Renaissance STAR assessment data. The request specified that fall and spring scores for all students assigned to participants' classroom be included, if available, to document evidence of learner growth and ascertain if students met expected levels of achievement.

As an additional measure of student learning growth, participants were also asked to submit de-identified data from pre and post assessments. It was left to participants' discretion what type of assessment to submit. This decision acknowledged the expertise and choice of participants, and that some

CASE STUDY EXAMINATION OF COMPLETER IMPACT

districts have formalized benchmarking processes and others do not. Submitted pre-post data included unit tests, results of standards-based, benchmarking scales, progress monitoring scores, and rubric-scores for first and final drafts of two English papers. One participant did not submit pre or post assessment data. Two candidates submitted copies of actual student work from which the research team processed results to serve as evidence of student learning; the other three candidates submitted outcomes of assessment measures.

Anonymous student engagement surveys using Schlechty's (2002) Levels of Engagement were distributed to students in participants' classrooms as another indicator of student success. Academic engagement is associated with higher academic achievement and lower dropout rates (Glanville & Wildhagen, 2007; Klem & Connell, 2004). Surveys were adapted for readability for elementary students. Students were asked to look back over the past week and choose the option that best described the way they felt about the class and work they were asked to do, rating their level of engagement on the five-option scale: (a) *Authentic*: high attention/high commitment; (b) *Ritual*: high attention-low commitment; (c) *Passive*: low attention-low commitment; (d) *Retreatism*: no attention-no commitment; or (e) *Rebellion*: diverted attention-no commitment. For elementary completers, surveys were distributed by a participants' colleague with whom the students had a prior relationship (e.g., counselor or another teacher). This occurred in the participant's classroom without them being present, and survey options were read out loud to students. The survey was distributed as a paper copy and returned as de-identified, scanned copies to the data manager, who then compiled results and provided them to the analyzing researchers. Secondary education program completers were asked to choose a class and administer an electronic version of the survey twice approximately two weeks apart while teaching the same unit. Surveys were administered through a link shared with students using the Qualtrics survey platform. Surveys were completed anonymously and submitted directly to the data manager. Mean scores from both administrations was calculated and aggregate results were provided to participants.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

Teacher interview.

Targeted interview questions were drafted and revised by the research team to gain explanation and personal views of participants on factors related to student learning outcomes (Yin, 2014). Questions about the impact of teaching were generated around the outcomes the program was designed to achieve, the pre-existing codes from the InTASC standards of Learner and Learning, Content, Instructional Practice, and Professional Responsibility (interview protocol available upon request). Questions within each code were designated as inquiry about student learning and teacher impact as qualitative evidence of factors that impact student growth.

Data Collection

Data collection was directed by a protocol timeline for consistency across cases. As quantitative data sources were submitted via email, results were coded for anonymity, de-identified by the data manager, and provided to the researchers conducting analysis. Phone interviews with participants were completed by a member of the research team not employed by the participant's preparation program nor involved in their training; interview questions were provided to participants in advance. Participants were notified by email when data collection concluded with information about the completion timeline and member checking.

Data Analysis

Data analysis occurred in four phases: document review of quantitative data, analysis of qualitative interview data, a case-by-case analysis, and final cross-case consideration. Descriptive statistics were used to define quantitative data from achievement/growth scores, pre-post assessment results, and student engagement surveys. A template was created to populate results for each data source for consistency. Data from interviews were systematically analyzed using a combination of strategies from Hill, Thompson, and William's (1997) *A Guide to Consensual Qualitative Research* (CQR) and an item-by-item analysis to identify core ideas in an inductive manner. The lead researcher assumed the role of initial evaluator of interview data and the second researcher as an auditor of the initial evaluator's review. Again, a template was created to organize participants' responses for each of the items on the

CASE STUDY EXAMINATION OF COMPLETER IMPACT

interview protocol. Classification began via pre-existing codes aligned with the four InTASC constructs. Through constant comparison and reconceptualization codes were confirmed using categorization (Creswell, 2007; Yin, 2014). An independent examination was conducted by the auditing researcher to determine consensus of findings.

Within each case the constant comparative method of qualitative analysis was used (Glaser & Strauss, 1967) to construct themes through continuous comparison of data (Merriam, 1998). The initial evaluator completed a cross analysis of findings for each case to determine emerging patterns related to the contribution of TPP graduates to student outcomes. Guidelines of thematic analysis were used to ensure reliability (Braun & Clark, 2006) in the review of results for each case and across cases. This allowed for a constant comparison throughout data in formulating the final propositions about P-12 student learning.

Trustworthiness

Trustworthiness was established in several ways. A case study protocol developed from best practices in educational research (Merriam, 1998; Yazan, 2015; Yin, 2014) was followed, and replication logic was used across all cases (Yin, 2014). In addition, constructivist concepts of the *Reflective Experiential Teacher* framework and the InTASC constructs grounded design decisions and instrument selection. Data from multiple measures served to capture different dimensions of student growth and achievement. A data manager, who did not conduct analysis, maintained the research database and coded all data to maintain anonymity. Interviews were conducted by a researcher who was training in interviewing techniques, was not employed by the TPP, and had not participated in preparing the participants for teaching. Standardized interviews were conducted individually, and each interview followed the structured protocol. Finally, member checking occurred for respondent validation of results.

Case Study Findings

Multiple sources of data for examining the growth and achievement of students in the classrooms of six TPP completers and their perception of impact on students were examined. The convergence of data is an attempt to describe student learning growth and the teachers' contribution to the learning, as

CASE STUDY EXAMINATION OF COMPLETER IMPACT

well as to inform the TPP of graduate outcomes in a data-based programmatic decision model. Results are described first according to each case and then cross-case findings are presented; see Table 2.

[Table 2 here]

Case 1: Jamie

A majority of Jamie's 4th grade students exhibited growth in learning across several content areas including reading, language arts, and math according to multiple sources of evidence. On the engagement survey, 57% of Jamie's students self-reported they were engaged in learning either authentically (36%) or ritually (21%). Jamie submitted NWEA Measures of Academic Progress (MAP) student progress reports which included scores in math, reading, and language arts. Spring Rasch Unit Scale (RIT) scores, grade level means, and percentile scores were included. The RIT score represents a student's achievement level at any given moment against criterion norms. Jamie's students achieved above average for reading with 57.1% at or above the norm grade level mean; 21.5% of students demonstrated improvement over the course of the year. In the area of language arts 64.3% met or exceeded grade level means and 5.7% demonstrated growth. Student scores in math indicated 14% were at or above grade level means and 14.3% demonstrated growth. As the data showed, a limited number of students demonstrated growth according to the MAP tests in reading, language arts, and math.

Jamie's students demonstrated academic improvement over the school year according to pre-post assessment data. Jamie submitted de-identified AIMSweb oral reading fluency scores for each student as pre and post assessments. The AIMSweb reading curriculum-based measurement is an individually administered, standardized test of oral reading. When comparing fall scores to spring scores, Jamie's students scored as follows: 0% change in at risk students, 20% decrease for students who fell below target, 6.6% increase for on target students, and a 13.3% increase for students who were performing above target. By the spring semester, all students had demonstrated marked improvement in oral reading fluency, increasing the number of words read correctly.

Instructional reading levels, using leveled reading indicators from A-Z (A correlated with the start of kindergarten and Z the end of 5th grade) were also submitted as a pre-post assessment of student

CASE STUDY EXAMINATION OF COMPLETER IMPACT

learning growth. Fourth grade student norms for fall are levels P-Q, and the spring target level is S. All students were on target or above target for oral reading fluency at the end of the school year. In the spring, student reading levels had changed as follows: below target = 6.6% increase, on target = 6.7% increase, and above target = 13.3% decrease. The lowest level reported for Jamie's 4th grade students was an O (n = 1), reading at approximately a 3.5 grade level, and the highest was W (n = 3), reading at approximately a 6.5 grade level. In addition, the range of students' scores increased in words read correctly from 8 additional words to 79 additional words. However, the number of proficient readers as indicated by instructional reading levels decreased slightly from 93.3% to 86.7%.

Finally, phone interview results for Jamie indicated that confidence in content knowledge directly impacted student learning. Core ideas of factors impacting student learning included: strategy preference, confidence level of content, reflection, and personal goal setting. Jamie noted, "I'm pretty knowledgeable with math and science. I can see if [students] are going wrong. I struggle with language arts. Students will ask a question. If I don't know it, I'll say, 'I'll look that up and get back to you.'" Jamie also shared specific examples of how collaboration with students, colleagues, and administrators impacted their content knowledge. For example, Jamie stated, "I learn a lot from peers and coworkers" and "when my supervisor is observing. I go and talk with him about improvement." Jamie's interview statements and achievement results indicate an overall positive impact on student learning outcomes.

Case 2: Terry

According to the engagement survey results, Terry's 4th grade students were engaged and experienced learning growth over the year. In fact, 83.5% (n = 15) of Terry's students reported an authentic (39%) or ritual (44.5%) level of engagement. Terry provided de-identified NWEA MAP spring class reports for reading, language arts, science, and math, RIT scores for each student, and percentile indicators. Report summaries provided a class mean in comparison to norm grade level means, as well as the number of students at or above norm grade level mean RIT. Terry did not provide fall to spring scores for comparison, however the reports did indicate the total students at or above norm grade level means. According to the class reports, 58% of students in Terry's classroom during were at or above the

CASE STUDY EXAMINATION OF COMPLETER IMPACT

norm grade level mean in reading, 58% in language arts, 68% in science, and 42% in math. Students performed highest in the area of science given the four test categories. Terry did not submit pre-post assessment information for review.

Codes related to impact factors from Terry's interview included confidence in content, and collaboration with colleagues, students, and supervisor. Terry indicated that specific strategies, such as whole brain teaching, impacted student learning. Terry stated, "I began using [whole brain teaching] in the second half of the school year and I noticed significant changes...it doesn't work for everyone, not for every teacher, but it works for me." Terry also made note that confidence in content impacts the ability to share content knowledge with students, "I do have a grasp on content areas, not just give content from the textbook. I can break it down into smaller parts." Terry also displayed confidence in the ability to use reflection to improve teaching stating, "I'm always thinking about what we should change for next year. It is good to show [students] that teachers constantly reflect too." Terry also identified the importance of setting goals to improve teaching and learning.

Case 3: Reese

Reese's 4th grade students were engaged, showed learning growth, and understood expectations in the classroom. Of 16 students, 87.5% of them reported authentic (37.5%) and ritual (50%) levels of engagement. Data also indicated that Reese's students showed significant achievement gains in both reading and math. Reese submitted achievement data in the form of STAR assessment growth and progress reports for reading and math. STAR formerly stood for Standardized Tests of Achievement in Reading but has recently expanded to content areas other than reading. Progress report data were collected in both the fall and the winter, and percentile scores and increase in percentile rank were included. According to the de-identified data, 80% of Reese's students were at or above the 50th percentile in reading, while 93.7% were above the 50th percentile in math. Reese's students increased 73.3% in percentile ranks for both reading and math.

Reese's students showed substantial growth on the teacher made pre and post assessments for English Language Arts (ELA) and math. Reese submitted a spreadsheet with the results of pre and post-

CASE STUDY EXAMINATION OF COMPLETER IMPACT

instruction results. Scores were presented as 1 = little understanding of a concept/cannot demonstrate any mastery, 2 = partial mastery, 3 = meets targets, and 4 = exceeds targets. Prior to instruction, 100% of students performed below target (11 students demonstrated a level two score and five students demonstrated a 2.5) on the ELA assessment. After instruction, assessment, and re-teaching, 87.5% of students (n = 14) attained a level three on the post assessment. On the math pre assessment all students (n = 16) scored below target, however 93.7% (n = 15) of students attained a level 3 score on the post assessment.

Reese's interview indicated a wide variety of influences that positively affect student learning which included: understanding learners needs, designing assignments, gaining content knowledge and experience, reflection, collaboration with supervisor, and student understanding of expectations. Reese mentioned that understanding learners' needs informs design of developmentally appropriate learning activities and specifically stated, "If I understand where [students] are at, I can create activities for them that are challenging enough, but not too challenging to where it is overwhelming" and that knowing students' needs "helps me match the level of rigor of assignments to learners." Reese also mentioned that confidence with content knowledge, collaborating with administrators, and reflecting on teaching impacts student learning. Another important factor that Reese mentioned, in regards to impacting student learning, was the importance of students knowing the teacher's expectations. Reese stated, "Students understand what I expect. I want them to understand where we are going to go." These statements indicate there are several factors that positively impacted student learning in Reese's classroom.

Case 4: Stacy

A majority of Stacy's 5th grade students exhibited growth in learning across several content areas including reading, language arts, math, and science, according to multiple sources of evidence. On the engagement survey, 72.8% of Stacy's students indicated they were engaged either authentically (27.3%) or ritually (45.5%). Stacy submitted NWEA MAP student progress reports which included scores in reading, math, language arts, and science. RIT scores, grade level norms, and percentage of students who demonstrated growth were included in the report that spanned from spring 2017 to winter 2018.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

According to the de-identified scores 50% of students in Stacy's class were at or above grade level norms in reading, while 91.7% of students were at or above grade level norms in language arts. According to math results 75% of Stacy's students were at or above grade level norms, while 83.3% met or exceeded grade level norms for science. Overall, most of Stacy's students demonstrated growth in reading (75%), language arts (83.35%), and math (75%) with less growth in science (41.7%).

Although not all students' scores increased, most of Stacy's students showed growth on the pre-post assessment. Stacy submitted scanned copies of a 20-question pre and posttest in math. There was not a data table or other form of analysis submitted. The tests were labeled as cumulative versions A and B from the Saxon Math Intermediate 5th grade curriculum. Average scores increased 17% from pre to post assessment. The greatest increase from pre to posttest was 6 points and the average increase was 3.44 points. All but one of Stacy's students increased their scores on the math posttest. Additionally, results of Stacy's phone interview indicated that understanding students' needs has an impact on student learning. Stacy noted the importance of setting expectations on the first day of school, "I work to set [expectations] on the first day. It's a very big deal to set them on the first day" and "[students] know what I expect. I am firm and consistent." Stacy also explained that written reflection is used to improve instruction.

Case 5: Taylor

Most of Taylor's students showed improvements in several areas of achievement. Results of the student engagement survey show 87.6% of Taylor's students indicated they were engaged in learning, either authentically (66.3%) or ritually (21.3%). Taylor provided the spring 2017 to spring 2018 de-identified NWEA MAP student progress reports for reading and language arts. The reports included RIT scores, a percentile indicator, a class mean in comparison to norm grade level means, as well as the number of students at or above the norm grade level mean RIT. According to MAP results, 55.2% of students were at or above the norm grade level mean in reading and language arts. Additionally, 79.1% of students demonstrated growth in reading while 85.2% demonstrated growth in language arts. Pre and post assessment results indicate increases in achievement for Taylor's students in reading, language arts,

CASE STUDY EXAMINATION OF COMPLETER IMPACT

and writing. Taylor submitted teacher-made, pre-and post-instruction rubric scores for first and final drafts of three papers written in class (i.e., informative, compare/contrast, and argumentative). However, there were only two data points (pre and post) for a select number of papers. Average scores on the informative paper increased by 11.2%. Average scores for the compare and contrast paper also increased by 8.4%, with 100% of Taylor's students increasing their overall scores from the first to final draft. The greatest increase for the informative paper was seven points and three points for the compare and contrast paper. On average, students' scores increased 2.1 points on both the informative and compare and contrast papers.

Finally, phone interview results for Taylor indicated that understanding learners' needs, reflection, confidence, and collaboration with colleagues directly impacted student learning. Taylor noted that reflection helps improve instruction and specifically stated, "I use [reflection] to go over what I failed, or how each class is different." Taylor also explained that collaborating with professors from the educator preparation program and colleagues was used to improve student performance. Overall, results from multiple sources indicated that Taylor had a positive impact on student learning outcomes.

Case 6: Alex

A majority of Alex's high school science students exhibited growth in learning in general science, and 82.7% of Alex's students indicated they were engaged in learning, either authentically (33.3%) or ritually (49.4%) on the engagement survey. Alex provided class and individual student progress reports for the 2017-2018 academic year for the NWEA MAP test in science, which included RIT scores for each student and a percentile indicator. Results indicated 74% of students in Alex's classroom were at or above the grade level norm, and 87% of students demonstrated growth. Alex submitted teacher-made, pre-and post-instruction student data for the topic of ecology. A spreadsheet with the results of the pre and post assessments was submitted. Average scores on the ecology test increased by 21.8% from pre to post assessment. The median scores for pre and post assessment were 16 and 24, respectively. Scores on the pre assessment ranged from 4-30 points while the post test scores improved ranged from 6-32.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

Alex's phone interview indicated a variety of factors impacted student learning: understanding learners' needs, designing assignments, making real-life connections, planning, assessment, reflection, confidence, collaboration, and modeling lifelong learning. Alex described knowledge of learner development and how that impacts student learning, "I think understanding what level students are at helps me to determine what level assignments will work with them. It helps me match the level of rigor of assignments to learners." Alex noted the importance of planning and assessment as factors that impact student learning:

The more I plan and am structured the better the outcomes...daily I know where [students] are at and I can see if [a topic] needs retaught or presented in a different way....every single day I review the exit tickets and reflect on how to improve, to better fit what [students] need.

Alex went on to explain that collaborating with colleagues improved confidence in content knowledge, "I have PLCs [professional learning communities] that are content based. Other science teachers and I build the curriculum together and decide how [students] move through and what vocabulary they need to know." Finally, Alex described being a role model who demonstrates the importance of being a lifelong learner who values education, impacts student learning. Multiple measures of student learning growth indicated that a majority of Alex's students exhibited growth in learning in the content area of science.

Cross-Case Findings

A cross-case review was conducted to look for patterns across findings related to P-12 student learning and impact factors (Tables 8). The single case findings and cross-case analysis answer the research question, "Do P-12 students in the classrooms of TPP completers demonstrate expected levels of learning and development?" Results of multiple measures indicated students in these completers' classrooms overall did demonstrate expected, or better than expected, levels of learning and development. The analysis provides student data to inform TPPs of potential areas for continuous program improvement based on results of student learning in classrooms and teacher perceptions of influences on those student outcomes.

Student Achievement/Growth Scores

CASE STUDY EXAMINATION OF COMPLETER IMPACT

To demonstrate levels of learning and development, participants were asked to submit teacher-linked student achievement data as evidence of learner growth and to ascertain if students met expected levels of achievement. An acceptable achievement outcome is student performance at or above the 'average' student score in a norm group, the median score (Thum & Hauser, 2015). A majority of students in completers' classrooms demonstrated acceptable levels of academic achievement according to the provided measures; a majority of students were at grade level norms near or above the 50th percentile. This occurred for completers' students with two exceptions in elementary math. Of Jamie's students, only 14% were at grade level norms for 4th grade on the MAP assessment. Results for Terry's students indicated 42% tested at grade level norms, slightly under the 50th percentile (see Table 2). To note, more than 50% of students scored at grade level for reading and language arts for both Jamie and Terry, and also in science for Terry. Both secondary educators demonstrated relatively high percentages of students who recorded growth in academic achievement. For secondary educators, scores are considered only in relation to the subject area they taught. Of the students in Taylor's English classroom, 79.1% demonstrated growth in reading and 85.2% demonstrated growth in language arts; 87% of students in Alex's high school science classroom demonstrated growth on the MAP test. Overall, a majority of students taught by TPP graduates met expected norms and demonstrated growth over time after receiving instruction from the TPP graduate according to the achievement tests, with the exception of Jamie's students who demonstrated limited growth according to this measure.

Student Pre and Post Assessment Data

A majority of students in completers' classrooms demonstrated learning progress according to teacher-selected pre and post assessments. In fact, 100% of students in the classrooms of Jamie, Reese, Alex, and Taylor demonstrated growth. In Stacy's classroom only one student did not demonstrate growth on the math pre-post assessment. The student's score decreased by one incorrect question. Terry did not submit comparative pre-post data.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

Student Engagement Surveys

In total, 126 students taught by TPP graduates completed the engagement survey. A majority of students self-reported the highest two levels of engagement, either authentic or ritual ($n = 102$). The level of engagement with the highest percentage for each participant is as follows: Jamie-authentic engagement (36%), Terry-ritual engagement (44.5%), Reese-ritual engagement (50%), Stacy-ritual engagement (45.5%), Taylor- authentic engagement (66.3%), and Alex-ritual engagement (49.4%). The most commonly report level of engagement was ritual engagement. The graduate with the lowest level of authentic engagement was Alex (high school science) with 33.3% authentic engagement. The highest level of authentic engagement was self-reported by Taylor's 8th grade English students with 66.3% authentic engagement. Out of all students surveyed, students indicated passive compliance, six indicated retreatism, and three students self-reported a level of rebellion. Terry, Reese, Stacy, and Alex had no students report rebellion. No students in Reese's classroom reported the two lowest levels of retreatism or rebellion.

Teacher Interviews

Participants were interviewed to collect data related to perceptions about factors related to their ability to impact student learning. A cross-case analysis was conducted to identify commonalities across cases and answer the research question, "What factors do graduates perceive impact students' learning and development?" Cross-case interview analysis revealed three main themes aligned with the InTASC construct of Professional Responsibility: utilizing collaboration as a tool for improvement, the impact of reflection on teaching practices, and the role self-efficacy plays in student learning.

Utilizing collaboration as a tool for improvement. Interview results indicated a majority of participants (5 of 6) collaborate with colleagues on a regular basis to improve student performance. Several participants cited specific examples of the changes they implemented as a result of collaborating with supervisors. Additionally, participants mentioned the importance of collaboration with a variety of stakeholders, including parents, friends, TPP professors, and community members. The purpose of collaboration was consistent across cases as well. Most cases (4 of 6) revealed that participants used

CASE STUDY EXAMINATION OF COMPLETER IMPACT

collaboration with colleagues to brainstorm ideas that supported student learning. It is also worth mentioning that two elementary participants and both secondary education participants explained their use of collaboration with colleagues in making curricular decisions. In addition to collaboration, reflection among and between colleagues was evident in the cross-case analysis.

The impact of reflection on teaching practices. The use of reflection was cited in a majority of cases (5 of 6) as a way for teachers to improve their teaching and become more effective. Two participants mentioned their use of daily reflection to improve teaching and therefore, positively impact student learning. Similarly to collaboration and its connection to a variety of stakeholders, several participants stated that reflection is not done alone. Participants mentioned the importance of involving students, colleagues, and supervisors in the reflection process. This level of professional responsibility for student learning through collaboration and reflection directly impacts teachers' self-efficacy.

The role self-efficacy plays in student learning. As results of the interview analysis indicated, all participants reported that their teaching is effective. Each participant listed a variety of factors that impacted their ability to teach effectively. These factors included educational support within the school, importance of setting expectations, confidence in content knowledge, and understanding students' unique learning needs. The majority of participants (5 of 6) acknowledged the importance of creating a classroom community where students' interests are valued and relationship building is essential to learning. Positive student-teacher relationships were cited as factors that impact teacher effectiveness and student success. Conversely, several of the participants listed factors that have limited their self-efficacy and ability to teach effectively. These factors ranged from personal matters, such as medical issues, to lack of administrative support, and an unpreparedness to face the challenges of diverse populations and culture differences. All participants did agree that their TPP training successfully prepared them to effectively meet the needs of their learners.

Overall, data from cross-case interview analysis supported the idea that collaboration, reflection, and self-efficacy impact graduates ability to teach effectively. This, in turn, impacts student learning. Participants reflected on the variety of professional responsibilities inherent in teaching. It was evident

CASE STUDY EXAMINATION OF COMPLETER IMPACT

that graduates recognize their role in the classroom and how the effectiveness of their instruction impacts students. These results are important as TPP's examine the role they play to develop educators' ability to teach effectively.

Discussion

Teacher preparation programs can acquire important information from the graduates they follow into the field. Analysis of student learning in teachers' classrooms can assist TPPs to make data-driven decisions about improving training program quality. Patterns across the six cases answer the research question, "How can a TPP improve program quality based on results of P-12 student learning in classrooms in which graduates are teaching?". These findings also provide a model for evaluating preparation efforts. Results indicated three main considerations for program improvement: pre-service training opportunities for authentic engagement strategies, effective student data utilization, and the influence of collective efficacy.

From Ritual to Authentic Engagement

Self-reported levels of engagement indicated ritual engagement as the most common for students in participants' classrooms with the range of responses for ritual engagement from 21-50% of students. If upwards of 50% of students moved from ritual to authentic engagement, as with one participant in this study, collective student learning could increase significantly. Ritual engagement indicates a compliant classroom, one that is orderly and in which most students appear to be working. This can make it easy for teachers to conclude learning is taking place even when the work is not actually meaningful to students. Often there are extrinsic outcomes of value, such as grades, that keep students from becoming disengaged or passive. Schlechty (2002) acknowledged that students can lack authentic engagement but still do well on classroom and state level exams because they learn what they need to do well on these tests and please those in authority. It is estimated that by high school 40-60% of students are disengaged (Klem & Connell, 2004), but they may be earning acceptable marks in school and remain compliant. A more engaged classroom would benefit passive students as well as those experiencing retreatism or rebellion, and teachers could focus more time on designing and teaching engaging lessons.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

It is important for TPPs to consider how they prepare teachers to identify the presence or absence of engagement and ways to guide students in ritual engagement to more authentic emotional, behavioral, and cognitive engagement. Teachers need to know what kind of student actions identify authentic engagement, the “look fors”. Schlechty (2011) defined indicators present when a student is authentically engaged: the student is attentive to the task, persistent, voluntarily commits time, energy, and effort, and finds the work of value. Additionally, teachers need to reflect on their own actions and responses to engage students, evaluating their strategies for effectiveness. In a study conducted by Digamon and Cinches (2017), results indicated levels of engagement can be explained by teacher related factors, such as the activities given to the class (specifically cooperative group work), how the teacher employs instructional strategies, and how students experience those strategies. This consideration guides the EPP to reflect on the knowledge, skills and disposition of the InTASC construct of Instructional Practice that TPPs are designed to achieve.

Slechty (2002) also defined the qualities of engaging work that can help teachers move students to authentic levels of engagement. First, the work students do and the product or performance they create, must be significant to them; these products often require an investment of personal ideas. Authentic engagement requires clear articulation of standards for success. Students need to know exactly what is expected so they can decide if it is significant to them. For students to engage in this way, they also need a classroom that is emotionally and intellectually safe, a classroom where there is freedom from adverse consequences for failure and in which their choices, as well as efforts, are affirmed. Engagement is encouraged when students work interdependently with others (e.g., cooperative learning) and when the work has authenticity to lived experiences and real-life application. Finally, novelty and variety can add the unusual or unexpected element that triggers authentic versus ritual engagement. These qualities that generate authentic student engagement, and the skills and disposition needed for a teacher to implement them, illustrating the intricacy that is teaching and learning.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

Evaluating programs for these kinds of experiences could serve to improve better prepare completers for the responsibilities of the classroom. In addition, targeted practice with specific engagement strategies in simulated and real classrooms, paired with timely feedback, should be embedded in methods courses. It has been noted that a greater amount of practice teaching benefits teachers in their first year (Boyd et al., 2008). It would be interesting to teacher educators to further examine what teacher actions and/or behaviors students identify as engaging in the classrooms of graduates with the highest levels of engagement. Student interviews or focus groups may well serve to expand understanding to this end and assist efforts to improve preparation.

Student Data and Implications for Program Improvement

Examining P-12 student outcomes leads a preparation program to consider implications for the preparation experience. In general, a majority of students taught by TPP graduates in this study met expected norms and demonstrated growth over time. A causal assumption using student data to judge the TPP would indicate the program prepared teachers reasonably well. However, measuring the effectiveness of a teacher is complicated, and one single factor or test cannot adequately determine effective teaching; student performance should not be the determining or predominate factor (NWEA, 2013). Even though it is relatively well understood that teachers “make more of a difference in determining the quality of education than any other school-controlled resource” (Shepard, 2012, p. 8), the strongest variant in learning outcomes remains the student. As Hattie (2015) noted, what students bring to the classroom is the greatest source of variance in learning, in fact, the student constitutes about 50% of variance.

They differ greatly, they bring different attributes and prior knowledge, they have different motivations and purposes for learning, they study in varied ways, some are collaborators some are loners, they have a manifold of likes and dislikes, and they can be bright or struggling. (p. 87)

Although caution remains regarding achievement scores as estimates of teaching effectiveness, considering student results can serve to inform preparation programs along with a combination of other reliable measures.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

Student outcomes in this study do draw attention to the need for multiple measures, particularly in the case of Jamie's classroom. Jamie's students had contrasting results when comparing reading achievement test scores to pre-post assessment results. According to NWEA testing, 57.1% of students were at or above the grade level norms in reading. However, according to the AIMSweb progress monitoring, 86.7% of students were on or above target for instructional reading level (i.e., comprehension), and 100% were on or above target for oral reading fluency when comparing fall to spring results. This difference is substantial; lower student results could require intervention and educational services, while higher scores would indicate no instructional concern. This example from Jamie's classroom demonstrates how essential it is to triangulate results and look for congruence of evidence. When indicators diverge, further explanations should be sought. While teacher selected pre-post assessment data was provided to compare reading scores, it would be of interest to exam curriculum-based measurements for math as well given the low rate of students meeting grade level norms (only 14%) and the discrepancy in reading scores.

Jamie's results support pre-post classroom data that tracks student progress as a better indicator of teacher impact than achievement tests. In fact, students demonstrated higher levels of growth on pre-post assessments compared to testing data for all participants (Terry did not submit comparative growth information). That achievement data did not coincide well with norm-referenced achievement data is substantial. It makes the case that the power of achievement data is not in its potential to evaluate teacher effectiveness, but to provide a teacher and their colleagues an opportunity to assess student readiness, track growth, and plan instruction and interventions accordingly. As such, it may be more informative to a TPP to hear how the teacher analyzed the achievement and pre-post results, what they thought about the data, and what they did about the results than to examine the results themselves. This is, in essence, cognitive task analysis, the process of identifying, analyzing, and structuring the knowledge and skills experts apply when they perform complex tasks (Crandall, Klein, & Hoffman, 2006). Teaching is, after all, one of the most complicated human activities.

In response to findings of this study, a TPP could more explicitly utilize cognitive task

CASE STUDY EXAMINATION OF COMPLETER IMPACT

analysis during preparation regarding student growth measures and teacher data literacy. A TPP could also include teacher-selected pre and post assessments and corresponding analysis in required courses on educational assessment. Furthermore, candidates should be specifically instructed on methods to measure student improvement and growth instead of comparing students to other students via norms. Candidates should learn to gauge the effect they have on each students' learning by using effect size calculations (i.e., $\text{effect size} = \frac{\text{average post test score} - \text{average pretest score}}{\text{average standard deviation}}$). This could be a requirement during clinical experience that provides the TPP evidence that candidates measure P-12 students' progress and their own professional practice. A TPP that replicates this study could also adapt the interview protocol to include a line of questioning about the growth measures and what interpretation and instructional decisions the teacher made as a result.

Collective Teacher Efficacy

Through participant interviews the themes of collaboration, reflection, and self-efficacy were substantiated, and when considered jointly the impact of collective teacher efficacy on student learning emerged as a major theme. Collective teacher efficacy is defined as, "the perception of teachers in a school that the efforts of the faculty as a whole will have a positive effect on student learning" (Brinson & Steiner, 2007, p. 1). According to Hattie (Visible Learning, 2018) self-efficacy was found to have an effect size of 0.92 indicating a high likelihood of impact on student achievement, and collective teacher efficacy was found to have an effect size of 1.57, the strongest correlation with student achievement of all influences. When asked about influences on student learning, participants discussed working collaboratively to align expectations with student needs, using reflection to improve outcomes, and a confidence that their efforts make a difference in how much students learn. Teacher interview responses centered on professional accountabilities, revealing collaboration and shared responsibilities for learning across stakeholders. They described working with students, other teachers within and outside their own school, supervisors, related service providers, parents of students, their own parents who were educators, community members, spouses, TPP professors, friends, and graduate school professors in their efforts to ensure student learning.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

According to Hattie (2015), about 20-25% of total learning variance is in the hands of the teacher, whose instructional knowledge and skills do make a difference. Hattie acknowledged that nearly all methods a teacher uses have a positive impact on learning, but some methods have a higher impact on outcomes than others (p. 81). In a school context, there is interdependence of teacher actions to reach the common goal of positive student outcomes. Participants in this study demonstrated an understanding of these professional characteristics of teaching. Participant experiences coincide with the findings that collective teacher efficacy has a strong connection with student achievement, not the individual teacher, the effectiveness of their strategies, nor the quality of their initial teacher training program. Teachers with strong perceptions of efficacy put more effort into planning, seeking new ideas, and persevering (Jerald, 2007), and collective efficacy has been connected to positive consequences on student performance, parent relationships, teacher commitment, as well as negating adverse effects of low socioeconomic status (Brinson & Steiner, 2007).

Study Limitations

There is not a perfect model for measuring student academic achievement in relation to educator effectiveness, or the quality of training program in which they were prepared. Limitations have been identified even though case study is well situated for both investigating complex educational phenomenon and advancing research of teacher preparation related to student outcomes. Efforts were made to address bias through methods and discussions, yet bias remains an inherent issue in case study research. This inquiry was an attempt to understand the collective cases, as a whole, as well as the TPP (Baxter & Jack, 2008); as such, causation of student learning growth is again cautioned. Saturation of data for thick description is limited due to purposeful selection of the six cases. Participants were not representative of all completers, and as such, findings and applications are not necessarily generalizable for the whole preparation program or beyond; limitations stem from studying one TPP and a limited variety of grades and subject areas. The six cases in this study were elementary and secondary teachers; investigation of additional program areas has been initiated.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

Conclusion

Results from these case studies suggest that graduates are making a positive impact on the learning and development of students in their classrooms. Even so, it is not always clear whether student achievement is due to the teacher, their training or other complex, intervening factors. The study offers an invitation for further dialogue and research to better understand the relationship of program graduates' teaching, student learning, and teacher preparation program effectiveness. This study could be repeated on a larger scale to analyze additional cases as well as include graduates from additional program areas, and those who may be employed in substantially different contexts from where they were prepared (e.g., inner-city or international locations). It would also be informative to further study perceptions of graduates and their supervisors in judging self-efficacy and collective teacher efficacy, perhaps in combination with student growth data or work samples. The process for evaluating the new teacher workforce, and the programs which prepare them, would be well served by additional case study and continued conversations amongst preparation programs about the impact their graduates are making. But foremost, teacher preparation programs must remain committed to ensuring every learner has a competent, effective, and reflective educator.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

References

- American Psychological Association [APA]. (2014). *Assessing and evaluating teacher preparation programs: APA task force report*. Washington, DC: Author.
- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L, ...& Shepard, L. A. (2010). Problems with the use of student test scores to evaluate teachers. EPI Briefing Paper #278. Economic Policy Institute.
- Bandura, A. (1989). Social cognitive theory. In R. Vasta (Ed.), *Encyclopedia of human behavior* (vol. 4, pp. 71-81). New York, Academic Press.
- Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report*, 13(4), 544-559.
- Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wycoff, J. (2008). *Teacher preparation and student achievement: Working paper 14314*. Cambridge, MA: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w14314>
- Braun, V., & Clark, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77-101.
- Brinson, D., & Steiner, L. (2007). *Building collective efficacy: How leaders inspire teachers to achieve*. Issue Brief. Washington, DC: The Center for Comprehensive School Reform and Improvement.
- Chung, R. R. (2008). Beyond assessment: Performance assessments in teacher education. *Teacher Education Quarterly*, 35(1), 7-28.
- Chung Wei, R., & Pecheone, R. L. (2010). Assessment for learning in preservice teacher education: Performance based assessments. In M.M. Kennedy (Ed.), *Teacher assessment and the quest for teacher quality: A handbook* (69-132).
- Cochran-Smith, M., & Zeichner, K. (2005). *Studying teacher education: The report of the AERA panel on research and teacher education* (Eds.). Washington, DC: American Educational Research Association.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

- Council of Chief State School Officers [CCSSO]. (2013). *InTASC model core teaching standards and learning progressions for teachers 1.0: A resource for ongoing teacher development*. Washington, DC: Author.
- Council for the Accreditation of Educator Preparation [CAEP]. (2015). *Questions and answers on standard 4, program impact*. Washington, DC: Author.
- Crandall, B., Klein, G., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. Cambridge, MA: Massachusetts Institute of Technology.
- Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five traditions* (2nd ed.). Thousand Oakes, CA: Sage.
- David, J. L. (2010). Using value-added measures to evaluate teachers. *Educational Leadership*, 67(8), 81-82.
- Dewey, J. (1938). *Experience and education*. New York: Touchstone.
- Digamon, J., & Cinches, M. F. C. (2017). Schlechty's student engagement continuum in the work team experience: A pilot study. *Journal of Institutional Research South East Asia* 15(3), 5-18.
- Everson, K. C., Feinauer, E., & Sudweeks, R. (2013). Rethinking teacher evaluation: A conversation about statistical inferences and value-added models. *Harvard Educational Review*, 83(2), 349-370.
- Frey, T. J., Andres, D. K., McKeeman, L. A., & Lane, J. J. (2012). Collaboration by design: Integrating core pedagogical content and special education method courses in a pre-service secondary education program. *The Teacher Educator*, 47(1), 45-66.
- Gansle, K. A., Noell, G. H., & Burns, J. M. (2012). Do student achievement outcomes differ across teacher preparation programs? An analysis of teacher education in Louisiana. *Journal of Teacher Education*, 63(5), 304-317.
- Glanville, J. L., & Wildhagen, T. (2007). The measurement of school engagement: Assessing dimensionality and measurement invariance across race and ethnicity. *Educational and Psychological Measurement*, 67, 1019-1041.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

- Glaser, B. G. & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine Pub. Co.
- Goldhaber, D., Liddle, S. & Theobald, R. (2012). *The gateway to the profession: Assessing teacher preparation programs based on student achievement: CEDR working paper 2012-4*. University of Washington, Seattle, WA.
- Hattie, J. (2015). Teacher-ready research review: The applicability of visible learning to higher education. *Scholarship of Teaching and Learning in Psychology, 1*(1), 79-91.
- Hill, C. E., Thompson, B. J., & William, E. N. (1997). A guide to conducting consensual qualitative research. *The Counseling Psychologist, 25*, 517-571.
- Humphrey, D., Wechsler, M., & Hough, H. (2008). Characteristics of effective alternative programs. *Teachers College Record, 110*(1), 1–63.
- Jerald, C. D. (2007). *Believing and achieving*. Issue Brief. Washington, DC: The Center for Comprehensive School Reform and Improvement.
- Kennedy, M. (2016). Parsing the practice of teaching. *Journal of Teacher Education, 67*(1), 6-17.
- Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health, 74*, 262-273.
- Marzano, R. J., Frontier, T., & Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching*. Alexandria, VA: ASCD.
- Maxwell, J. A. (2005). *Qualitative research design: An interactive approach* (2nd ed.). Thousand Oaks, CA: Sage.
- Merriam, S. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.
- Noell, G. H., Brownell, M. T, Buzick, H. M., & Jones, N. D. (2014). Using educator effectiveness measures to improve educator preparation programs and student outcomes (Document No. LS-1). Retrieved from: <http://cedar.education.ufl.edu/tools/literature-syntheses/>

CASE STUDY EXAMINATION OF COMPLETER IMPACT

North Dakota Department of Public Instruction [ND DPI]. (2015). *North Dakota teacher evaluation guidelines*. Bismarck, ND: Author.

Northwest Evaluation Association [NWEA]. (2013). *Measures of academic progress: A comprehensive guide to the MAP K-12 computer adaptive interim assessment*. Portland, OR: Author.

Regional Educational Laboratory (REL) Central. (2014). *Research that relates teacher preparation to teacher and student outcomes: A literature summary* (Limited distribution report for the Educator Effectiveness Research Alliance). Retrieved from relcentral@marzanoresearch.com

Schlechty, P. C. (2002). *Working on the work: An action plan for teachers, principals, and superintendents*. San Francisco, CA: Jossey-Bass.

Schlechty, P. C. (2011). *Engaging students: The next level of working on the work*. San Francisco, CA: John Wiley & Sons.

Shepard, L. A. (2012). *Evaluating the use of tests to measure teacher effectiveness: Validation as a theory-of-action framework*, National Council on Measurement in Education, Vancouver, British Columbia, April 14, 2012. Author.

Tatto, M. T., Richmond, G., & Andrews, D. J. (2016). The research we need in teacher education. *Journal of Teacher Education*, 67(4), 247-250.

Thum, Y. M., & Hauser, C. H. (2015). *NWEA 2015 MAP norms for student and school achievement status and growth*. NWEA Research Report. Portland, OR: NWEA.

Toch, T., & Rothman, R. (January, 2008). *Rush to judgment: Teacher evaluation in public education*. Washington, DC: Education Sector.

Tschannen-Moran, M., & Johnson, D. (2011). Exploring literacy teachers self-efficacy beliefs: Potential sources at play. *Teaching and Teacher Education: An International Journal of Research and Studies*, 27(4), 751-761.

Visible Learning. (2018). *250+ influences on student achievement* [PDF file]. Author. Retrieved from: https://us.corwin.com/sites/default/files/250_influences_10.1.2018.pdf

CASE STUDY EXAMINATION OF COMPLETER IMPACT

Yazan, B. (2015). Three approaches to case study methods in education: Yin, Merriam, and Stake. *The Qualitative Research Report*, 20(2), 134-152.

Yin, R. K. (2014). *Case study research: Design and method* (5th ed.). Thousand Oaks, CA: Sage.

Running head: CASE STUDY EXAMINATION OF COMPLETER IMPACT

Table 1
Participants and Context

	Program	Experience	Continuing Education	School Context
Jamie	Major: Elementary Education Minor: Special Needs	2 years-1st year 3rd grade; 1 year 4th grade	5 Continuing Ed. Credits	New Teacher Mentoring: None Total enrollment: 230 Gender: Male (126); Female (104) Ethnicity: Black (3), Asian/Pacific Islander (0), Hispanic (0), American Indian or Alaska Native (2), White (39), No Response (1) Special Programs: EL (0), Free/Reduced Meals (72), Special Education (31)
Terry	Major: Elementary Education Minor: Science Endorsement: Middle School	2 years-both 4th grade	21 Graduate Credits Master's program-STEM	New Teacher Mentoring: Available Total enrollment: 257 Gender: Male (142); Female (115) Ethnicity: Black (4), Asian/Pacific Islander (1), Hispanic (0), American Indian or Alaska Native (2), White (241), No Response (9) Special Programs: EL (9), Free/Reduced Meals (32), Special Education (26)
Reese	Major: Elementary Education Minor: Early Childhood Education	2 years-both 4th grade	District Professional Development Days	New Teacher Mentoring: Available Total enrollment: 307 Ethnicity: Black (1%), Asian/Pacific Islander (0%), Hispanic (1%), Native American (1%), White (96%) Special Programs: EL (N/A), Low Income (11%), IEP (8%)
Stacy	Major: Elementary & Early Childhood Education-Double major; Psychology	3 years-2 years at 2 nd grade; currently 5th grade	2 Continuing Ed. Credits	New Teacher Mentoring: Available Total enrollment: 88 Ethnicity: Black (0%), Asian/Pacific Islander (4%), Hispanic (12%), Native American (4%), White (85%) Special Programs: EL (N/A), Low Income (29%), IEP (14%)
Taylor	Major: History & English Education Double Major	3 years-8 th grade English	27 Graduate Credits Master's program-English Education	New Teacher Mentoring: Available Total enrollment: 517 Ethnicity: Black (1%), Asian/Pacific Islander (1%), Hispanic (1%), Native American (35%), White (61%) Special Programs: EL (N/A), Low Income (15%), IEP (N/A)
Alex	Major: Composite Biology Education Endorsement: Middle School	2 years-grades 9-12 science	3 Graduate Credits 2 Continuing Ed. Credits Master's program-Education	New Teacher Mentoring: Available Total enrollment: 255 Ethnicity: Black (1%), Asian/Pacific Islander (1%), Hispanic (32%), Native American (2%), White (64%) Special Programs: EL (N/A), Low Income (42%), IEP (18%)

Note. Information compiled from the state school database, TPP graduate survey, document review and completer interview protocol.

CASE STUDY EXAMINATION OF COMPLETER IMPACT

Table 2

Cross-Case Findings

	Students At or Above Grade Level Norms	Pre-Post Assessment Demonstrated Growth	Highest Two Levels of Student Engagement	Interview: Impact Factor Themes	
	%	%	%		
Jamie	Reading	57.1	Reading: 100	57.0	
	Math	14.0			
	ELA	64.3			
Terry	Reading	58.0	N/A*	83.5	
	Math	42.0			
	ELA	58.0			
	Science	68.0			
Reese	Reading	80.0	Math: 100	87.5	Collaboration
	Math	93.7			
Stacy	Reading	50.0	Math: 89	72.8	Self-efficacy
	Math	75.0			
	ELA	91.7			
	Science	83.3			
Taylor	Reading	55.2	English Paper: 100	87.6	
	ELA	55.2			
Alex	Science	74.0	Ecology Test: 100	82.7	

Note. Terry did not submit pre-post assessment data