Executive Summary of Case Studies
Spring 2017

Introduction
- **Why we conducted case studies**
- **Background on how the research group was formed, meetings, discussion, development of research protocols and materials used in the research**

SUNY Cortland collected impact and application of knowledge/skills disposition data because there is an absence of such data from the New York State Education Department.

A Professional Learning Team (PLT) conducted case studies in order to both examine program completer impact on student learning (CAEP 4.1) through the review and discussion of available artifacts and also to observe the effective application of professional knowledge, skills, and dispositions (CAEP 4.2) with the same participants through structured observation protocols and pre- and post-observation interviews.

We also hope to use this set of case studies as a pilot research project to inform the systematic and ongoing investigation of program completers' impact on their students.

Dean Andrea Lachance proposed a Case Study Research design and sought funding and research team members from different departments to assure a representative sample from across the unit. A team of 6 researchers (2 from Adolescence Education, 2 from Physical Education and Health, 2 from the School of Education) convened on February 23rd, 2017 to decide on initial protocols for recruiting program completers. Additional meetings were held to determine interview protocols connected to CAEP components 4.1 and 4.2 and observation protocols to use based on already validated observation rubrics (e.g. Danielson, 2011) in use in New York State's Annual Personnel Performance Review (APPR). An IRB was submitted to the IRB Chair at SUNY Cortland and approved on 3/17/2017. Research was conducted soon after with each researcher interviewing two (three, in one case) program completers.

Overview of procedures and methods for all 13 case studies

The research team described in the Institutional Review Board (IRB) application a descriptive, multiple case study research design of 13 program completers across P-12 content and grade levels to include semi-structured interviews and structured observations. A purposive sampling approach was used to focus on information-rich cases that produced data to inform the investigation of impact and effectiveness across a representative sample of program completers.

The plan included collecting observation data using questions and prompts adapted from Charlotte Danielson's (2011) *Framework for Teaching Evaluation Instrument*, as applied within New York State for teacher observation and evaluation. The instrument used within Danielson's framework has been validated as a tool for observing teaching practice across multiple classrooms.

Upon IRB approval, the research team sent out an initial recruitment email to program completers from targeted SUNY Cortland teacher education programs, including the Childhood/Early Childhood Education, Foundations and Social Advocacy (special education), Health and Physical Education, Adolescence English Education and Adolescence Math Education programs. The initial recruitment email set out basic parameters for the study (see Appendix C, IRB), including that each participant share artifacts that demonstrate impact on student-learning growth.

The list of artifacts requested from each participant is detailed in Appendix D of the IRB Application, and included such items as APPR data from previous year's final scores, portfolios with Student Learning Outcomes (SLO) data, lesson plans, student work samples, teacher reflections, pre- and post-teaching assessments for individuals, small, and whole groups, teacher-generated curriculum guides, block plans, scope and sequence plans, communications with parents/community, efforts to differentiate instruction and student choices, and other artifacts selected by participants. Participants were expected to sample from this list and were not required to have most or all artifacts.

Each researcher conducted one initial interview to review artifacts, and then returned to conduct one structured observation, first completing a pre-observation interview and following up with a post-observation interview with each participant.

Data was compiled within forms that included Interview 1 and Pre- and Post-observation questions. Data also included collecting the structured observation tools and completing a Case Study Template to summarize and synthesize the information for each case.

Cross-case analysis was conducted in writing and at a final meeting on July 12th, 2017. A structural narrative analysis was conducted in stages where the group focused on themes within and across cases and then research team members checked these themes within and across the case study templates provided.

List of 13 teacher-participants, their grade level(s), subject area, and SUNY Cortland program from which they graduated

Participant Code	Grade Level(s)	Subject Area(s)	SUNY Cortland Program
4101	7-12	Physical Education	B.S.Ed. in Physical Education
4102	7-8	Health	B.S.Ed. in Physical Education; M.S.Ed. in Adapted Phys Ed + NYS Health Certification
4103	K-5	Physical Education	B.S.Ed. in Physical Education
4104	7-12	Physical Education	B.S.Ed. in Physical Education; MST in Health
4105	7-8	Math	B.A. and M.S.Ed. in Adolescent Education - Math
4106	9-10	Math – Pre-Algebra	M.A.T. in Teaching Math 7-12
4107	7	ELA	B.A. in 7-12 Adolescent Education in English
4108	6 (middle school setting)	ELA	B.A. in 7-12 Adolescent Education in English
4109	3	Gen Ed (Multi-Subject)	B.S. in EDD; M.S.Ed. in Literacy
4110	3	3rd grade Gen Ed (no Soc Studies; focus on science)	M.S.T. in Childhood Education
4111	2-3	Special Ed Math/ELA	B.S. in Inclusive Special Education
4112	K-3	Special Ed Math/ELA	B.S. in Inclusive Special Education
4113	8	Advanced Algebra	B.A. 7-12 Adolescent Education in Math

Context and description of the P-12 learning data that was examined (CAEP 4.1 rubric)

The P-12 learning data that was examined included lesson plans, student work samples, formative assessments, announced and unannounced observation rubrics filled out by administrators with aggregated and disaggregated scores, pre-assessments used for Student Learning Outcomes (SLOs), progress monitoring data related to fitness, literacy, and other skills collected manually and through online/computer-based applications, teacher reflections, writing samples, exit tickets, student handbooks with compilations of worksheets, and portfolios with collections of student work and other assessment data.

The context of this P-12 learning data includes data collected by 13 completers working in rural, suburban, urban, small village, town, and city settings, within and one outside of New York State. The completers represent teachers of physical education, health, elementary and middle/high school math, multiple subjects in primary and middle elementary grades, literacy interventions, and middle/high school literature and composition. The participants are all practicing teachers who hold full-time positions and who graduated from a SUNY Cortland program within the last 7 years. The majority of participants were in their first two years of teaching.

The population of students taught by these 13 program completers included kindergarteners through twelfth graders, many of whom receive free and reduced lunch. Students are diverse in racial, ethnic, and socio-economic status background, and some are English language learners.

Description and explanation of the representativeness of the data (CAEP 4.1 rubric)

All participants shared lesson plans and some type of formative assessments. Participants also shared data in the form of progress monitoring (e.g. health statistics, phonemic awareness skills), exit tickets, administrator observations, some Student Learning Outcome (SLO) pre-assessments, student-created goals, some classroom curriculum-based pre- and post-assessments, and some written work samples. There was a variety in forms of assessment though few samples that showed student growth across time.

While all researchers did ask about standardized assessments connected to the state's APPR process, very few participants were able to show student growth in artifacts connected to the APPR process or to assessments that districts and states regularly collect from K-12 students.

Analysis of Data

Evidence that demonstrated completers contribute to an expected level of student-learning growth draws from small samples of students and demonstrates some growth in discrete skills like phonemic awareness (participants 4111, 4112) or in fitness statistics (participants 4101, 4103, 4104). Few completers could show both pre- and post-assessments and most shared artifacts like lesson plans and formative assessments that did not also demonstrate the outcomes of those lessons and how the formative assessments were used to inform instruction and connect to student-learning outcomes. Even when participants had online/computer-based systems to track things like fitness statistics (e.g. FitnessGram) or independent work from students (e.g. *Classworks*), few showed how they did track growth in what their students were learning.

Most participants could produce exit tickets and end products that met criteria stated in lesson plan objectives, rubrics or even in goals set by students themselves (participants 4107, 4108, 4109, 4110, 4111, 4112). One participant pulled worksheets together into one student handbook to show student learning within one unit of study (participant 4102).

Some participants shared anecdotes of how they taught skills related to movement, for example, in a classroom setting and then observed their students using those skills taught outside of the classroom setting (participants 4101, 4103, 4104). Other anecdotes include observing greater quantities of writing (participant 4112) and the ability of students to explain what was taught to their peers (participants 4105, 4106, 4113) or the ability of students to set their own goals and track them (participant 4109).

Many of the participants are using school- and district-prescribed programs and curriculum and seemed hesitant to use individual student data to influence the pace of instruction. Participants felt they must follow the district's or school's pacing guide and that prescribed guide and curriculum had more of an impact on participants' decision-making than their assessments of students' learning (participants 4107, 4108, 4109, 4110)

Overview of evidence that completers effectively apply the professional knowledge, skills, and dispositions that the preparation experiences were designed to achieve (CAEP 4.2)

This section focuses on themes from the evidence that are related to each of the four Domains on the Structured Observation Rubric, which is aligned to Danielson's (2011) framework used to observe in-service teachers in New York State. Please see the Observation Ratings document for disaggregated average scores in each domain for each participant.

Planning and Preparation (3.18/4 average across participants)

Effective application of planning skills was demonstrated and documented within the observation rubrics across the 13 cases. Some participants not only created their own curriculum but were also considered curriculum leaders (participants 4101, 4102, 4103, 4104). Among these participants, only one was given some curriculum materials in the form of a set of textbooks (participant 4102) and so had to construct the student learning experience independent of any guidance.

All participants demonstrated they could plan using standards and clear objectives, and some participants had a highly prescribed set of behavioral as well as academic objectives they were expected and required to follow. There was a notable absence of planning for assessment of cognitive and affective domains where participants teaching physical education often focused instead on planning to track fitness outcomes from their students, possibly due to fitness tracking being the form of standardized assessment most accessible to the participants (participants 4101, 4103, 4104).

Participants could plan for multiple abilities and could give students choices and accommodations, some based on individualized goals (participants 4101, 4102, 4103, 4104, 4111, 4112), as demonstrated on written lessons sent previous to observations. Examples of accommodations and adaptations included changing distance, speed, size of equipment (participants 4101, 4103, 4104) and levels of texts (participant 4112).

Weaknesses include a lack of clear planning for ongoing assessment and, at times, the lack of the ability to articulate criteria for assessments in place.

Classroom Environment (3.48/4 average across participants)

Most participants had clearly posted objectives, a performance that is emphasized in preparation programs and understood to be valued by school building leaders (participants 4101, 4102, 4103, 4104). These objectives were sometimes read by participants' students in the form of "I can" statements at the beginning of lessons that were observed (participants 4111, 4112).

Participants demonstrated the ability to develop rapport with their students, many working with students beyond and outside of the classroom environment in activities like coaching, running clubs, and team sport activities (participants 4101, 4103, 4104) and on field trips (participant 4112). Participants were strong in using highly structured modules and protocols to facilitate activities like student debates and discussion (e.g. in Socratic Seminars) where students clearly shared their own ideas based on the curriculum and texts they were reading (participants 4107, 4108).

Participants showed strengths in setting up well-organized, literacy/content-rich classroom environments with a range of manipulatives and text materials (participants 4109, 4110, 4111, 4112).

Instruction (including Assessment) (3.03/4 average across participants)

Participants demonstrated effective application of constructivist and social learning theory approaches in how they encouraged student voice (participants 4107, 4108), student goal-setting (participants 4109, 4110), and peer to peer teaching and explanation of concepts (participants 4105, 4106, 4113).

In artifacts related to observations from administrators, no participant received an Ineffective rating throughout the rubrics and all improved over time in the observation assessments shared with the researchers. If participants were once rated as Developing, they worked their way up to Effective and Highly Effective, often within a year (participants 4105, 4106, 4111, 4113).

Participants often used multi-modal instruction that included a variety of materials, activities, and pacing and the use of peers to scaffold learning (participants 4107, 4108, 4111, 4112). Most participants set clear behavioral expectations and managed time and space and students' questions with skill (participants 4107, 4108, 4111, 4112).

While some participants constructed their own curriculum, the majority followed a prescribed set of curricular and behavioral expectations and did so effectively (participants 4107, 4108, 4109, 4110, 4111). Behavioral expectations were especially consistent in their application (participants 4107, 4108, 4112). Across settings and diverse groups of students, researchers observed that students responded well to highly structured and consistently applied expectations. Notable was how participants were able to use the rapport they had with students to demonstrate that the prescribed curriculum had value and was positive and participants were able to build students' enthusiasm for the activities and content at hand, balancing encouragement with drawing out students' ideas (participants 4107, 4108, 4111, 4112). In post-observation interviews, participants confirmed that they were expected to conform to the curriculum and behavioral expectations that were prescribed in their school and district and noted that they would most likely have to leave/lose their position if they did not follow the prescribed curriculum (participants 4107, 4108, 4109, 4110).

The majority of participants had full command of their classrooms and were able to manage all corners of the classroom (participants 4107, 4108) and to clearly communicate what was expected as an end product and outcome of the lesson. Teaching students to write persuasive paragraphs, for example, often mimicked a test-taking task versus a writing-process-oriented progression of tasks (participants 4107, 4108).

The same weakness evident in planning was evident in instruction – related to assessment. When asked about how they knew their students are learning – and what their impact on that learning was – most participants discussed how they met individual students' needs and rarely focused on helping students through the processes encouraged in Common Core math concepts or in process-based writing. While participants were able to project a rubric for students to read criteria, some participants were unable to explain the criteria in a clear manner or to connect it to student growth over time. When discussing different outcomes for different students, some participants could group students by how those students performed and reflected less often on the choices they as teachers were making to help impact those different performances.

Professional Responsibilities (3.41/4 average across participants)

The majority of participants recognized the need to continue learning, to use standards that were instilled in their programs, and to find curriculum and text materials that matched their students' needs (participants 4101, 4102, 4103, 4104, 4111).

Most participants demonstrated effective record-keeping and continuous seeking of professional development opportunities, including pursuing master's degrees, additional certifications and training, and collaborative problem-solving. This was especially true for participants who found gaps in their preparation – in literacy interventions, for example (participants 4111, 4112).

In reflecting on their instruction, the majority of participants identified the need to keep working on assessment, though only the more experienced participants reflected on how to move away from compliance-based curriculum, assessment, and grading. Some participants suggested their own students impacted their teaching approach the most, while more experienced teachers sought to observe other master teachers and use professional development groups to observe and learn beyond what their students need in the moment of teaching.

Themes in the area of professional responsibilities are mostly connected to finding a way to meet individual students' needs while also working within curriculum that is either not yet aligned to a cohesive set of expectations or that is highly prescribed by the school and district. The inability of some participants to actively reflect on and show evidence of impact on student learning is another theme and potential gap in this area.

Trends

Schools are not holding all teachers accountable to the same measures. Even within the APPR process in New York State, it seems that special education teachers use very general school-wide SLO assessments and some physical education teachers are evaluated by ELA scores, for example. The current evaluation system for teachers seems to provide little reliability with great variety from district to district. This makes educator preparation program research of program completer impact difficult, as there is no consistent form of data to provide evidence of student-learning growth.

While P-12 school systems are not yet providing reliable measures of teachers' impact and effectiveness, SUNY Cortland is able to demonstrate that outcomes shown within other unit assessments like the Exit, Alumni, and Employer surveys, as well as the Student Teaching Evaluation (Revised), are consistent and also connect to the outcomes found within this case study research. Across these assessment tools, strengths in planning and instruction and weaknesses in assessment clearly emerge.

There was also continuity seen from preparation programs to observations of completers within this study and completers' focus on connecting with and encouraging students and their enthusiasm for and confidence within peer to peer learning, a love of literature and learning, movement and health, student voice, and expectations for appropriate behaviors.

Other trends that emerged within the findings include questions related to teacher autonomy, and whether or not SUNY Cortland teaches its candidates how to integrate best practices of assessment within a prescribed curriculum. While preparation programs provide instruction in assessment, it is possible that the lack of authentic assessment tasks and responsibilities for a whole classroom/set of students prevents preservice candidates from learning in-depth about how to enact meaningful assessment until they arrive in their own classrooms.

The idea of a developmental continuum for completers in their first years and then later in their practice emerged in discussions and analysis of the case study data. There was a marked difference in how one teacher could articulate impact and another teacher in the same grade level and school struggled to connect growth in student-learning to the teacher's actions and impact. This was also reflected in how some participants (especially first-year teachers) showed a lack of understanding of the state-wide teacher evaluation system (APPR) and others could explain and integrate that process into reflection on their own impact and practice.

When discussing how completers use what is taught in their preparation, participants show they are skilled enough to either create their own curriculum or to deal with prescribed curriculum; and they survive. Questions were raised within this study about whether or not program completers continue to adhere to the core of what was taught in their preparation programs, related to problem-solving, decision-making, judgment, and best practices for teaching and assessing conceptual learning.

Program Changes

The research team discussed how some of the gaps that emerged in this case study research (and in other unit assessments) are already being addressed, to include the expansion of grade levels in physical education curriculum courses, the addition of more literacy credits in the dual certification childhood and special education program, and the use of the edTPA across preparation programs to move the focus from teacher inputs to student outcomes.

Other potential program changes include the consideration of a new teacher mentoring program with completers, though school districts in New York State are already mandated to provide this. Resources in the form of standards, curriculum, and classroom management guidance could still be offered and the challenge of keeping in touch with program completers has partially been addressed with the collection of Exit Surveys that request contact information from graduates before they leave campus. The idea of

creating a New Teacher Support Center where resources are maintained and advertised could form part of the resources that program completers need once they complete their initial preparation program.

The question of a developmental continuum for teachers continues to emerge. In the area of assessment, researchers raise the question of whether or not preservice and even first year teachers are ready to learn in-depth about forms of process-based assessment. Instead of the integration of assessment into curriculum courses, it is possible that preservice candidates need entire courses on assessment, and then need to take assessment courses again once they are in their first years of teaching. Considering that teachers need multiple opportunities to study and learn about assessment will require a paradigm shift in how educator preparation programs apply their resources for assessment.

Sustainability of Research with Program Completers

In considering the sustainability of this approach to assessing program completer impact and effectiveness, the idea of recruiting faculty and asking them to cycle through this kind of research project – providing compensation to faculty for each observation/case completed – has the potential to make this kind of research an addition to faculty as well as program development. The methods used in this study were considered valid to this point and the case study approach of collecting multiple forms of interview and observation data with a cross-case analysis helps bring a depth of understanding to the evidence at hand.

One thing to consider is timing. If the research period were spread out across a longer period of time with more follow-up, this would allow researchers to better know the context of the program completers and also give researchers more time to check in with their team. A suggestion is to put a call out to graduating students who were just in student teaching to join the research study, in order to get an earlier start, and possibly to include some mentoring of that candidate/completer for a year.

Questions have also been raised about how to teach and model pre- and post- assessments in our own courses. The cycle of teaching, assessing, and then re-teaching is a hard cycle to simulate in teacher preparation. One improvement may be to have clinically rich programs where we observe candidates and P-12 students in schools, and then teach based on what both faculty and candidates observed.