

Leveraging Case Study Research: A Mechanism to Measure Teaching Effectiveness

Sarah K. Anderson
Mayville State University

Brittany D. Hagen
Mayville State University

Andrea L. Dulski-Bucholz
Mayville State University

Christopher M. Whitsel
North Dakota State University

Kayla D. Smith
Mayville State University

Ann M. Willeson
Mayville State University

This article shares findings from a descriptive, multiple case study assessing graduate outcomes 2.5 years after the completion of a teacher education program. Case study was used as an alternative to value-added measures to holistically examine complex attributes of effective teaching. Mixed methods data collection included graduate and supervisor surveys, self and supervisor evaluation of skills and dispositions, interviews, and classroom observations. Results indicated participants effectively applied the knowledge, skills, and dispositions the program was designed to achieve in the areas of Learner and Learning, Content, Instructional Practice, and Professional Responsibility. Three major recommendations emerged: 1) the necessity to develop established proficiency levels for new teachers related to diverse learners, 2) the need for shared responsibility of outcomes and targeted induction support, and 3) support for supervisor evaluation as a viable mechanism for education program accountability of teaching effectiveness of graduates.

Introduction

Teaching is one of the most complex human activities (Ball & Forzani, 2011). Due to its fluid and relational nature, researchers recognize that evaluating teaching effectiveness presents multiple challenges (Chung Wei & Pecheone, 2010; Noell, Brownell, Buzick, & Jones, 2014). Among these challenges is increasing accountability (Tatto, Richmond, & Andrews, 2016), requiring educator preparation programs (EPPs) to measure and report graduates' knowledge, skills, and dispositions as well as the impact these elements have on P-12 student learning (Council for the Accreditation of Educator Preparation (CAEP), 2013). These three facets of effective teaching are delineated in order to examine the complexity of teaching quality; yet

knowledge, skills, and dispositions are intricately interdependent and cannot be truly evaluated in isolation.

Several options for EPPs to measure effective teaching of graduates from their programs have been suggested (CAEP, 2015). The first, supported by both federal regulatory and accrediting agencies, is through value-added measures which seek to determine what educational achievement a teacher has added to student learning by measuring the difference in student test scores over time (Tatto et al., 2016). The research base presents conflicting justifications in using these measures to identify quality teaching. Critics suggest value-added measures are not always a strong indicator of teaching effectiveness and should not be used in isolation from other measures of knowledge, skills, and dispositions (Baker et al., 2010; Briggs & Domingue, 2011; Everson, Feinauer, & Sudweeks, 2013; Toch & Rothman, 2008) or “erroneous conclusions about teachers” could be made (Marzano, Frontier, & Livingston, 2011, p. 103). Further, value-added measures are not readily available to all EPPs, as many state systems are not currently in place to connect student test scores to individual teachers (CAEP, 2015).

In a study that estimated effects of preparation program features on teachers' value-added measures to student test-score performance, Boyd, Grossman, Lankford, Loeb, and Wycoff (2008) found indicators that pre-service preparation can influence the effectiveness of teachers, particularly those in their first year. These researchers surveyed perceptions of all first-year teachers in New York City representing graduates from over 30 teacher preparation programs. Results linked the quantity of practice teaching during preparation to directly benefiting teacher effectiveness. However, uncertainty remained concerning the extent to which value-added measures of student achievement are actually valid measures of student learning or of teachers' impact on learning. Cochran-Smith and Zeichner (2005) argued for stakeholder caution when using value-added data for program evaluation due to the complex variables involved in linking outcome measures to preparation. Acknowledging these issues, EPPs are encouraged to build evidence around their choice of measures and make arguments for those decisions to demonstrate that accreditation standards are met (CAEP, 2015). As an alternative approach, case study methodology has been suggested as an option to gather evidence of graduate outcomes (CAEP, 2016). Case study research has been used extensively to study teaching (Merriam, 1998), largely focusing on individual P-12 students to address specific issues, as well as to identify and explain problems of practice and school culture, historical cases, and sociological cases focused on constructs of educational phenomena.

In 2014, the Regional Educational Laboratory (REL) Central conducted a review of 1,891 publications that associated educator preparation with both teacher and student outcomes (REL Central, 2014). The literature summary report indicated 56 studies where research connected teacher preparation to any type of teacher or student outcomes, and only three employed a case study design. Since case study has been suggested as an option for reliable EPP evidence of program quality, these three studies were of particular interest. The first study included performance assessment of two pre-service elementary teachers through surveys and focus groups (Chung, 2008). The second involved 16 pre-service teachers' use of universal design as assessed through lesson plans, unit assessments, and reflection (Frey, Andres, McKeeman & Lane, 2012), and the final study examined characteristics of seven alternative certification programs appraised as effective (Humphrey, Wechsler & Hough, 2008). The review did not

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.

This case study describes the knowledge, skills, and dispositions that impact P-12 student learning, and the development of two teachers, 2.5 years after completing their pre-service training, as evidence of teacher education program quality. An additional aim for this study was the creation of a replicable protocol available for broader use across EPPs who similarly chose alternative program impact measures. The case study was part of a larger investigation of the role EPPs play in graduates' effectiveness and the function this effectiveness has on student learning, among a myriad of other complex factors (Visible Learning Plus, 2018). Data was collected to answer the question, "What knowledge, skills, and dispositions do EPP completers demonstrate within one to three years of graduation?" This study was designed to describe if graduates have a proficient level of teaching effectiveness to positively impact student learning, as well as to drive continuous improvement efforts for preparation programs.

Reflective Experiential Framework

The *Reflective Experiential Teacher* framework¹ was used to design methods and guide analysis. It encourages learning through reflection and contemplation of beliefs and experiences as knowledge (Dewey, 1938). These experiences occur through observation of learning as cooperating teachers model effective pedagogy, as well as practice teaching paired with targeted feedback, the essence of social cognitive theory (Bandura, 1989).

This constructivist framework incorporates the knowledge, skills, and dispositions of the ten core teaching standards of the Interstate Teacher Assessment and Support Consortium (InTASC) in the form of four general categories: Learner and Learning, Content, Instructional Practice, and Professional Responsibility (Council of Chief State School Officers (CCSSO), 2013). These categories "provide a scheme for the valid evaluation of any teacher's core competencies and the reliable means of recording and compiling overall teacher performance" (North Dakota Department of Public Instruction (ND DPI), 2015, p. 4). In addition, these categories represent the foundation upon which teaching practice and its effect on student achievement are based.

The acquisition of the competencies needed to become a professional educator requires critical thinking about experiences within social, cultural, and environmental contexts. These experiences influence teachers' prior knowledge and how they reflect on theory and evidence-based practices. Learning and growth takes place through completers' reflection on their experiences as well as the cyclical informal and formal teacher evaluation process in the workplace. This framework forms the system of concepts, assumptions, expectations, and beliefs that informed this study (Maxwell, 2005).

Methods

A descriptive, multiple case approach was utilized to describe two cases of graduates' knowledge, skills, and dispositions that their preparation experiences were designed to achieve.

¹ The framework is outlined in the handbook of the EPP from which the participants graduated, and so it was left intentionally uncited to protect anonymity.

The participants completed a teacher education program at a regional, 4-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. The cases were a part of a larger, mixed methods study about the impact of graduates on P-12 student learning and classroom instruction, which first required examination of graduate proficiency. Case study was selected because it allowed data collection from many sources to provide in-depth information and allowed for holistic examination of knowledge, skills, and dispositions as interrelated attributes (Kennedy, 2016). The design rigor of Yin (2014) along with the constructivist-education epistemology of Merriam (1998) complement each other and were used to meet the demands of a study on completer outcomes (Yazan, 2015). As a descriptive case study, there is not an attempt to infer causality of effective teaching.

Participants

Upon approval from the Institutional Review Board, participants were selected through purposeful sampling (Merriam, 1998); each demonstrated a completer perspective within a defined context and with enough information to portray an in-depth picture of two graduates from the largest program in the EPP, elementary education.

Recruitment criteria included: full-time teaching within an 80-mile radius of the EPP and program completion date from one to three years. Four graduates were contacted via email and invited to participate. Two responded and consented. Once graduates gave initial consent, the superintendent and principal of each participant’s school were contacted. Graduate participation was dependent on administrative agreement for supervisors to complete surveys and interviews; both principals agreed. Written informed consent from participants was gained prior to beginning the first online survey and was verbally reaffirmed prior to classroom observations and interviews. Descriptions of each participant, Terry and Jamie (pseudonyms), and their current teaching contexts are displayed in Tables 1 and 2.

Table 1
Participants

	Terry	Jamie
Program	Major: Elementary Education Minor: Science Endorsement: Middle School	Major: Elementary Education Minor: Special Needs
Licensure	Not submitted	2 year initial; elementary grades 1-6
Experience	2 years-both 4 th grade	2 years-1 st year 3 rd grade; currently 4 th grade
Continuing Education	21 graduate credits STEM Master’s program	5 continuing education credits

Note. Compiled from the Transition to Teaching survey (TTS), document review and completer interview protocol.

Table 2
Contexts

	Terry	Jamie
Mentoring	Available	None
Demographics 2016-2017	Total enrollment: 257 Gender: Male (142); F (115) Ethnicity: Black (4), Asian/Pacific Islander (1), Hispanic (0), American Indian or Alaska Native (2), White (241), Multiple/No Response (9) Special Programs: EL (9), Free/Reduced Meals (32), Special Education (26)	Total enrollment: 230 Gender: Male (126); F (104) Ethnicity: Black (3), Asian/Pacific Islander (0), Hispanic (0), American Indian or Alaska Native (2), White (39), Multiple/No Response (1) Special Programs: EL (0), Free/Reduced Meals (72), Special Education (31)
Context	<i>Agree:</i> Physically safe and secure place; faculty and staff have positive relationships with students' parents/guardians. <i>Tend to Agree:</i> Teachers respect dignity and worth of students.	<i>Agree:</i> Physically safe and secure place; faculty and staff have positive relationships with students' parents/guardians; teachers respect dignity and worth of students
Professional Environment	<i>Agree:</i> Teachers are continually learning and seeking new ideas to enhance practice. I receive valuable professional guidance from mentors or colleagues, and teachers have influence over establishing curriculum. The administration is responsive to needs of teachers.	<i>Tend to Agree:</i> Teachers are continually learning and seeking new ideas to enhance practice. <i>Tend to Disagree-</i> The administration is responsive to needs of teachers. <i>Disagree:</i> I receive valuable professional guidance from mentors or colleagues, and teachers have influence over establishing curriculum.
Resources	<i>Agree:</i> Teachers have time for planning with colleagues and appropriate instructional space, and teachers have curricular materials and supplies that are appropriate for students' needs and necessary technology resources.	<i>Agree:</i> Teachers have time for planning with colleagues and appropriate instructional space. <i>Tend to Agree:</i> Teachers have curricular materials and supplies that are appropriate for students' needs. <i>Tend to Disagree:</i> Teachers have the necessary technology resources.
The Profession	<i>Agree:</i> I am as happy about teaching as I thought I would be, and rewards of teaching are worth the efforts required by my preparation program.	<i>Tend to Agree:</i> I am as happy about teaching as I thought I would be, and rewards of teaching are worth the efforts required by my preparation program.

Note. Information from the Transition to Teaching Survey (TTS); school demographics from state database.

Instruments

Data was collected over two months using seven sources: 1) Transition to Teaching Survey (TTS), 2) Supervisor Survey (SS), 3) Skills of Teaching Observation Tool (STOT), 4) Disposition Evaluation, 5) classroom observation, 6) teacher and supervisor interviews, and 7) document review. This included interview questions about teaching impact, factors impacting effectiveness, and preparation.

Transition to Teaching Survey and Supervisor Survey. These surveys are part of the EPP's state-wide common metrics project. The surveys were developed using a rigorous process that included multiple psychometric analyses, focus groups, pilot testing, revision, and alignment with accreditation standards by the Network for Excellence in Teaching (NExT Consortium, 2016). The 46-item TTS and 45-item SS are aligned to the InTASC Standards, therefore items and sections can be compared. The entire TTS and SS were administered however, only relevant portions were utilized in this study. Because the NExT Consortium holds survey copyright, alterations were not permitted. Responses established school context and participant demographics and queried participants to rate how well prepared they felt across domains of teaching including: Diverse Learners, Learning Environment, Instructional Practice and Professionalism. The SS asked supervisors to assess the quality of graduates' abilities in these domains.

Skills of Teaching Observation Tool (STOT). The STOT evaluation included 34 indicators of performance in the areas of Learner and Learning, Content, Instructional Strategies and Professional Responsibility. Actionable descriptors indicated levels of proficiency. The STOT (available at "ndacte.org") was completed by each participant and their supervisor to evaluate performance of observed teaching skills in their own classrooms.

Disposition Evaluation. The 19-item Disposition Evaluation was completed by participants and supervisors. The evaluation was developed by EPP faculty (from the participants' university) to measure values, commitments, and ethics influencing behaviors toward students, families, colleagues, and communities (CCSSO, 2013). Construct validity was ensured through alignment with InTASC standards. An underlying assumption is that dispositions frame the decisions teachers make in classroom interactions.

Classroom observation. The fieldwork researcher, an educational sociologist contracted from outside the EPP, conducted observations of participants' knowledge, skills, and dispositions in their own classrooms. Observation represents a "firsthand encounter with the phenomenon of interest rather than a secondhand account of the world obtained in an interview" (Merriam, 1998, p. 94) and can substantiate findings. Each participant was observed one time for approximately 90 minutes. Copious field notes were recorded: one note column for contextual factors, the other for interactions amongst participant and students. Notes included classroom descriptions, items written on the board, materials used, teacher statements, and student actions. Handwritten notes were typed soon after observation and later were coded to identify themes.

Teacher and supervisor interviews. Interview questions were drafted and revised by the research team (See Appendix A). Similar to the TTS and SS, questions were generated around

pre-existing codes of Learner and Learning, Content, Instructional Practice, and Professional Responsibility. Questions were included for inquiry about student learning and program impact.

Document review. Document review was included to corroborate and augment other sources of evidence (Yin, 2014). Participants were asked to submit digital copies of the following documents: two most recent district-level supervisor evaluations, teaching license, and current transcripts.

Data Collection

Data collection was guided by the study protocol timeline. Interactions with participants were conducted by the data manager and fieldwork researcher, neither of whom taught the participants or conducted analysis. Data collection followed a pattern of initial request, reminder email if not completed within one week, and a phone call within one additional week if needed. As sources were completed, results were coded as Participant 1 or Participant 2 by the data manager and provided to the researchers conducting analysis.

First, the TTS and SS were sent simultaneously via email with instructions and password protected survey links. Upon completion of surveys, links to the disposition evaluation and STOT were distributed to participants and supervisors; these were completed online in the EPP's quality assurance system. Participants were provided a unique login and password; the institutional account was password protected. After completion of quantitative survey data, the fieldwork researcher arranged classroom observations where field notes were collected and submitted online to the data manager. Phone interviews with participants and supervisors were also completed by the fieldwork researcher. Interview protocols were provided in advance (see Appendix A) and interviews occurred after observations to not influence observable practices. Participants were requested not to discuss research processes until the study concluded. In the final step of data collection, participants were requested to submit scanned copies of documents. Participants were notified by email when data collection concluded, with information about the completion timeline and member checking.

Data Analysis

Analysis of quantitative and qualitative data was conducted using similar processes. Two researchers independently analyzed data using structured coding forms to ensure consistency and organized data for the constant-comparative process. Descriptive statistics were used to describe data for each case. Within each case, the constant comparative method of qualitative data analysis was used (Glaser & Strauss, 1967) to construct codes, categories, subcategories, or themes through continuous comparison of data (Merriam, 1998). Guidelines of thematic analysis were used to ensure reliability (Braun & Clark, 2006). Classification began via pre-existing codes aligned with instrumentation: Learner and Learning, Content, Instructional Practice, and Professional Responsibility. Through constant comparison and reconceptualization, codes were confirmed using categorization (Creswell, 2007; Yin, 2014). Next, a search for patterns among these categories was used to identify themes. A consensus on overarching themes was reached, and single-case results were sent to participants for member checking. Data was re-analyzed

across cases using the framework of Morse (1994): comprehending, synthesizing, theorizing and recontextualizing.

Trustworthiness

Research adhered to a case study protocol developed from best-practices in educational research (Merriam, 1998; Yazan, 2015; Yin, 2014). Additionally, the theory of constructivism through the EPP's *Reflective Experiential* model was used as a broad foundation for design decisions. Triangulation captured dimensions of data and cross-validated findings through multiple sources of quantitative and qualitative evidence. Replication logic was used across both cases (Yin, 2015). To reduce the risk of bias, a database was maintained by a data manager who did not conduct analysis, and who also coded data prior to analysis to maintain anonymity. Furthermore, interviews and observations were conducted by a researcher who was not employed by the EPP and who had not participated in preparing the participants for teaching. Finally, member checking occurred for respondent validation of results.

Results

Multiple sources of data for measuring participants' effective knowledge, skills, and dispositions of teaching are reported in the results. Convergence of data is an attempt to describe participants' teaching practice through a holistic lens and integrate sources to answer the research question: "What knowledge, skills, and dispositions do EPP completers demonstrate within one to three years of graduation?"

Transition to Teaching Survey and Supervisor Survey

Participants completed the TTS, and the SS was completed by the participants' respective supervisors. Participants answered the following question, "To what extent do you agree or disagree that your teacher preparation program prepared you to do the following?" When completing the SS, supervisors answered, "To what extent do you agree or disagree that this teacher does the following?" Results for Terry, Jamie, and their supervisors are displayed in Table 3. Both participants and their supervisors rated overall agreement that participants exhibit the knowledge, skills, and dispositions of effective teaching with individual patterns of strengths and weaknesses.

Table 3
Transition to Teaching Survey (TTS) and Supervisor Survey (SS) Results

	Rating	Terry		Terry's Supervisor		Jamie		Jamie's Supervisor	
		<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Instructional Practice 21 Items	A	14	67	7	33	7	33	13	62
	TA	7	33	14	67	11	53	8	14
	TD	0	0	0	0	3	14	0	0
	D	0	0	0	0	0	0	0	0
Diverse Learners 9 Items	A	0	0	0	0	0	0	1	20
	TA	5	56	9	100	4	44	4	80
	TD	4	44	0	0	5	56	0	0
	D	0	0	0	0	0	0	0	0
Learning Environment 9 Items	A	9	100	3	33	3	33	9	100
	TA	0	0	5	56	3	33	0	0
	TD	0	0	1	11	3	33	0	0
	D	0	0	0	0	0	0	0	0
Professionalism 7 Items (TTS) 6 Items (SS)	A	4	57	0	0	0	0	6	100
	TA	2	29	4	67	6	86	0	0
	TD	1	14	2	33	1	14	0	0
	D	0	0	0	0	0	0	0	0
Total 46 Items (TTS) 45 Items (SS)	A	27	59	10	22	10	22	29	71
	TA	14	30	32	71	24	53	12	29
	TD	5	11	3	7	12	25	0	0
	D	0	0	0	0	0	0	0	0

Note: A = Agree, TA = Tend to Agree, TD = Tend to Disagree and D = Disagree. Participant rated preparation; supervisor rated current performance. Jamie's Supervisor marked four items as "not able to observe" in the area of Diverse Learners. The four categories of the surveys are not directly aligned with the four categories of effective teaching due to copyrighted content.

Terry's survey results. Terry rated general agreement with teaching preparation. The category of highest agreement was Learning Environment; Instructional Practice was second highest, and third was Professionalism; the category with the highest rate of disagreement was Diverse Learners. Starting with the highest rating, the supervisor indicated Learning Environment, Instructional Practice, Diverse Learners, and then Professionalism. The supervisor *disagreed* with one individual item, that Terry "helps students regulate their own behavior"; Terry marked *agree* on this item.

Terry and the supervisor rated all Instructional Practice items with *agree* or *tend to agree*. Terry typically reported higher rates of agreement for Instructional Practice than the supervisor. In the area of Professionalism, there was a marked difference between Terry's rating and the supervisor's response. The supervisor noted *tend to disagree* on items such as "collaborating with colleagues" and "using feedback from colleagues to support professional development."

There were no common ratings between the supervisor and Terry in this area. Terry rated being prepared to work with Diverse Learners with the highest rate of disagreement of all sections. The self-rating was lower than the supervisor rating; the supervisor rated 100% of items *tend to agree*. Terry disagreed with several statements related to differentiating instruction; the supervisor ranked these same statements as *tend to agree*. Terry and the supervisor marked the same rating for 33.3% of items.

Jamie's survey results. Jamie indicated highest agreement in the Learning Environment and Instructional Practice categories, followed by Professionalism. The category with the highest rate of disagreement was Diverse Learners. The supervisor indicated *agree* or *tend to agree* on all items. In order of highest rating, the supervisor indicated Learning Environment, Professionalism, Instructional Practice, and lastly Diverse Learners.

In the category of Learning Environment, disagreement was indicated by Jamie with preparation in the areas of: “effective communication skills and strategies,” “helping students regulate their own behavior,” and “organizing the physical environment for instruction.” The supervisor rated *agree* with these same statements. In the area of Instructional Practice, ratings between Jamie and the supervisor were consistently *agree* and *tend to agree*, other than Jamie's ranking in “ability to analyze appropriate assessment types,” “use digital and interactive technologies,” and “engaging students in the use of interactive technologies” where *tend to disagree* was indicated.

Jamie did not mark *agree* on any items related to Professional Responsibility. Six were marked *tend to agree*, with *tend to disagree* on one item: “seek out learning opportunities that align with my professional development goals.” The supervisor provided a more positive rating, marking *agree* with all items. Jamie self-rated preparation for working with Diverse Learners with more *tend to disagree* than the other areas but did not mark any items as *disagree*. The supervisor, by contrast, rated only 20% of the items as *agree* and marked *not able to observe* on four items in this category (40%). Jamie and the supervisor marked the same rating for 31.1% of items.

Dispositions and STOT

On these measures, Terry and Jamie each indicated that they demonstrated the professional dispositions and skills of effective teaching, and their supervisors concurred (see Table 4 and 5). For both cases, dispositions were rated as more proficient than teaching skills.

Terry's Disposition and STOT: Terry's overall Disposition rating was 3.3 out of 4.0; the supervisor rated 3.05, slightly lower than Terry. Terry's lowest self-rated disposition was “organization.” One item, “timeliness,” received an unsatisfactory rating by the supervisor. The supervisor and participant were in agreement on 31.5% of items.

On the STOT rating for teaching skills, Terry self-rated every item as *partially proficient*. One item had corresponding ratings for Terry and the supervisor, “collaboratively designs instruction.” This item ranked lowest on the supervisor STOT. Of the supervisor's rating, 97% fell in *proficient* or *distinguished* categories, whereas no self-ratings for Terry were higher than *partially proficient*.

Jamie's Disposition and STOT. Jamie self-rated all professional dispositions as either *distinguished* or *proficient* as did Jamie's supervisor. Jamie self-rated a total Disposition score of 3.5; the supervisor rated an almost equal score of 3.42. The supervisor and Jamie were in agreement on 36.8% of items.

Jamie's results on the STOT evaluation indicated a self-rating of 3.01. The highest rated item was "upholds legal responsibilities as a professional educator." Jamie's supervisor gave a slightly lower rating overall. The highest rated item by the supervisor was "collaborates with parent/guardian/advocate to improve student performance," which was the lowest rated item by Jamie. The supervisor rated *emerging* on six items: (1) guide learners in using technology appropriately, safely, and effectively, (2) guides mastery of content through meaningful learning experiences, (3) integrates culturally relevant content, (4) accesses resources to build global awareness, (5) engages students in self-assessment strategies, and (6) uses technology appropriately to enhance instruction. Agreement between the supervisor and Jamie occurred on 41.2% of items.

Table 4
STOT Results: Teaching Skills (1-4 scale)

	Terry	Terry's Supervisor	Jamie	Jamie's Supervisor
Learners & Learning	2.50	3.39	2.94	2.72
Content	2.50	3.43	3.07	2.28
Instructional Practice	2.50	3.38	2.92	2.79
Professionalism	2.50	3.00	3.25	3.75
STOT Rating	2.50	3.32	3.01	2.75

Table 5
Dispositions Results (1-4 scale)

	Terry	Terry's Supervisor	Jamie	Jamie's Supervisor
Learners & Learning				
Teacher awareness	4	3	4	3
Sensitivity to diversity	3	4	3	3
Rapport	3	2.5	3	4
Attitude toward learners	4	4	4	4
Total	3.50	3.38	3.50	3.50
Instructional Practice				
Organization	2	3	3	3
Flexibility	3	3	4	4
Assessment	3	4	3	3
Total	2.67	3.33	3.33	3.33
Professionalism				
Timeliness	3	1	3	4
Attendance	4	3	4	4
Dress and appearance	3	3	4	3
Attitude and composure	3	2.5	3	3
Initiative	3	3	3	3
Ethics and confidentiality	3	2.5	3	4
Communication	3	3	4	3
Cooperation/collaboration	3	2.5	4	4
Self-reflective	4	3	3	3
Responsiveness to feedback	4	3	4	4
Lifelong learner	4	4	4	3
Ways to contribute	4	4	3	3
Total	3.42	2.88	3.50	3.42
Overall Disposition Rating	3.30	3.05	3.47	3.42

Note. The disposition evaluation did not contain any items in the area of Content.

Document Review

Both participants submitted their teaching licenses, current transcripts, and supervisor evaluations for review. License and transcript information was utilized to describe the context for teaching and learning in Table 1. In Figure 1, a visual re-creation of the forms for supervisor evaluation submitted by each participant demonstrate content of the thematic analysis.

Terry's review. Terry submitted one document, the *Teacher Evaluation Report*, utilized by the school district. It was based on administrative observations, feedback, conferencing and professional interactions.

Terry's Teacher Evaluation Report Form	Jamie's Goals, Observation and Anecdotal Notes
<p>The following Teacher Evaluation Report is based on administrative observations, feedback to the teacher, conferences, and related professional interactions. It is understood that these observations and interactions constitute the formative/supervisory process and this report is the summative process.</p> <p>I. Observations: Date _____ Time _____ Length _____ Lesson _____ Follow-Up _____</p> <p>II. Evaluation of Current Level of Teacher Effectiveness:</p> <ol style="list-style-type: none"> 1.) Clear Learning Goals. Domain 1-Element 1 2.) Classroom Rules and Procedures. Domain 1-Element 4 3.) Chunking Content into "Digestible Bites". Domain 1-Element 9 4.) Demonstration of With-it-ness". Domain 1-Element 33 5.) Celebrating Success. Domain 1-Element 3 6.) Organizing the Physical Layout of the Classroom. Domain 1-Element 5 <p>III. Recommended Area for Future Growth:</p> <p>IV. Administrator's Comments:</p> <p>V. Evaluation: Based on the supervisory process, the performance of _____ is considered:</p> <p>_____ Proficient _____ Partially Proficient _____ Developing (A detailed growth plan is attached. Continuation of contract is dependent upon successful completion of this growth plan) _____ Novice (Continuation of contract is not recommended.)</p> <p>VI. Comments by the Teacher: (Optional)</p> <p>_____ Principal _____ Teacher _____</p> <p>The signature indicates the teachers has read the report. It does not necessarily indicate concurrence.</p> <p>_____ Date _____ Date _____</p>	<p>Personal Goals:</p> <ol style="list-style-type: none"> 1. Learning Goals 2. Scales 3. Student Engagement 4. Routines and Expectations 5. Pick and Element. Element 9: Chunking Content into "Digestible Bites" 6. AdvancEd Goals for Improvement <hr/> <p>Classroom Observation Teacher _____ Evaluator _____ Date: _____ Time: _____ Subject: _____</p> <p>Strategies:</p> <ul style="list-style-type: none"> Learning Goals Feedback Relevant Lesson Tracking Progress Planning Evident Technology Students Engaged Transitions Teacher Movement Positive Tone Rules/Procedures Layout/Organization Celebrating Success Areas of Strength Areas for Improvement: <hr/> <p>I visited _____'s classroom on _____ at _____ time.</p> <p>Typed Notes</p> <p>_____ Principal _____ Teacher _____ date _____</p> <p>Teacher comments:</p>

Figure 1. Submitted Supervisor Reports and Forms

Evidence of effective teaching was documented qualitatively by the supervisor in designated domains of the Marzano Evaluation System (Marzano et al., 2011):

- 1) Clear learning goals (e.g., I can statements with a scale for level of performance, activities match learning goals)
- 2) Classroom rules/procedures (e.g., establishes and reviews expectations for rules and procedures, monitors students, uses "whole brain" methods)
- 3) Chunking content into "digestible bites," (i.e., breaks input experiences into small chunks based on student needs and monitors appropriateness)
- 4) Demonstrating "withitness," (that is awareness of students and managing issues); "is aware of when students are getting off task and what rule they need to be reminded of... has become a more confident teacher this year... is able to manage issues faster so that they don't get out of hand,"
- 5) Celebrating success relative to learning goals, (e.g., gains relative to learning goals, it seems that students are always celebrating something, and students want to make progress)
- 6) Organizing the physical layout of the classroom (e.g., facilitates movement, use of centers, alternative seating and lighting).

According to the supervisor's evaluation, Terry is considered a *Proficient* teacher; proficient was the highest level of performance on the evaluation.

Jamie's review. Jamie submitted four documents for review: personal goals, one classroom observation form, and two typed reports of notes from supervisor classroom visits. Personal goals were established in reference to the Marzano's Teacher Evaluation Model (2011) and to school accreditation.

The supervisor observation included 16 observation areas (see Figure 1). The evaluation contained qualitative supervisor statements such as "yes," "good," "N/A," "ok," and "none." Seven of these statements were followed up with a comment. For example, in the area of teacher movement, the observation stated "yes-went from group to group." Areas of strength were noted as rapport with students, patience, improvement in classroom control, and taking advice. Areas for improvement were to relocate the "I Can" statement location and to use a timer on the interactive whiteboard as a visual reminder.

In the second submission of Jamie's anecdotal classroom notes, the supervisor descriptively recorded what was occurring in the classroom. Reference of Jamie's established goals was included in the notes. In the area of learning goals, "[Jamie] had the "I Can" statement for the lesson on the board. [Jamie] read the "I Can" statement to the children then had them repeat it to [Jamie]. Student engagement was also noted, "[Jamie] started the lesson by using a clap-clap-clap-strategy... The students were involved in the lesson. They were engaging each other in the task of finding right ingredients to put in the mixture and then measuring the ingredient to add to the bowl." Routines and expectations were also documented by the supervisor, "The children know the expectations of the teacher and they understand the rules."

There was no conclusive rating in the submitted supervisor evaluation to indicate level of performance. General qualitative comments indicated the supervisor was satisfied with Jamie's teaching effectiveness: "Your class control has grown tremendously since year one. Good job-I appreciate how you take advice," "there is a positive feeling or tone in the room," "many students with unique needs both academically and behaviorally. I commend [Jamie] for doing a lesson like this based on that assessment," "the layout and organization of the room is good."

Classroom Observation Field Notes

To confirm and substantiate results, classroom observations occurred with Terry and Jamie. Frequency of pattern coding using the categories of Learner and Learning, Content, Instructional Practice, and Professional Responsibility are displayed in Table 6.

Terry's classroom observation. Terry's classroom observation occurred during reading and math instruction. During the observation, Terry demonstrated knowledge, skills, and dispositions in all four areas. Examples to illustrate observations are found in Table 6. The observation began with students lining up for leveled reading groups; some were going to different classrooms and some joined the remaining students from Terry's class. Students formed groups on the playground, and Terry checked in, monitoring groups as they worked. Students worked on an activity summarizing likes/dislikes about books read as a class. Upon returning to the classroom,

students added their personal likes and dislikes on poster size papers distributed around the room. Terry worked with individual students during this time and managed the groups. Next, students transitioned to math work stations. Some students gathered at the front of the room with individual whiteboards, where Terry instructed a small group about converting fractions. Some gathered at a computer in the back of the room, and other students received tablets for practice with individualized worksheets; students rotated through stations.

Table 6
Observation Field Notes: Frequency for Pre-Existing Codes

Terry							
Learner & Learning		Content		Instructional Practice		Professional Responsibility	
Transitions	9	Instructional strategies	1	Instructional strategies	5	Collaboration	1
Positive learning environment	6	Supplementary resources	1	Assessment	4	Professional development	1
Management	5	Content knowledge	1	Instructional choice	4		
Learners' needs	4			Technology	3		
Expectations	2						
Relationships	1						
Engagement	1						
Total	28		3		16		2
Jamie							
Learner & Learning		Content		Instructional Practice		Professional Responsibility	
Positive learning environment	8	Content knowledge	1	Assessment	7	Collaboration	3
Transitions	6			Instructional choice	6		
Learners' needs	5			Technology	1		
Expectations	3						
Instructional strategies	2						
Management	2						
Relationships	2						
Engagement	1						
Motivation	1						
Total	30		1		14		3

Jamie’s classroom observation. Jamie’s observation also occurred during reading and math instruction. Jamie demonstrated skills in each of the four areas of effective teaching during the observation (see Table 6). According to the field notes, students arrived in the classroom after gym and had snacks. Next, students worked independently on a memory book activity. While the

students worked, Jamie played music and monitored individual students. Students transitioned to math instruction, which was done in flexible grouping; some students left the classroom, and other students entered. Jamie did a “fist-bump” with each student as they left or entered the room and gained their attention before instruction using the statement, “5, 4, 3, 2, 1, lips locked, eyes on me.” A one-minute timed math assessment occurred and Jamie distributed individualized worksheets to each student. During math time, three adults were engaged in the classroom, including a para-educator and the principal. When Jamie noticed students struggling with a common item, the students’ attention was directed to a whiteboard easel where Jamie led the whole group through a problem. Students finished worksheets at different times and moved onto math activity pages.

Teacher and Supervisor Interviews

The fieldwork researcher conducted phone interviews with the participants and their respective supervisors using the established interview protocols (see Appendix A). Responses provide insight and examples of knowledge, skills, and dispositions of effective teaching. Closed coding yielded frequency of codes according to the four areas of effective teaching.

Terry’s interviews. Frequency of codes for Terry and Terry’s supervisor are displayed in Table 7. Significant statements offer explicit examples of effective teaching documented in other data sources, including the responses of Terry’s supervisor, which are presented alongside Terry’s answers.

Learner & Learning. The results of Terry’s interview indicated agreement that student learning is supported by building relationships with students and families. Terry confirmed that a student’s “home life contributes to learning,” stating specific examples of migrant students who attend school in Mexico in the winter and need instructional materials in both English and Spanish. The supervisor detailed that Terry is always looking for ways to support students, and added that additional academic support is also provided during recess and behavior plans are used to provide parents with feedback. In regards to the learning environment, Terry stated the classroom environment is “not a typical classroom”; there is flexible seating and lighting. The supervisor confirmed that Terry “tries to find new ways of doing things.” Both agree that diligent work in the beginning of the year occurred to evaluate how students learn best, and that Terry plans instruction accordingly.

Content. Terry displayed confidence in the ability to “have a grasp on content areas,” providing specific examples of using supplementary materials other than the textbook, making cross-curricular connections, chunking content into smaller parts, and making learning relevant to students’ everyday lives. Terry’s supervisor noted a specific illustration of content knowledge application and appropriate methods, stating that Terry “takes kids outside to study forces” using the playground to study friction, momentum, gravity, push, and pull. Both highlighted Terry’s ability to “find new ways to engage students” in the content.

Instructional Practice. Terry’s responses on instructional practice combined instruction, planning, and assessment. Terry stated, “I typically use formative assessment more than summative. I’m constantly asking them how they are doing and checking their

understanding.” Terry detailed use of pre-assessment, differentiation, adjusting groups based on assessment results, and students tracking their own progress. Terry’s supervisor added that Terry is “always differentiating” and searching for ways to engage students in learning something new.

Professional Responsibility. When reflecting on professional responsibility, Terry stated “I’m always thinking about what we should change for next year.” Terry engaged students in self-reflection and modeled that teachers are learners too. Terry stated that collaboration with other teachers was helpful in deciding what students should know, understand, and do. Contrarily, Terry’s supervisor stated that collaboration is something that Terry is working on, and is one “weak area.” The supervisor also indicated there is a perception that Terry “doesn’t collaborate well with others.”

Jamie’s interview. Frequency of codes for interview results for Jamie and Jamie’s supervisor are displayed in Table 8. Significant statements captured during the interview offer explicit examples of teaching documented in other data sources.

Learner & Learning. Results of Jamie’s interview included responses related to the ability to get to know students and how they learn best as well as designing lessons tailored to learning styles. Jamie’s supervisor supported this claim, confirming that Jamie identifies students’ different abilities, strengths, and weakness, and makes adjustments to the curriculum to support needs. Jamie mentioned the importance of using outside resources to make lessons “relevant” and developing instruction that is “hands on and very active.”

Jamie shared the idea of spending time in the beginning of the year to ensure students are following expectations. The beginning of the school year is more structured as students learn and practice expectations. Jamie’s supervisor described the classroom environment as “fairly open” and “inviting for students to interact with each other.” The supervisor explained that Jamie’s management skills have grown over time and Jamie continues to be “fair and firm” when managing the classroom environment.

Content. Jamie indicated a sense of confidence in teaching math and the ability to direct students when they have misconceptions, as well as an ability to explain things in “multiple ways.” Less confidence was self-reported in Jamie’s ability to teach language arts. However, Jamie modeled how to research ideas and questions that were not well understood for students. Jamie believed that understanding content comes with time, but can also come naturally. Jamie’s supervisor responded to Jamie’s ability to understand and apply content, exemplified by going into a unit “very well-prepared.”

Instructional Practice. Responses established the use of assessment and a variety of strategies to drive instructional practice. During reading, Jamie indicated a variety of assessments are used at the beginning of the school year including AIMSweb reading fluency (NCS Pearson, 2017) and NorthWest Evaluation Association (NWEA) Measures of Academic Progress (MAP). These assessment results are used to place students in leveled groups. Jamie stated that assessment “is critical for every subject.” The supervisor explained that collaboration within the school supports the collection, use, and interpretation of assessment data. Jamie also explained that assessment results are used to differentiate within the classroom and that all students “learn in a different

way.” Instructional examples from the interview included vocabulary sorting, multiple choice quizzes, peer teaching, and mixed groups. In the area of technology in the classroom, Jamie identified the use of speakers, document cameras, tablets, and an interactive whiteboard. Jamie is described by the supervisor as “tech savvy.”

Professional Responsibility. Several examples of professional responsibility were identified. Jamie explained the importance of reflecting on teaching in order to improve instruction and discussed the effects of supervisor observation on improvement: “When my supervisor is observing, I will talk about improvement. The last time we talked I was told I had improved on classroom setup and organization. I appreciate the feedback.” Both Jamie and the supervisor alluded to the importance of being part of a collaborative team. Jamie said it “makes it easier when you plan with others” and the supervisor recognized that Jamie “is collaborating with someone every week.” The supervisor stated collaboration is “one of Jamie’s strengths as a young teacher.”

Table 7
Teacher and Supervisor Interviews: Frequency of Pre-Existing Codes: Terry

Terry							
Learner & Learning		Content		Instructional Practice		Professional Responsibility	
Learners' needs	5	Application of content	4	Assessment	8	Professional development	6
Responsive to diverse backgrounds	5	Content knowledge	3	Learners' needs	6	Reflection	6
Engagement	4	Modeling	3	Instructional supports	5	Collaboration	5
Instructional supports	3	Instructional strategies	3	Instructional choices	3	Instructional choices	2
Management	3	Instructional supports	2	Instructional strategies	2	Instructional supports	2
Positive Learning Environment	3	Learners' needs	1			Learners' needs	1
Total	23		16		24		22

Terry’s Supervisor							
Learner & Learning		Content		Instructional Practice		Professional Responsibility	
Engagement	5	Content knowledge	4	Assessment	4	Lack of collaboration	9
Positive learning environment	5	Engagement	3	Instructional supports	4	Professional development	3
Responsive to diverse backgrounds	5	Instructional strategies	3	Learners' needs	3		
School supports	3	Professional development	1	Instructional strategies	2		
Instructional choice	3						
Management	2						
Total	23		11		13		12

Table 8
Teacher and Supervisor Interviews: Frequency of Pre-Existing Codes: Jamie

Jamie							
Learner & Learning		Content		Instructional Practice		Professional Responsibility	
Positive learning environment	5	Application of content	3	Learners' needs	8	Professional development	9
Instructional choice	4	Instructional strategies	3	Instructional strategies	5	Collaboration	6
Learners' needs	4	Content knowledge	2	Instructional supports	4	Feedback from supervisor	2
Management	4					Reflection	2
Engagement	3						
Responsive to diverse backgrounds	1						
Total	21		8		17		19

Jamie's Supervisor							
Learner & Learning		Content		Instructional Practice		Professional Responsibility	
Learners' needs	6	Content knowledge	2	Instructional supports	6	Professional development	4
Positive learning environment	5			Learners' needs	3	Collaboration	3
Management	3			Feedback from supervisor	1	School supports	1
Instructional choices	2			School supports	1		
Responsive to diverse backgrounds	1						
Total	17		2		11		8

Cross Case Results

Overall, both participants agreed they were prepared to teach. In the area of Instructional Practice, Jamie self-rated slightly lower than Terry. This category was the highest rated for both participants (see Table 3). Terry and Jamie were more critical when they rated their preparation for Instructional Practice, marking themselves lower than their supervisors rated their performance in the classroom. Most items for both participants related to their preparation for Professional Responsibility were rated with overall agreement, either ratings of *agree* or *tend to agree*. The area of Professionalism had the most notable difference in ratings from participant self-reported preparation to supervisors' observation of implementation with the supervisors' rating higher.

On the survey response items associated with Diverse Learners, both participants rated a higher tendency to disagree the preparation program prepared them for teaching diverse learners. Although there was a higher rate of *tend to disagree*, no items were marked as *disagree* by or for either participant. Specific items of disagreement included differentiating instruction for students

with mental health needs, teaching English Learners, and accessing resources for students with diverse needs. Terry tended to disagree that the EPP prepared graduates to differentiate instruction for students with IEPs and 504 plans, while Jamie agreed with being prepared in that particular area. Overall, participants rated lower than their supervisors rated when teaching practices were observed.

There was no distinct pattern between participants or in comparison to their supervisors' responses on the Dispositions Evaluation other than an overall response of Distinguished or Proficient on the 19 professional disposition indicators. Terry had an overall score of 3.3 on a 4.0 scale indicating performance between proficient and distinguished, and Jamie self-rated with an overall score of 3.5. There was also no distinct pattern between participants and supervisor responses on the STOT other than proficient performance of teaching skills. The thematic analysis of observation field notes, graduate interviews, and supervisor interviews indicated knowledge, skills, and dispositions within the four areas of effective teaching were demonstrated by both participants during a classroom observation and confirmed in the interviews with examples and explanation.

Discussion

According to multiple data sources, both participants demonstrated effective teaching. Two EPP graduates participated in this case study approximately 2.5 years after completing a teacher education program. Results indicated both teachers effectively applied the professional knowledge, skills, and dispositions the program was designed to achieve. Upon reflecting on their experiences, each participant demonstrated overall proficiency of teaching in the areas of Learner & Learning, Content, Instructional Practice, and Professional Responsibility. Participants also demonstrated personal strengths and weaknesses. Content Knowledge, Instructional Practice, and Professionalism were three areas in which no significant weakness were noted by Terry, Jamie, or their supervisors. Opportunities for growth and continued dialogue were identified in the area of Learner & Learning. Three major recommendations emerged that warrant further discussion: working with diverse learners, appropriate and shared responsibility for teacher outcomes, and mechanisms for measuring teacher effectiveness.

Diverse Learners

Of all survey items, questions related to Diverse Learners ranked amongst the lowest rated. Specific topics included differentiating instruction, mental health needs, teaching English Learners (EL), and accessing resources for differentiation. However, results of classroom observations and supervisor interviews indicated participants exhibited the skills necessary to be responsive to diverse backgrounds, and learners' needs were met through differentiation.

These conflicting findings could be due to supervisors noticing only a glimpse of differentiation in the classroom, as observations are typically short and the instruction during that time may or may not require differentiation of process, product, or content. Observation assumes that the supervisor will get an accurate picture of the teacher's effectiveness in the classroom. However, teachers use interventions in a variety of ways and at chosen times and frequencies, so differentiation is not guaranteed to occur during the observed lesson. Instead, differentiation is

meant to be responsive to the needs of the learners. The discrepancy could also be attributed to teacher self-efficacy in meeting learners' diverse needs (Bandura, 1977).

Additionally, Terry and Jamie are both employed in a state that recently passed new legislative requirements for eight hours of mental health training for teachers every two years. This mandate could have resulted in an increased sense of intensity and immediacy, and perhaps a perception of personal ineptitude, to both of them regarding their interactions with diverse learners. This study's findings in this area are consistent with aggregate data from other teacher preparation programs across North Dakota, where they work (North Dakota Association of Colleges for Teacher Education, 2017). New teachers consistently indicate they feel less prepared to work with students with diverse needs (Eberly, Joshi, Konzal & Galen, 2010). Because the area of Diverse Learners was identified by graduates as an area of challenge across multiple teacher education programs in North Dakota where this study was conducted, a state-level subcommittee is working on a response, with focus groups planned for further investigation and action.

Another factor that may have resulted in lower rankings is that skills needed to work with diverse learners are often developed after initial teacher preparation programs are completed (Eberly et al., 2010). In fact, initial training followed by coaching with corrective support has been found to be most helpful for teachers to build and maintain skills (Kretlow & Bartholomew, 2010). Upon graduation and subsequent hiring, teachers have more opportunities to practice and apply skills for teaching diverse learners in real-life classroom settings. The question might be, do graduates lack the knowledge and skills necessary to successfully meet the needs of diverse learners or do they lack opportunities to develop and apply their knowledge and skills?

Through data analysis and ongoing communication with the graduates' EPP, several changes arose that had the potential to influence candidates' knowledge, skills, and dispositions in the area of Diverse Learners. First, the participants' preparation program did not have required coursework in teaching ELs during the participants' time in the program. Coursework was added as a requirement after participants had exited the program. Second, significant changes in curriculum occurred, including the identification of diversity objectives, expansion of diverse field experiences, curriculum mapping, and hiring of an institutional Director of Diversity and Inclusion. Furthermore, the EPP's elementary education program has since been revised to include an embedded minor in special education and pedagogy for responsiveness to adverse childhood experiences. These factors, among others, likely led to lower self-ratings of preparedness on items related to Diverse Learners, and informed the EPP of opportunities for continuous improvement that aligned with addressing the changing demographics of schools.

A final area of consideration is that participants' responses on the surveys were more reflective of their current classroom performance rather than their level of preparedness. This dichotomy presents challenges when making the connection back to the educator preparation program (American Psychological Association, 2014), which leads to the question: What is the best way to bridge the gap between the shared responsibility of preparation programs and administrators/schools in preparing teachers to work with diverse learners? School partners and EPPs must establish mutually agreed upon expectations of in-service teachers at different stages in their careers since teacher development does not end when teachers graduate from training programs.

Shared Responsibility

School administrators and EPPs must work together to evaluate effectiveness of new teachers and help inform best practices for continuous improvement. Educator preparation programs must rely on multiple measures of beginning teacher effectiveness to provide a fair and well-rounded picture of a teacher's strengths and challenges. Defining and measuring beginning teacher effectiveness are both the work of EPPs and of P-12 schools. Administrators have a holistic view of teacher effectiveness and employ teachers who have graduated from multiple programs. Inclusion of supervisors' voices alongside teachers in this study emphasized how programs can use administrator feedback to inform programs and better prepare teachers. Responding to this shared responsibility has the potential to improve the new teacher workforce and ultimately, P-12 student achievement.

But what does this look like in action? The way in which this shared responsibility is carried out can be a vague and idealistic concept, as is sometimes the case with annual advisory board meetings or undocumented conversations that occur during the daily cycle of teaching and learning. Educator preparation programs can help districts identify areas for focused induction and specialized mentorship support. Areas identified in this study included teaching students who are ELs, from diverse backgrounds, and have special education or mental health needs. Communication and the sharing of information can inform improvement efforts for accrediting the EPP and improving both current teacher supervision/evaluation models and overall school improvement systems.

As these cases identified, some aspects of teacher training are addressed at the initial licensure level and require support to advance proficiency. Barnes and Smagorinsky (2016) noted that beginning teachers are expected to be highly proficient after a few semesters of coursework and practica and a semester of student teaching; this often leads to fragmented understanding of how to teach (p. 342). This seems to be the case for participants in this study. Shared responsibility brings the opportunity for EPPs and schools to establish mutually agreed upon expectations and actionable descriptors of proficiency at different career levels. Expectations of professionalism are being utilized (CCSSO, 2013; Danielson, 2013; Marzano, 2017), but the difference in degree of proficiency among graduates, well-prepared novice teachers, and veteran teachers remains unclear amongst stakeholders. This can create a gap between employer expectations and EPP completion requirements, and could compromise the efficacy of professional development and mentoring support schools provide.

Teaching expertise is developed over time in a nonlinear fashion. The early years of teaching give way to increasing complexity and sophistication of practices as understanding and application are no longer compartmentalized and separate from lived, classroom experiences. Advanced application and refined implementation divide the novice teacher from the distinguished (CCSSO, 2013). It is clear that "teachers need time to process new ideas, consolidate skills, and begin to make changes to their teaching practice" (Kutaka et. al, 2017, p. 150). EPP's are well-situated to serve as induction support by working with graduates through their growth process. Professional development should occur in areas defined by graduates as weaknesses in preparation and by supervisors in the lack of implementation. This may result in P-12 administrators advocating the value of the EPP outcomes data for their schools in addition

to EPPs' accreditation and improvement processes. To systematically develop teacher expertise in this way, Marzano, Frontier and Livingston (2011) acknowledged five necessary conditions: "a well-articulated knowledge base for teaching, focused feedback and practice, opportunities to observe and discuss expertise, clear criteria and a plan for success, and recognition" (p. 4).

A focus on advancing these complexities of teaching is reciprocally beneficial as newly licensed teachers require "re-education" in many states, and EPPs are positioned to meet the need of granting transcribed credit, fulfilling the common mission of institutions of higher education, and enhancing relationships within communities. Building teacher capacity to improve instruction and student outcomes can be realized through this targeted collaboration and increased voice of supervisors. Sharing information and best practices regarding the ways we prepare and support beginning teachers across institutions is fundamental to systemic progress.

Mechanism for Measuring Teacher Effectiveness

Leaders in educational research have called out the need to test assumptions of improving overall quality of education based on common evaluative metrics (Tatto et al., 2016). Adequately determining teaching effectiveness requires a multitude of valid measures. Examples of common measures include reports from teachers, reports from administrators, student outcome measures, observation data, performance assessment, graduate surveys, supervisor surveys, summative grades, standardized test scores, social/emotional outcomes, and classroom observation rubrics (ND DPI, 2015). This case study sought to include multiple sources of evidence to answer the question about teachers' knowledge, skills, and dispositions and explore the practicality of research using such measures. Few studies have been published that provide EPPs with a clear process for how this might be accomplished through case study, yet this replicable protocol, or portions of it, does provide EPPs with possibilities.

The TTS brought the voice of the teacher on perceived level of preparedness, while the aligned SS survey added the supervisor's opinion of observed implementation. The STOT asked the teacher and the supervisor to rate observable teaching skills, and the Disposition Evaluation explored the "habits of professional action and moral commitments that underlie the performances" (CCSSO, 2013, p. 60). These measures were easily administered through protected digital systems and analyzed through technology-based applications. Additionally, these measures were already embedded as part of longitudinal tracking of candidate growth from program admission to graduation. The measures were administered in this study as another data source at nearly three years of in-service teaching experience. This may serve in the future to explore predictive elements across measures. Classroom observations gave evidence of implementation, a rich description of the learning environment, and examples of effective practices. Interviews provided an opportunity for both the teacher and supervisor to explain responses and allowed for triangulation. These measures met professional standards of research and technical quality, yet were notably cost and time intensive. Even given these multiple measures, the researchers acknowledge that "education is mandatory but learning is not... teachers cannot succeed, then, unless their clients choose to learn" (Kennedy, 2016, p. 11). Selected mechanisms must always be considered with this persistent challenge in mind.

Results of this case study (as part of a larger investigation to describe teaching effectiveness, connections to P-12 student learning, and impact of the preparation program) can speak to the viability of case study as a feasible suggestion for accountability evidence (CAEP, 2016). A complete case study with several components of effective teaching, each examining different perspectives/frames of reference, did assess outcomes of teacher education. While time intensive and not solely representative of the program, results supplement other program impact evidence and have proved meaningful for improvement decisions. However, case study design is only worthwhile to use if the benefit exceeds the cost in terms of EPP resources and capacity as well as usefulness for informing improvement.

One source of data in particular presented itself as a feasible way for EPPs to efficiently and effectively gauge the performance of their graduates while limiting burdens of cost and time. Supervisor evaluation is already used to rate overall performance and enhance teacher effectiveness (Marzano et al., 2011) as required by the Every Student Succeeds Act (ESSA). While formats and processes differ across schools, this option leverages professional expertise of administrators and a processes already in place. These teacher evaluation models should specify differentiated performance levels (e.g., non-proficient, developing proficiency, proficient, exemplary) as well as identify strengths and areas for growth. Based on supervisor evaluation methods in this study and the magnitude of common statewide metrics utilized within these cases, it makes sense to request that graduates consent to voluntary, self-submitted evaluations as evidence of teaching effectiveness along with survey completion. This would honor concerns of privacy protection. Teacher advocacy groups are rightfully apprehensive that teacher growth and improvement would be hindered because of publicized evaluative measures, yet voluntary submission would provide EPPs with important information without leveraging open access parameters. Case study research can serve as a means for deeper exploration to address patterns or particular cases of interest paired with overall program impact, as well as leverage partnerships with P-12 schools and expertise of school administrators.

Study Limitations

Although case study is particularly situated for investigating complex educational phenomenon and advancing research on accountability measures, limitations have been identified. Efforts were made to address bias, yet bias remains an inherent issue in case study research. Furthermore, this inquiry was an attempt to understand the collective cases as a whole, not the various parts of the case or the contributing factors that influence the case (Baxter & Jack, 2008); as such, causation is not implied. Saturation of data for rich, thick description is limited due to purposeful selection of only two cases. Participants were not representative of all graduates; findings are not generalizable for the EPP or beyond. Lastly, the two cases in this study were both elementary teachers; investigation of early childhood and secondary graduates is underway.

Conclusion

Due to the complex nature of teaching, educational researchers struggle to connect the impact of teacher education programs to observed teacher classroom performance after graduation. It is not always clear whether a teacher's knowledge, skills, and dispositions, or the achievement of students in their classrooms are due to teacher training or other intricate, intervening factors.

Three avenues may strengthen understanding of the effectiveness of program graduates. First, this study could be repeated on a larger scale to analyze additional cases as well as include secondary education and early childhood education graduates. Second, future research could connect results of this study to research on available student growth measures, specific to teachers' own classroom experiences (e.g., teacher-made pre and post-assessments or student-growth percentiles), in order to provide a holistic picture of how teaching skills impact student learning outcomes. Third, it would be informative to further study the perceptions of graduates and their supervisors in judging EPP success when considering how well P-12 students learn from graduates of respective preparation programs. The implications for evaluating a new teacher workforce, and the programs which prepare them, would be well served by these sources of acceptable evidence.

Author Notes

Sarah K. Anderson is Associate Professor of Education at Mayville State University.

Brittany D. Hagen is Assistant Professor of Education at Mayville State University.

Andrea L. Dulski-Bucholz is Dean and Chair of the Division of Education and Associate Professor of Education at Mayville State University.

Christopher M. Whitsel is Associate Professor of Sociology at North Dakota State University.

Kayla D. Smith is Data Management and Student Placement Coordinator at Mayville State University.

Ann M. Willeson is Professor of Education at Mayville State University.

Correspondence concerning this article should be addressed to Sarah K. Anderson at sarah.anderson2@mayvillestate.edu.

References

- American Psychological Association (APA), APA Task Force Report (2014). *Assessing and evaluating teacher preparation programs*. Retrieved from <https://www.apa.org/ed/schools/teaching-learning/teacher-preparation-programs.pdf>
- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., Rothstein, R., Shavelson, R. J., & Shepard, L. A. (2010, August). *Problems with the use of student test scores to evaluate teachers. (EPI Briefing Paper #278)*. Washington, DC: Economic Policy Institute.
- Ball, D. L., & Forzani, F. M. (2011). Teaching skillful teaching. *Educational Leadership*, 68(4), 40-45.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215.
- Bandura, A. (1989). Social cognitive theory. In R. Vasta (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). New York, NY: Academic Press.
- Barnes, M. E., & Smagorinsky, P. (2016). What English/language arts teacher candidates learn during coursework and practica: A study of three teacher education programs. *Journal of Teacher Education*, 67(4), 338-355.
- 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, September). *Teacher preparation and student achievement* (Working Paper No. 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(2), 77-101.
- Briggs, D., & Domingue, B. (2011, February). *Due diligence and the evaluation of teachers: A review of the value-added analysis underlying the effectiveness rankings of Los Angeles Unified School District teachers by the Los Angeles Times*. Boulder, CO: National Education Policy Center. Retrieved from <https://files.eric.ed.gov/fulltext/ED516008.pdf>
- 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* (pp. 69-132). San Francisco, CA: Jossey Bass.

- Cochran-Smith, M., & Zeichner, K. (Eds.). (2005). *Studying teacher education: The report of the AERA panel on research and teacher education*. Washington, DC: American Educational Research Association.
- Council for the Accreditation of Educator Preparation. (2013, February). *CAEP accreditation standards*. Washington, DC: Council for the Accreditation of Educator Preparation.
- Council for the Accreditation of Educator Preparation. (2015, September). *Questions and answers on Standard 4, Program impact*. Washington, DC: Council for the Accreditation of Educator Preparation.
- Council for the Accreditation of Educator Preparation. (2016, July). *When states provide limited data: Guidance on using Standard 4 to drive program improvement*. Washington, DC: Council for the Accreditation of Educator Preparation...
- Council of Chief State School Officers (CCSSO). (2013, April). *InTASC model core teaching standards and learning progressions for teachers 1.0: A resource for ongoing teacher development*. Washington, DC: Council of Chief State School Officers.
- Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five traditions* (2nd ed.). Thousand Oaks, CA: Sage.
- The Danielson Group. (2013). *2013 framework for teaching evaluation instrument*. Princeton, NJ: Danielson. Retrieved from <https://www.danielsongroup.org/framework/>
- Dewey, J. (1938). *Experience and education*. New York, NY: Touchstone.
- Eberly, J. L., Joshi, A., Konzal, J., & Galen, H. (2010). Crossing cultures: Considering ethnotheory in teacher thinking and practices. *Multicultural Education, 18*(1), 25-32.
- 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.
- Glaser, B. G. & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago, IL: Aldine Pub. Co.
- Humphrey, D., Wechsler, M., & Hough, H. (2008). Characteristics of effective alternative programs. *Teachers College Record, 110*(1), 1-63.
- Kennedy, M. (2016). Parsing the practice of teaching. *Journal of Teacher Education, 67*(1), 6-17.

- Kretlow, A. G., & Bartholomew, C. C. (2010). Using coaching to improve fidelity of evidence-based practices: A review of studies. *Teacher Education and Special Education, 33*(4), 279-299.
- Kutaka, T. S., Smith, W. M., Albano, A. D., Edwards, C. P., Ren, L., Beattie, H. L., ... & Stroup, W. W. (2017). Connecting teacher professional development and student mathematics achievement: A 4-Year study of an elementary mathematics specialist program. *Journal of Teacher Education, 68*(2), 140-154.
- Marzano, R. J. (2017). *The new art and science of teaching*. Alexandria, VA: ASCD.
- 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, CA: Jossey-Bass Publishers.
- Morse, J. M. (1994). Emerging from the data: Cognitive processes of analysis in qualitative inquiry. In J. Morse (Ed.), *Critical issues in qualitative research* (pp. 23-43). Menlo Park, CA: Sage.
- NCS Pearson, Inc. (2017). *aimswEBPlus technical manual*. Bloomington, MN: NCS Pearson, Inc. Retrieved from <https://cdn2.hubspot.net/hubfs/559254/AimswEBPlus/Docs/aimswEBPlus%20Technical%20Manual.pdf>
- Network for Excellence in Teaching (NExT) Consortium. (2016). *Our operating principles*. Retrieved from Network for Excellence in Teaching website: <http://www.nextteachers.org/our-operating-principles>
- 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 University of Florida, Collaboration for Effective Educator, Development, Accountability, and Reform Center website: <http://cedar.education.ufl.edu/tools/literature-syntheses/>
- North Dakota Association of Colleges for Teacher Education (NDACTE). (2017). *North Dakota common metrics 2017 Transition to Teaching Survey: North Dakota aggregate report*. Unpublished internal document.
- North Dakota Department of Public Instruction (ND DPI). (2015, May). *North Dakota teacher evaluation guidelines*. Bismarck, ND: K. Baesler.

- Regional Educational Laboratory (REL) Central. (2014). *Research that relates teacher preparation to teacher and student outcomes: A literature summary*. Unpublished limited distribution report for the Educator Effectiveness Research Alliance.
- Tatto, M. T., Richmond, G., & Andrews, D. J. (2016). The research we need in teacher education. *Journal of Teacher Education*, 67(4), 247-250.
- Toch, T., & Rothman, R. (January, 2008). *Rush to judgment: Teacher evaluation in public education*. Washington, DC: Education Sector.
- Visible Learning Plus. (2018). *250+ Influences on student achievement*. [PDF file]. Retrieved from https://us.corwin.com/sites/default/files/250_influences_10.1.2018.pdf
- 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.

Appendix A: Interview Protocol for Graduates

Demographic Information

1. Please state your current teaching position and how long you've been employed.
2. At what university did you complete your teacher preparation program?
3. What was your major and/or minor program of study?
4. What teaching license(s) do you currently hold?

The Learner and Learning:

5. Describe the factors that contribute to your students' learning.
6. Explain how your knowledge of learner development (cognitive, linguistic, social, emotional, and physical) impacts your students' learning?
7. How do you use your knowledge of students' socioeconomic, cultural, and ethnic differences to meet their learning needs?
8. How does your classroom environment promote student learning and engagement?
9. Explain how your classroom management skills impact student learning?

Content:

10. Describe how you guide students to master content through relevant, real-life learning experiences.
11. How does your knowledge of content impact your students' learning?

Instructional Practice:

12. How does your use of assessment data contribute to your students' learning?
13. How do your practices of differentiating instruction contribute to student learning?
14. Describe your use of technology to support student learning?

Professional Responsibility:

15. Describe how you use reflection to improve your instruction.
16. Describe how you collaborate with others, both in-school and outside of school, to improve student performance.
17. List the professional development opportunities you've participated in since graduation and how they have impacted your students' learning?

EPP Impact:

18. Are there any factors that limit, or have limited, your ability to teach effectively? Please explain.
19. Was there anything that was repetitive or lacking in your teacher preparation?
20. What other factors (besides your knowledge, skills, dispositions, or EPP training) influence your students' achievement?
21. Do you think your training at [university] was effective? Please explain.
22. Do you perceive your EPP preparation as relevant to the responsibilities you confront on the job? Please explain.
23. Is there anything else you would like me to know?