



Public Schools of North Carolina
State Board of Education
Department of Public Instruction

Report to the North Carolina General Assembly

Advanced Teaching Roles Pilot
Evaluation - Preliminary Report
SL 2016-94 (HB 1030)

Research Organization:

*The William and Ida Friday Institute for Educational Innovation
North Carolina State University*

Date Due: June 15, 2018

Report # 88

DPI Chronological Schedule, 2017-2018

STATE BOARD OF EDUCATION

SBE VISION: Every public school student, through access to needed resources and rigor, will graduate ready for post-secondary education and work, prepared to be a globally engaged and productive citizen.

SBE MISSION: The State Board of Education will use its constitutional authority to lead and uphold the system of public education in North Carolina that guarantees every student in this state an opportunity to receive a sound basic education.

WILLIAM COBEY

Chair: Chapel Hill – At-Large

ERIC DAVIS

Vice Chair: Charlotte – At-Large

DAN FOREST

Lieutenant Governor: Raleigh – Ex Officio

DALE FOLWELL

State Treasurer: Raleigh – Ex Officio

MARK JOHNSON

Secretary to the Board: Raleigh

BECKY TAYLOR

Greenville – Northeast Region

REGINALD KENAN

Rose Hill – Southeast Region

AMY WHITE

Garner – North Central Region

OLIVIA OXENDINE

Lumberton – Sandhills Region

ALAN DUNCAN

Greensboro – Piedmont-Triad
Region

GREG ALCORN

Salisbury – Southwest Region

TODD CHASTEEN

Blowing Rock – Northwest Region

WAYNE MCDEVITT

Asheville – Western Region

PATRICIA N. WILLOUGHBY

Raleigh – At-Large

NC DEPARTMENT OF PUBLIC INSTRUCTION

Mark Johnson, State Superintendent :: 301 N. Wilmington Street :: Raleigh, North Carolina 27601-2825

In compliance with federal law, the NC Department of Public Instruction administers all state-operated educational programs, employment activities and admissions without discrimination because of race, religion, national or ethnic origin, color, age, military service, disability, or gender, except where exemption is appropriate and allowed by law.

Inquiries or complaints regarding discrimination issues should be directed to:

Maria Pitre-Martin, Ph.D., Deputy State Superintendent
6307 Mail Service Center, Raleigh, NC 27699-6307 / Phone: (919) 807-3759 / Fax: (919) 807-4065

Visit us on the Web: www.ncpublicschools.org

M0518



Teacher Compensation Models and Advanced Teaching Roles Pilot Programs

Year 1 (2017-18) Preliminary Report

*D. T. Stallings, Robert Maser, and Trisha D. Steinbrecher
The William and Ida Friday Institute for Educational Innovation
North Carolina State University*

May 2018

A Note on the Timing of this Report

The Friday Institute for Educational Innovation was contracted by the North Carolina Department of Public Instruction to conduct the evaluation of the Teacher Compensation Models and Advanced Teaching Roles Pilots on October 12, 2017. Since the Friday Institute was not able to conduct any work on this evaluation ahead of the awarding of that contract, the Scope of Work did not include completion of a Year 1 preliminary report for Fall 2017. The first four contracted deliverables and/or benchmarks were scheduled for submission between December 2017 and April 2018; the current report is a summary of work completed to meet requirements for those deliverables/benchmarks.

Table of Contents

Executive Summary	1
Introduction and Background	3
Overview	3
Legislatively-Prescribed Goals for the Pilot Programs	3
Participation and Support	3
Revised List of Evaluation Questions and Data Sources	5
Evaluation Schedule	9
Year 1 (November 2017-June 2018) – Planning and Baseline Data Collection	9
Year 2 (July 2018-June 2019) – Review of Implementation and Early Results	9
Year 3 (July 2019-June 2020) – Mid-Pilot Results and Initial Recommendations	10
Projected Deliverables	10
Evaluation Update: Prospects for Quantitative Estimations of Pilot Program Impacts	12
Similarities and Differences across Pilot Programs	12
Quantitative Analysis Projections	13
Targets for Quantitative Evaluation	13
Limitations to Quantitative Analysis Options	14
Plans for Quantitative Analyses	15
Appendix A. State-Level and LEA-Level Logic Models	17
Appendix B. Strategic Staffing Theory of Change	24

Executive Summary

Overview

This preliminary report outlines progress made toward the development and execution of the evaluation of the Teacher Compensation Models and Advanced Teaching Roles Pilot Programs created in response to North Carolina General Assembly Session Law 2016-94, Section 8.7. Six Local Education Agency (LEA) proposals were selected by the North Carolina Department of Public Instruction (NCDPI) to participate in the pilot process: Chapel Hill-Carrboro City Schools, Charlotte-Mecklenburg Schools, Edgecombe County Schools, Pitt County Schools, Vance County Schools, and Washington County Schools.

To date, the evaluation team has completed a Year 1 assessment of the status of each pilot's current implementation and has established a plan for conducting limited quantitative outcomes analyses, given time, data availability, and resource challenges.

Even though the pilots are only loosely similar as a result of the experimental nature of the pilots, the evaluation team worked with representatives at NCDPI to develop an overall logic model to represent NCDPI's understanding of how the initiative as a whole ideally contributes to intended outcomes. The evaluation team also worked with each participating LEA to construct LEA-level logic models that reflect the planned actions and intents of each pilot.

Targets for Quantitative Evaluation

For initiatives like this one, there is high demand for a defensible, quantitative assessment of its direct impact on teacher and student outcomes. Potential quantitative indicators include cross-year comparisons of:

- Student proficiency (as measured at the subject level by EoG and EoC results and at the school level by overall School Performance Scores);
- Percentage of students performing at grade level in each subject (as measured by subject-level EoG and EoC scores);
- Teacher EVAAS values;
- Teacher retention/attrition figures; and
- Changes in overall school/LEA teacher quality (as measured by years of experience, EVAAS outcomes).

There are, however, at least four factors that significantly limit the scope of a meaningful quantitative analysis: 1) The small number of teachers impacted; 2) the differences across LEA proposals; 3) the lack of randomization in the identification of impacted teachers and students; and 4) the pilot timeline. The evaluation team hosted a quantitative analysis summit in April 2018 with independent experts to discuss reasonable quantitative options for this evaluation, given these limitations. Based on that consultation, the evaluation team developed a plan for the quantitative analysis component of the evaluation that includes the following elements:

1. *A Combined-Data Approach.* Summit participants recommended ***cross-LEA combined-data analyses*** as the primary analyses. Each LEA's initiative is different, so this type of quantitative evaluation places the focus on the impact of the *presence* of staffing and compensation flexibility—*not* on the impact of a specific implementation format.¹
2. *A School-Level Lens.* Summit participants also recommended focusing on ***school-level rather than individual teacher- and student-level outcomes***, for two reasons: 1) the wide array of teacher roles makes analyses at the teacher level less reliable; and 2) the number of teachers impacted in several of the participating LEAs (and thus the number of students) is very low. By measuring the impact of an LEA's plan on an entire school, rather than on an individual teacher's or student's performance, this approach focuses on overall change in school-level culture.
3. *A Focus on Teachers and Teaching as a Career.* Finally, summit participants recommended a focus on ***teachers and the teaching career***, rather than on short-term student outcomes, which—as suggested by past research and by the relatively short length of the pilots—are not likely to change significantly as a result of the pilots.²

¹ Based on similarities in their pilot designs, the evaluation team also will attempt to conduct analyses of combined data from a sub-set of LEAs (Edgecombe, Charlotte-Mecklenburg, Pitt, and Vance) in an attempt to provide additional insights about the impacts of one approach to implementation that is being applied across multiple locations. The team will attempt to analyze outcomes for each LEA separately, but only when numbers allow.

² Any changes in student outcomes over the course of the pilot period will be noted but are not likely to be firmly attributable to the presence of the pilot programs.

Introduction and Background

Overview

This preliminary report outlines progress made toward the development and execution of the evaluation of the Teacher Compensation Models and Advanced Teaching Roles Pilot Programs created in response to North Carolina General Assembly Session Law 2016-94, Section 8.7.

To date, the evaluation team has completed a Year 1 assessment of the status of each pilot's current implementation and has established a plan for conducting limited quantitative outcomes analyses, given time, data availability, and resource challenges.

Legislatively-Prescribed Goals for the Pilot Programs

Per Section 8.7(a) of the enacting legislation, the intent of the pilot programs is to:

1. Allow highly effective classroom teachers to teach an increased number of students by assuming accountability for additional students, by becoming a lead classroom teacher accountable for the student performance of all of the students taught by teachers on that lead classroom teacher's team, or by leading a larger effort in the school to implement new instructional models to improve school-wide performance.
2. Enable local school administrative units to provide salary supplements to classroom teachers in advanced teaching roles. Selection of an advanced teaching role classroom teacher and award of related salary supplements shall be made on the basis of demonstrated effectiveness and additional responsibilities.
3. Enable local school administrative units to create innovative compensation models that focus on classroom teacher professional growth and student outcomes.
4. Utilize local plans to establish organizational changes related to compensation in order to sustain evidenced-based teaching practices that have the capacity to be replicated throughout the State.

Participation and Support

The legislation outlines a plan that includes implementation of three-year pilots, to begin with the 2017-18 school year and conclude with the 2019-20 school year. Six Local Education Agency (LEA) proposals were selected by the North Carolina Department of Public Instruction (NCDPI): Chapel Hill-Carrboro City Schools, Charlotte-Mecklenburg Schools, Edgecombe County Schools, Pitt County Schools, Vance County Schools, and Washington County Schools.³

The initial allocation for the 2017-18 fiscal year was \$7,180,000, with an additional \$3 million (\$1 million recurring for three years, 2017-18 through 2019-20) distributed among the three largest LEAs each year of the pilot. The total appropriations for the three-year pilot program are

³ Twelve LEAs submitted proposals, all of which can be found here: <http://www.ncpublicschools.org/district-humanresources/>

\$10,180,000.⁴ The disbursement of funds across the six accepted pilot programs is detailed in Table 1.

Table 1. Distribution of State-Provided Funding for Pilots

LEA	Total Project Budget	Recommended Funding	Annual Recurring Funding			Total Funding
			2017-18	2018-19	2019-20	
Charlotte-Meck.	\$ 2,645,131	\$ 1,947,995	\$257,477	\$257,477	\$182,182	\$2,645,131
Pitt	\$ 4,810,169	\$ 2,161,613	\$492,596	\$492,596	\$542,547	\$3,689,352
Chapel Hill-Carrboro	\$ 2,258,952	\$ 1,096,732	\$249,927	\$249,927	\$275,271	\$1,871,857
Vance	\$ 898,000	\$ 898,000	NA	NA	NA	\$ 898,000
Edgecombe	\$ 1,002,210	\$ 943,480	NA	NA	NA	\$ 943,480
Washington	\$ 132,180	\$ 132,180	NA	NA	NA	\$ 132,180
Total	\$11,746,642	\$ 7,180,000	\$ 1,000,000	\$ 1,000,000	\$ 1,000,000	\$10,180,000

⁴ Contingent upon \$1 million recurring budget appropriations through the 2019-2020 fiscal year.

Revised List of Evaluation Questions and Data Sources

The evaluation questions that guide the evaluation (Table 2) correspond to questions posed in the authorizing legislation, as well as questions of interest to the implementing agency (NCDPI). The list of evaluation questions has been revised and re-organized from the draft list presented in the evaluation team's original proposal to better reflect information collected about each LEA's pilot program as a result of the first stages of the evaluation work. The list also now includes the measurable outcomes and indicators associated with each evaluation question.

Table 2. Evaluation Questions and Proposed Data Source(s)

Evaluation Question	Measurable Outcome	Indicator	Data Sources
<i>Q1. Do advanced teaching roles improve the quality of classroom instruction? [NCGA]</i>	Teachers ⁵ demonstrate quality classroom instruction. Teachers exhibit more VA growth than a) teachers at other matched local (same-district) or nearby (comparable neighbor district) schools and b) statewide growth averages.	Student proficiency School performance score School percentage of students performing at grade level in each subject Changes in teacher EVAAS	EVAAS Rubric-based observation data (collected by LEAs); teacher and principal focus groups and/or surveys
	Students exhibit increased interest and engagement in class.	Students report increased interest in class Teachers report increased student engagement	Survey Interview/Focus Group
<i>Q2. Do advanced teaching roles increase school-wide student growth? [NCGA]</i>	Students demonstrate greater academic growth relative to pre-initiative period.	Within-school, cross-year comparisons of: <ul style="list-style-type: none"> • Student growth • School growth score • School report card grade 	North Carolina administrative Data
	Students exhibit more growth than a) students at other matched local (same-district) or nearby (comparable neighbor district) schools and b) statewide growth averages.	Student growth measured against students in matched schools in the LEA or region Student growth measured against all other schools statewide	North Carolina administrative Data School-level annual report

⁵ Note: The evaluation team also will attempt to measure lead teacher and non-lead teacher performance changes separately, to determine changes in either group (as opposed to just changes in the overall group), as time, data, and funding allow.

Evaluation Question	Measurable Outcome	Indicator	Data Sources
<i>Q3. Do advanced teaching roles and/or related local-level salary supplements, either collectively or individually, increase attractiveness of the teaching profession? [NCGA]</i>	Teachers apply for and fill advanced roles	Teacher application and recruitment figures Teachers attribute attractiveness of the teaching profession (in part or in whole) to initiative	Extant administrative data New teacher surveys and focus groups; teacher preparation program surveys
	Lead teachers remain in advanced roles	Lead Teacher retention /attrition figures	Extant administrative data Teacher and principal focus groups and/or surveys
<i>Q4. Do the pilot programs provide recognition to high-quality classroom teachers? [NCGA]</i>	Schools/LEAs provide role-based incentives for lead teachers	Program incentives <ul style="list-style-type: none"> Financial Job-related (e.g., mentoring, PD facilitator, etc.) 	Pilot program theories of action/logic models/incentive schedules Teacher and principal focus groups and/or surveys
	Schools/LEAs recruit and hire/reassign high-quality teachers for advanced roles	Initiative recruitment/recognition plan Lead teacher recruitment/candidate pool quality measures (e.g., GPA, years of experience, prior EVAAS scores, etc.) compared to other teachers in school/LEA	Logic models Extant administrative data Teacher and principal focus groups and/or surveys
<i>Q5. Do the pilot programs support retention of high-quality classroom teachers? [NCGA]</i>	Programs sustain advanced positions	Program funding allocation and sustainability plans Number and type of advanced roles available to teachers each year	Pilot program theories of action/logic models Extant administrative data Teacher and principal focus groups and/or surveys
	Overall increase in high-quality teachers at school/LEA level	Change in overall school/LEA teacher quality (as measured by years of experience, EVAAS outcomes) over time	Extant administrative data

Evaluation Question	Measurable Outcome	Indicator	Data Sources
<i>Q6. Do the pilot programs provide assistance to and support retention of beginning classroom teachers? [NCGA]</i>	Lead teachers support new/beginning teachers (e.g., mentor, planning, model strategies, etc.)	Lead teachers/ administrators report provision of support to new teachers New teachers report receiving adequate support from lead teachers Lead teacher evaluations identify practices/actions that support beginning teachers	Pilot program theories of action/logic models Extant administrative data Teacher and principal focus groups and/or surveys
	New/beginning teachers remain in pilot school/LEA	New teacher retention/ attrition figures New teachers indicate a desire to continue teaching (short and/or long term)	Extant administrative data Teacher and principal focus groups and/or surveys
<i>Q7. In what other ways do these pilot programs impact high-quality experienced classroom teachers? [NCGA]</i>	Other unanticipated/ untracked program impacts (direct and indirect)	Teacher perceptions of impact related to the program Principal perceptions of impact related to the program	Teacher and principal focus groups and/or surveys
<i>Q8. What do the proposed pilot programs have in common? What are each pilot program's unique components?</i>	Participating LEAs and evaluation team complete state-level and program-specific logic models	Descriptions of program models, intended impact, and fidelity of implementation; Unique program elements highlighted	Pilot program theories of action/logic models; descriptions of similar or related prior initiatives
<i>Q9. As measured by the quantitative and qualitative outcomes of interest described above, which pilot program or programs appear to be the most successful?</i>	Measurable outcomes for Q1 through Q7 -- individually or collectively -- indicate successful outcomes for a specific pilot model or models	Comparative assessment of qualitative and quantitative results for Q1 through Q7	All data gathered and results generated for evaluation questions described above
<i>Q10. Which pilot programs appear to be most scalable? What resources would the state need to commit in order to successfully scale them?</i> <i>Q10a. Should the state consider scaling one or</i>	Program sustainability measured by cost (and availability) of resources to maintain roles and salary supplements ⁶	LEA projections for fiscal sustainability after pilot period (cost)	Extant state and local fiscal data

⁶ A rigorous benefits-costs analysis or cost-effectiveness analysis is not feasible on the current pilot timeline and evaluation budget

Evaluation Question	Measurable Outcome	Indicator	Data Sources
<i>more of the pilot programs?</i>	Individual successful program components identified for Q9 show evidence of scalability to other LEAs Overall successful pilot program(s) identified for Q9 show evidence of scalability to other LEAs	Pilot program components are not place-dependent (i.e., they do not require locale-specific inputs, can be adapted across LEA contexts) (flexibility) Via survey and focus groups, implementers indicate ease of implementation (minimum LEA capacity requirements)	All data gathered and results generated for Q1 through Q9 Extant state and local fiscal data
<i>Q11. What are the costs and benefits associated with establishing advanced teaching roles? To what extent does the return on investment in establishing new compensation models that correspond with these roles (as measured by the outcomes of interest described above) justify the investment?</i>	Teachers and administrators express support for continuing the pilot	Trends in teacher survey responses over pilot period Trends in administrator survey responses over pilot period Trends in teacher focus group responses over pilot period Trends in administrator focus group responses over pilot period	All data gathered and results generated for evaluation questions described above

Evaluation Schedule

Year 1 (November 2017-June 2018) – Planning and Baseline Data Collection

- Work with NCDPI to establish an initiative-level logic model, as well as working definitions for key terms (e.g., “high-quality teacher” and “experienced teacher”)
- Work with each participating LEA to establish an LEA-specific logic model for each pilot program that identifies how the pilot program is designed to meet the legislative intentions outlined above, as well as how the LEA thinks the pilot program will address LEA goals and intentions
 - Develop objective descriptions of each pilot program
 - Identify commonalities across programs with the potential to allow for larger-scale (cross-LEA) quantitative analyses
- Identify examples of similar programs in other North Carolina LEAs (and possibly nationally as well), both current and historical
- If possible (depending on pilot program structures), identify potential comparison schools and LEAs
- Develop all instruments for the evaluation (focus group protocols, surveys)
 - Via survey and focus groups, conduct initial qualitative data collection rounds in each LEA and at each licensure program in the state (to determine perceptions of pilot program appeal among beginning teachers)
- Determine feasibility for any rigorous quantitative outcomes studies; feasibility will depend on: a) numbers of teachers, students impacted by the initiatives; b) similarities of initiatives across LEAs (to increase numbers available for analyses); and/or c) absence of similar initiatives in LEAs before introduction of current pilots

Year 2 (July 2018-June 2019) – Review of Implementation and Early Results

- Continue carry-over work from Year One (quantitative and qualitative data collection)
- Assess Program components (recruitment, selection, program content, program delivery, etc.)
- Begin recording beginning teacher and experienced teacher retention rates
- Begin comparing historical and current teacher vacancy application rates in participating LEAs and in nearest-neighbor LEAs
- Analyze short-term (one years out) qualitative outcomes for participating teachers, schools, and students
- Retrieve, prepare, and analyze administrative and accounting data for first year of implementation (anticipated initial data availability: November 2018) for use in quantitative analysis

Year 3 (July 2019-June 2020) – Mid-Pilot Results and Initial Recommendations

- Continue work from Year Two (quantitative and qualitative data collection; formative program assessment)
- Analyze and interpret mid-term (two years out) qualitative outcomes for participating teachers, schools, and students
- To the extent possible,⁷ analyze and interpret final year (three years out) qualitative data
- Estimate costs for and feasibility of scaling the pilots that appear to be experiencing the most success
- Retrieve, prepare, and analyze data for second year of implementation (anticipated initial data availability: November 2019)
- Conduct quantitative data analyses for SY 2017-18 and 2018-19⁸ for inclusion alongside qualitative data analyses

Projected Deliverables

Table 3 outlines the projected deliverables for this evaluation, as well as the evaluation question(s) each will answer and the estimated delivery date.

Table 3. Evaluation Deliverables

Deliverable	Anticipated Evaluation Question(s) Addressed	Estimated Delivery Date
Updated evaluation plan – including presentation to State Board of Education	N/A	<i>December 2017 (summarized in this report)</i>
Initial overall state logic model	Will inform EQ 4, 6, and 8	<i>February 2018 (summarized in this report)</i>
Initial individual local logic models	Will inform EQ 4, 6, and 8	
Data collection instruments (surveys, focus group protocols)	Will inform EQ 1, and 3 through 11	
Brief: Formal preliminary assessment of quantitative evaluation possibilities	EQ 2, 5, and 6	<i>April-May 2018 (summarized in this report)</i>
Initial formative evaluation report for State Board	EQ 1, and 3 through 9	<i>October 2018</i>
Initial quantitative outcomes report for State Board	EQ 2, 5, and 6	<i>March 2019</i>

⁷ Note: Because funding for the evaluation ends in June 2020, the evaluation team will be able to analyze only some of the Year 3 qualitative data (data collected by or before February 2020).

⁸ Note: Funding for the evaluation ends before data for quantitative analyses of Year 3 will be available (November 2020).

Deliverable	Anticipated Evaluation Question(s) Addressed	Estimated Delivery Date
Second formative evaluation report for State Board	EQ 1, and 3 through 9	<i>October 2019</i>
Two-year qualitative and quantitative findings report for State Board (includes identification of most promising pilots, plus estimates of costs to scale those pilots)	All EQs	<i>May-June 2020</i>

Evaluation Update: Prospects for Quantitative Estimations of Pilot Program Impacts

Because most of the pilot programs are in their first year of operation, the evaluation team anticipates that there will be changes in implementation between this year and next that will impact the final form of each pilot. Therefore, the evaluation team will delay presentation of in-depth descriptions of each pilot program until the initial formative evaluation report (October 2018). In the meantime, we present here information about the emerging similarities and differences across pilots and how they inform the quantitative components of the evaluation plan.

Similarities and Differences across Pilot Programs

Even though the pilots are only loosely similar as a result of the experimental nature of the pilots, the evaluation team worked with representatives at NCDPI to develop an overall logic model to represent NCDPI's understanding of how the initiative as a whole ideally contributes to intended outcomes. The evaluation team also worked with each participating LEA to construct LEA-level logic models that reflect the planned actions and intents of each pilot.

All logic models are included in Appendix A. They aided in the development of data collection tools for the evaluation—including development of the quantitative models used to estimate numerically measurable impacts of the pilots. In addition, they also informed the development of a matrix that summarizes the major points of comparison across programs (Table 4).

Table 4. Common Pilot Program Features

	CHCCS	CMS	Edge-combe	Pitt	Vance	Washington
Components						
Professional devel.	✓	✓	✓	✓	✓	✓
Variable class sizes		✓	✓	✓	✓	
Teacher teams		✓	✓	✓	✓	
Teacher-Leader Roles⁹						
PD facilitator ¹⁰	✓			✓		✓
Coach ¹¹		✓	✓	✓	✓	✓
Co-teacher ¹²		✓	✓	✓	✓	
Mentor ¹³	✓					
Team leader ¹⁴		✓	✓	✓	✓	

⁹ Roles are not the same as *position titles*; the roles in this table are those identified by the evaluation team as being covered by one or more positions across LEAs plans—regardless of an LEA's title for the person who takes on a given role. Corresponding positions in each LEA are identified in footnotes.

¹⁰ CHCCS=PD Facilitator; Pitt=Facilitating, Multi-Classroom Teacher; WCS=Master Teacher

¹¹ CMS=Multi-Classroom Teacher; ECPS=Expanded Impact Teacher, Multi-Classroom Teacher; Pitt=Facilitating, Multi-Classroom Teacher; VCS=Multi-Classroom Teacher; WCS=Master Teacher;

¹² ECPS=Expanded Impact Teacher, Multi-Classroom Teacher, Reach Associate; CMS=Multi-Classroom Teacher, Reach Teachers; Pitt=Facilitating, Multi-Classroom Teacher; VCS=Expanded Impact Teacher, Multi-Classroom Teacher, Reach Associate

¹³ CHCCS=Mentor Teacher

¹⁴ CMS, ECPS, Pitt, VCS=Multi-Classroom Teachers

As indicated by the table and the logic models, there are differences in each LEA's implementation (as would be expected with independently-conceptualized and -run pilots); however, several of the pilots share at least a few components and roles in common (though sometimes only in nominal ways; differences exist in each LEA's implementation of each component and role)—in part because three LEAs are working with or have worked with a common third-party support provider.¹⁵ These commonalities are important to keep in mind when reviewing the planned quantitative component of the evaluation.

Quantitative Analysis Projections

Targets for Quantitative Evaluation

For initiatives like this one, there is high demand for a defensible, quantitative assessment of the direct impact of the initiative on teacher and student outcomes. As indicated in Table 2 (above), there are several evaluation questions that would benefit from quantitative analyses alongside the qualitative components of the evaluation; those questions and their relevant quantitative indicators are revisited in Table 5.

Table 5. Non-Financial Evaluation Questions with Potential Quantitative Analysis Components

Evaluation Question	Quantitatively Measurable Outcome	Quantitative Indicator
<i>Q1. Do advanced teaching roles improve the quality of classroom instruction?</i>	Teachers demonstrate quality classroom instruction. Teachers exhibit more VA growth than a) teachers at other matched local (same-district) or nearby (comparable neighbor district) schools and b) statewide growth averages.	Student proficiency (EoG, EoC) School performance score (80% achievement; 20% growth) School percentage of students performing at grade level in each subject (EoG, EoC) Teacher EVAAS
<i>Q2. Do advanced teaching roles increase school-wide student growth?</i>	Students demonstrate greater academic growth relative to pre-initiative period.	Within-school, cross-year comparisons of: <ul style="list-style-type: none"> ● Student growth ● School growth score ● School report card grade
	Students exhibit more growth than a) students at other matched local (same-district) or nearby (comparable neighbor district) schools and b) statewide growth averages.	Student growth measured against students in matched schools in the LEA or region Student growth measured against all other schools statewide

¹⁵ Public Impact (<http://publicimpact.com/>), which promotes an advanced teaching roles model called Opportunity Culture, is working with Edgecombe and Vance on their pilots, and they formerly worked with Charlotte-Mecklenburg on an earlier iteration of their model. Each LEA is working with at least one additional support provider, but only Public Impact has worked across multiple LEAs.

Evaluation Question	Quantitatively Measurable Outcome	Quantitative Indicator
<i>Q3. Do advanced teaching roles and/or related local-level salary supplements, either collectively or individually, increase attractiveness of the teaching profession?</i>	Lead teachers remain in advanced roles	Lead Teacher retention/attrition figures
<i>Q5. Do the pilot programs support retention of high-quality classroom teachers?</i>	Overall increase in high-quality teachers at school/LEA level	Change in overall school/LEA teacher quality (as measured by years of experience, EVAAS outcomes) over time

Limitations to Quantitative Analysis Options

A key purpose of the overall evaluation is to attempt to analyze as many of these quantitatively-measurable outcomes as is feasible. There are however, at least four factors that significantly limit the development of a meaningful quantitative analysis strategy: 1) The small number of teachers impacted; 2) the differences across LEA proposals; 3) the lack of randomization in the identification of impacted teachers and students; and 4) the pilot timeline.

1. Size of Impacted Teacher Population

The evaluation can report raw changes in teacher behavior (e.g., attrition rates) and teacher quality (e.g., via formal teacher evaluations), but, without sufficient numbers of impacted teachers, these measures may not be suitable for rigorous statistical analysis. This problem is confounded by the second issue.

2. Differences across LEA Proposals

As indicated by each pilot program's logic model (Appendix A) and the table of common program features across LEAs (Table 4), each pilot program is different from the others in certain ways, and many are very different—enough so that combining data from multiple pilots (allowing for stronger analyses of impacts on larger groups of participants and impacted students) in order to overcome the first limitation must be done with caution and with a full explanation of the caveats that apply to all results.

3. Randomization

Very often, programs like these pilots depend upon either teacher participant volunteers or teacher assignment to program participation based on one or more preconditions, or both. In most cases, students are not randomly assigned to the teachers who participate. None of these factors prevents evaluators from determining statistically significant *correlations* between program initiative components and certain outcomes, but all of them do prevent evaluators from determining *causation*.

4. *Lifespan of the Pilots*

Finally, at its heart, a differentiated pay/advanced roles plan, no matter how it is implemented, is about changing school culture for the long term. The evaluation team knows from studies of the impact of changing even just one school culture variable (for example, changing principal leadership) that schools often experience a regression in outcomes for at least a year before even highly successful program begin to show positive results. The evaluation of the state's statewide and local Race to the Top experiments with strategic staffing (2010-2014) also suggested that fully-realized impacts of an advanced teacher roles plan often will not materialize for several school years, after preliminary impacts on school culture and teacher turnover have paved the way for later impacts on student performance (a top-level Theory of Change produced as part of this work is included in Appendix B). With only a three-year window of implementation, it will be challenging to detect the potential full effects of any of these pilots, and it is not inconceivable that the initial short-term results will be negative, even if the longer-term prospects are potentially positive.

Plans for Quantitative Analyses

Given the limitations described above, as well as the limits imposed by the evaluation budget, and with acknowledgement of the Theory of Change suggested by the evaluation team's work on similar programs during the state's Race to the Top period, the team hosted a quantitative analysis summit in April 2018 with independent experts to discuss reasonable quantitative options for this evaluation. Based on that consultation, we developed the following plan for the quantitative analysis component of the evaluation.

A Combined-Data Approach. With acknowledgement of the caveats above, summit participants recommended ***cross-LEA combined-data analyses*** as the primary analyses. As noted above in the list of limitations, each LEA's initiative is different, so this type of quantitative evaluation places the focus on the impact of the *presence* of staffing and compensation flexibility—not on the impact of a specific implementation format.

To partially compensate for that, based on the similarities of their programs as indicated by Table 4, we also will attempt to conduct analyses of combined data from Edgecombe, Charlotte-Mecklenburg, Pitt, and Vance only. Doing so may provide additional insights about the impacts of one approach to implementation that is being applied in similar ways across multiple locations.¹⁶

Finally, we will attempt to analyze outcomes for each LEA separately, but only when numbers allow. Because of the low numbers of teachers and students impacted in some LEAs, these analyses may be possible only for some outcomes in Chapel Hill-Carrboro, Charlotte-Mecklenburg, and Pitt. In all cases in which we conduct LEA-level analyses, we will include qualitative data and findings for context.

¹⁶ Because of the major differences between Chapel Hill-Carrboro's program and the programs in the five other LEAs (e.g., the Chapel Hill-Carrboro program is the only pilot program that does not include school-based leadership roles; to be explored in greater detail in the October 2018 report), the evaluation team also may run separate analyses that exclude only Chapel Hill-Carrboro data.

A School-Level Lens. Summit participants also recommended focusing on *school-level rather than individual teacher- and student-level outcomes*, for two reasons: 1) the wide array of teacher roles makes analyses at the teacher level less reliable; and 2) the number of teachers impacted in several of the participating LEAs (and thus the number of students) is very low. In keeping with the idea of whole-school cultural change, this approach focuses on the impact of an LEA's plan on an entire school, rather than on its impact on an individual teacher's or student's performance. Indicators to be measured at the school level are outlined in the **Quantitative Indicator** column in Table 5 (above).

In the case of one participating LEA (Pitt County¹⁷), the evaluation team may be able to conduct school-level quantitative analysis of the same indicators that compare impacted versus non-impacted schools within an LEA. In all other cases, the team will analyze quantitative differences between matched schools in similar participating and non-participating districts. Some factors that may aid in the matching process include historical and current Teacher Working Conditions survey results, student proficiency and growth results, school size, school location, and school demographics.

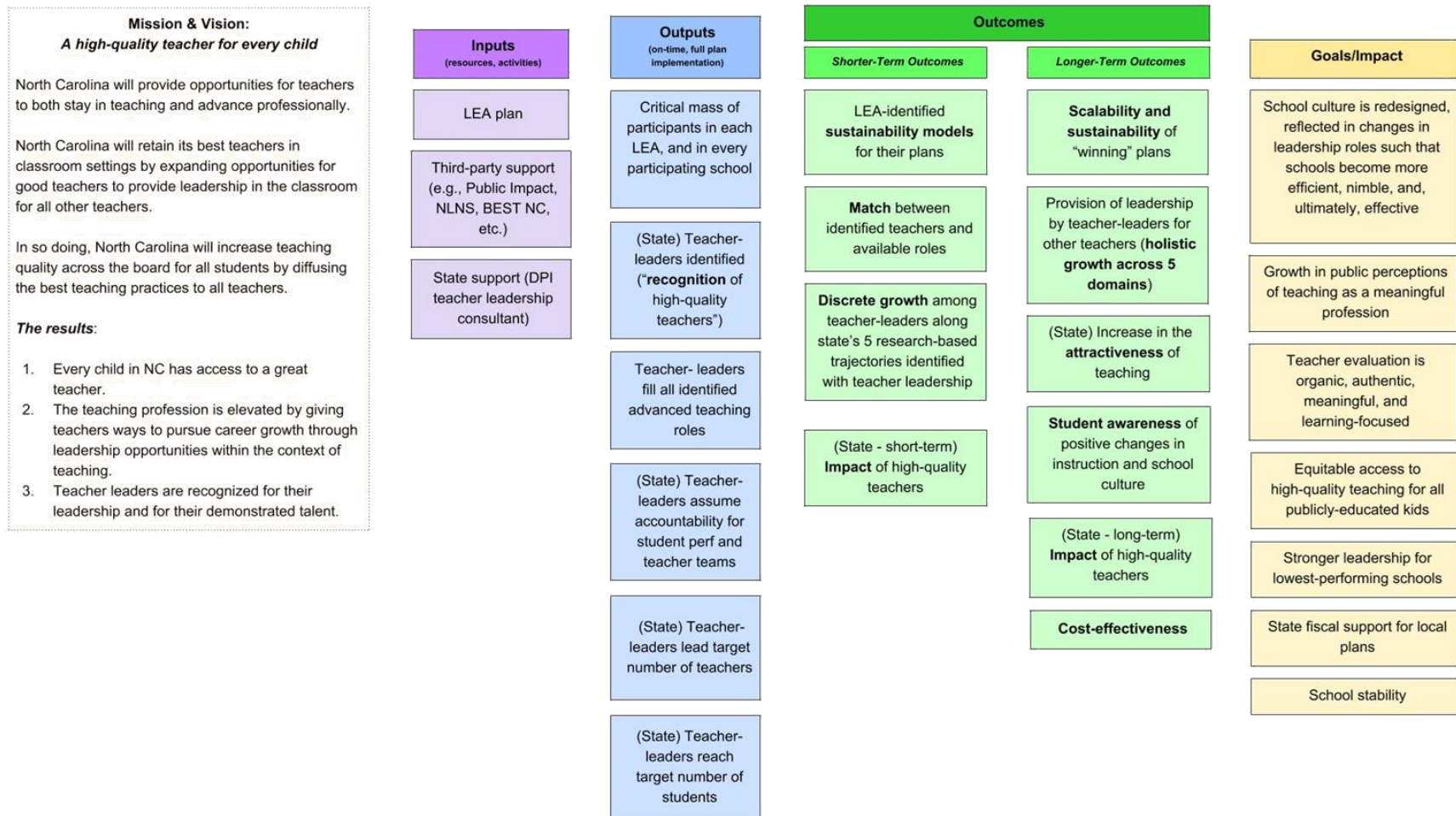
A Focus on Teachers and Teaching as a Career. Finally, summit participants recommended a focus on *teachers and the teaching career*, rather than on short-term student outcomes, which—as suggested by the Theory of Change model and by the relatively short length of the pilots—are not likely to change significantly as a result of the pilots (any changes in student outcomes over the course of the pilot period will be noted but are not likely to be firmly attributable to the presence of the pilot programs). When available, we will rely on EVAAS value-added scores to control for some of the challenges posed by each LEA's different approaches to teacher selection.

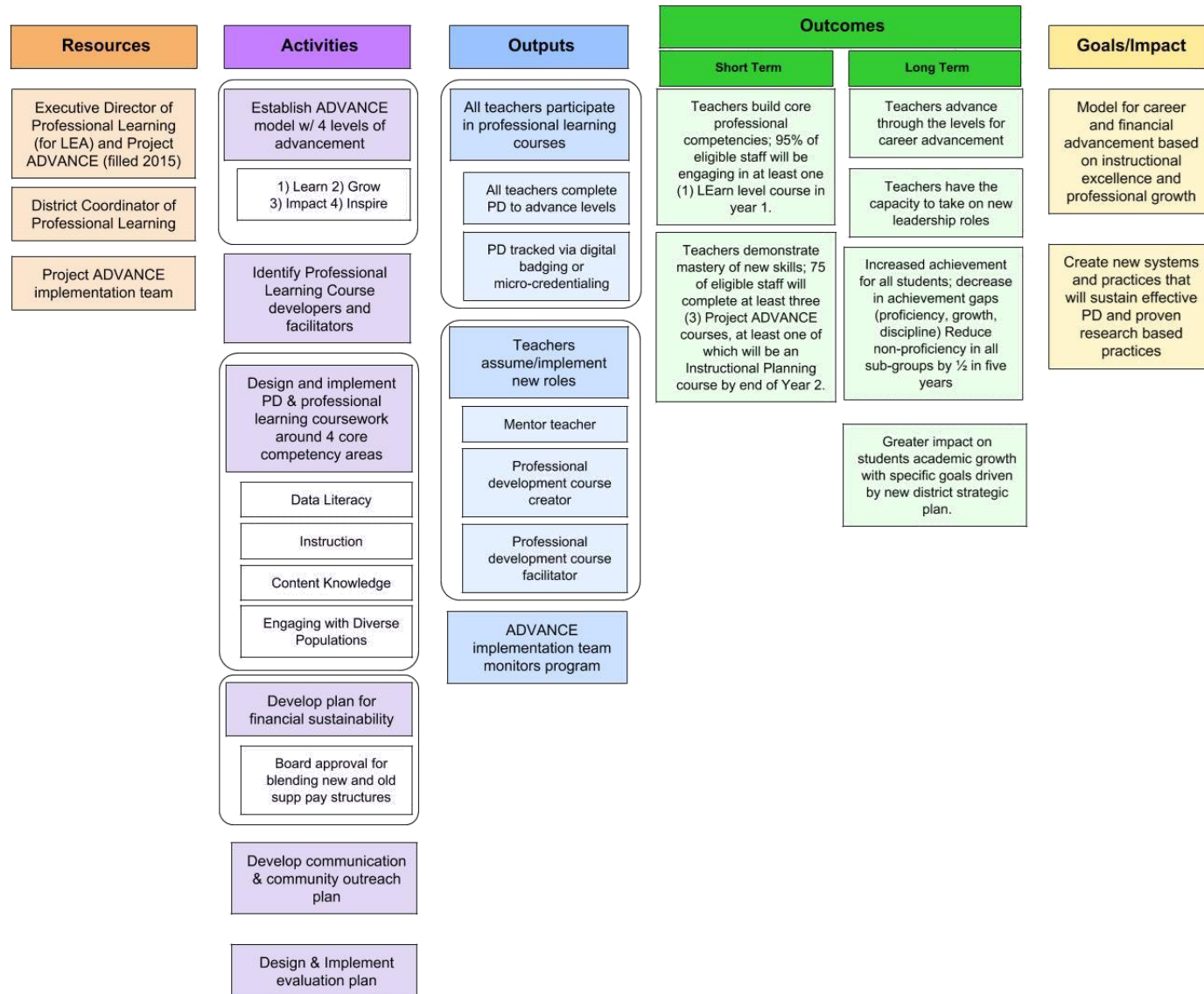
In addition, to gauge the overall appeal of career ladder programs to young professionals, the evaluation team will investigate the possibility of asking teacher licensure candidates in colleges of education to review short descriptions of each program (with a focus on role/position descriptions and salary/bonus schedules) and assess their relative appeal.

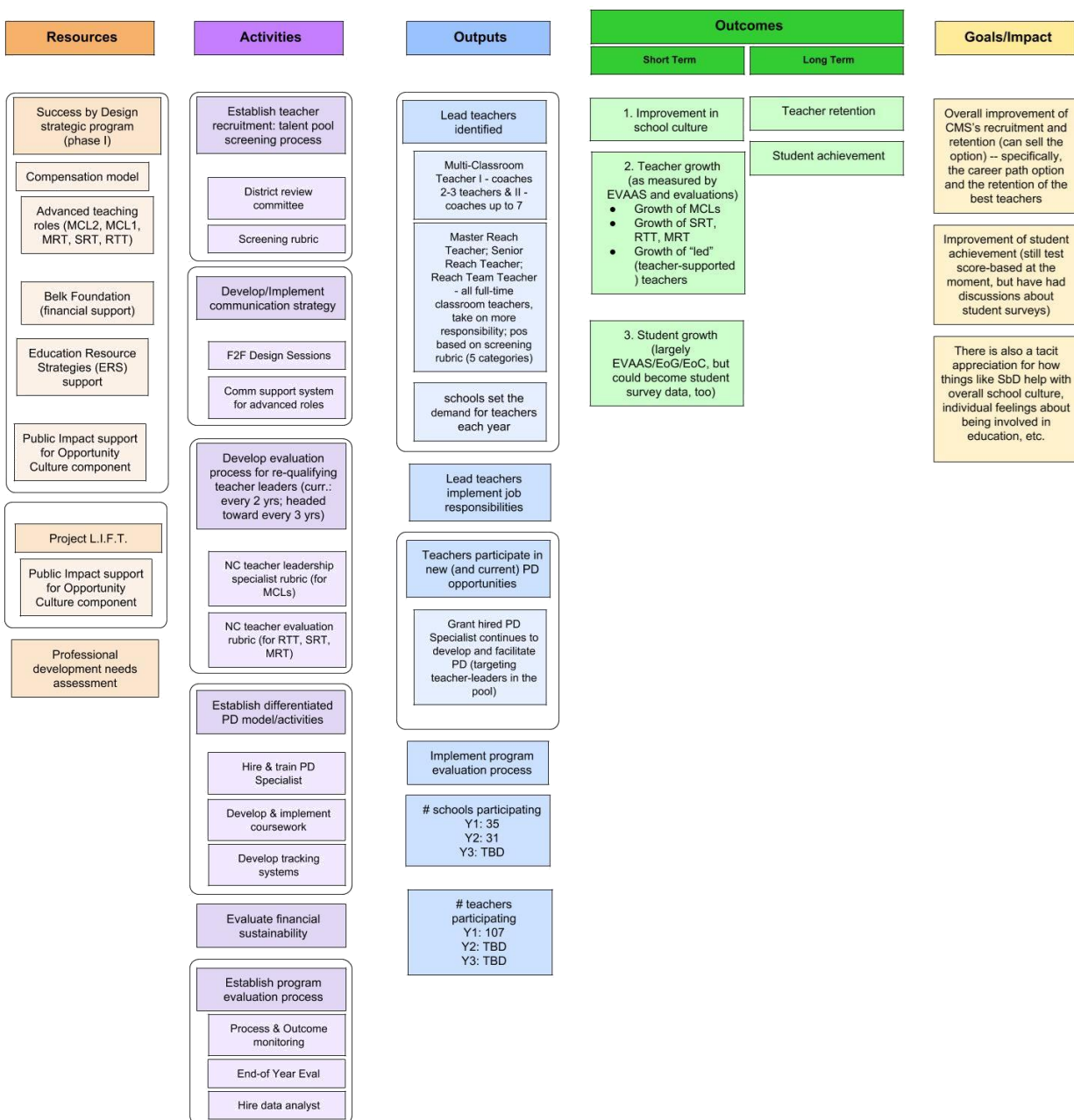
¹⁷ Charlotte-Mecklenburg's long history with differentiated pay and strategic staffing models may make it difficult to find appropriate comparison schools within the LEA that have not been impacted by similar models in the recent past.

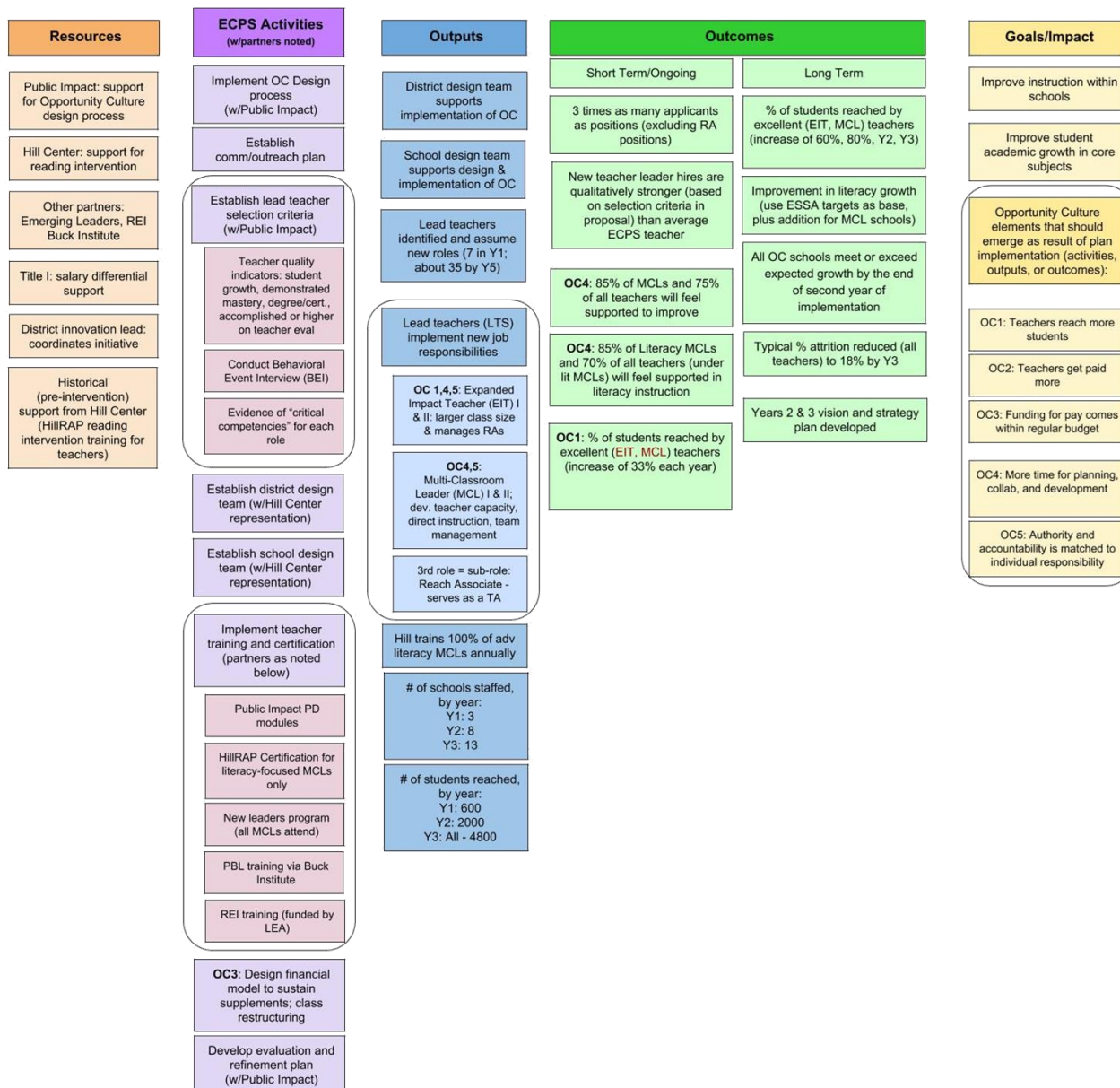
Appendix A. State-Level and LEA-Level Logic Models

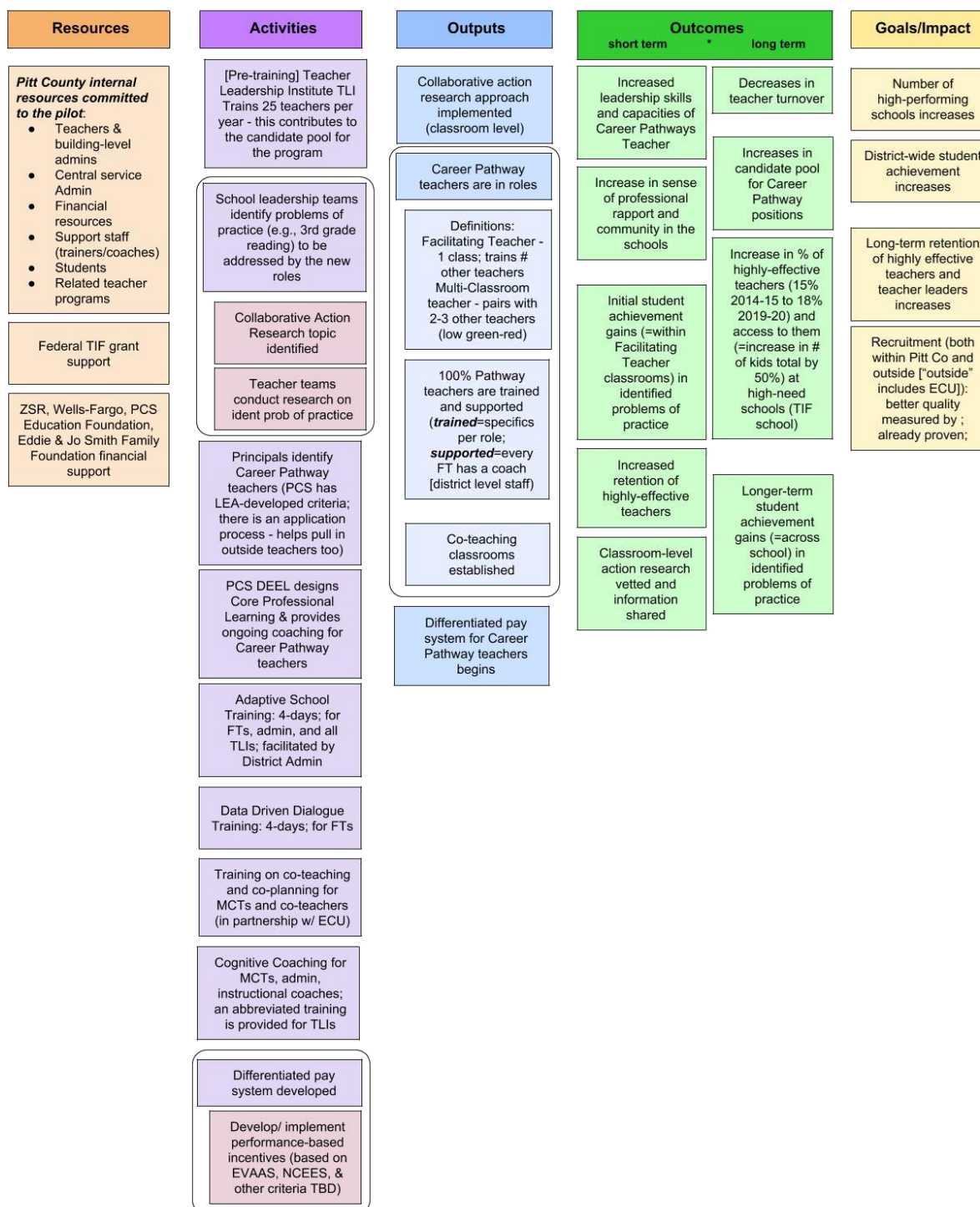
State-Level

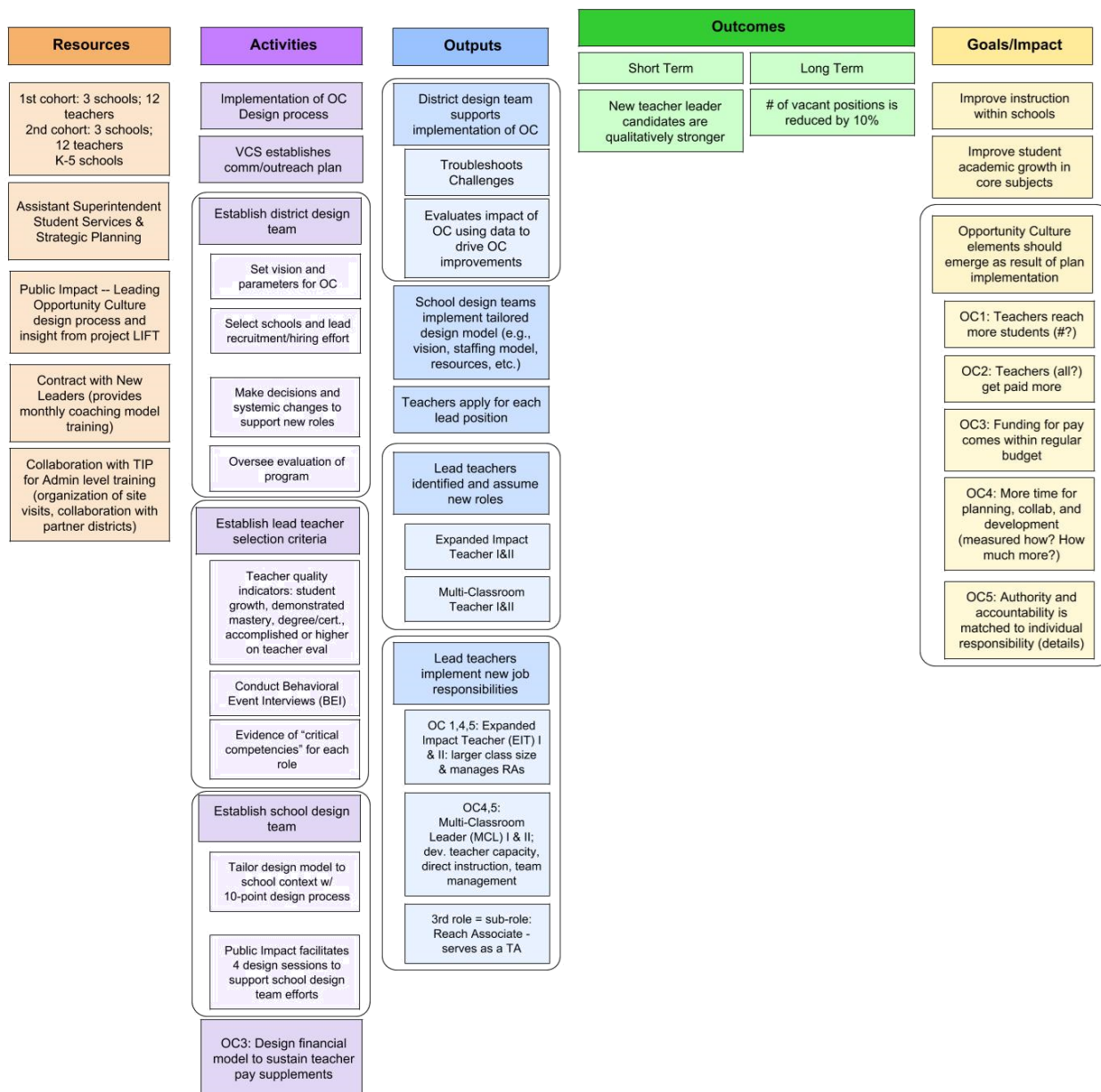


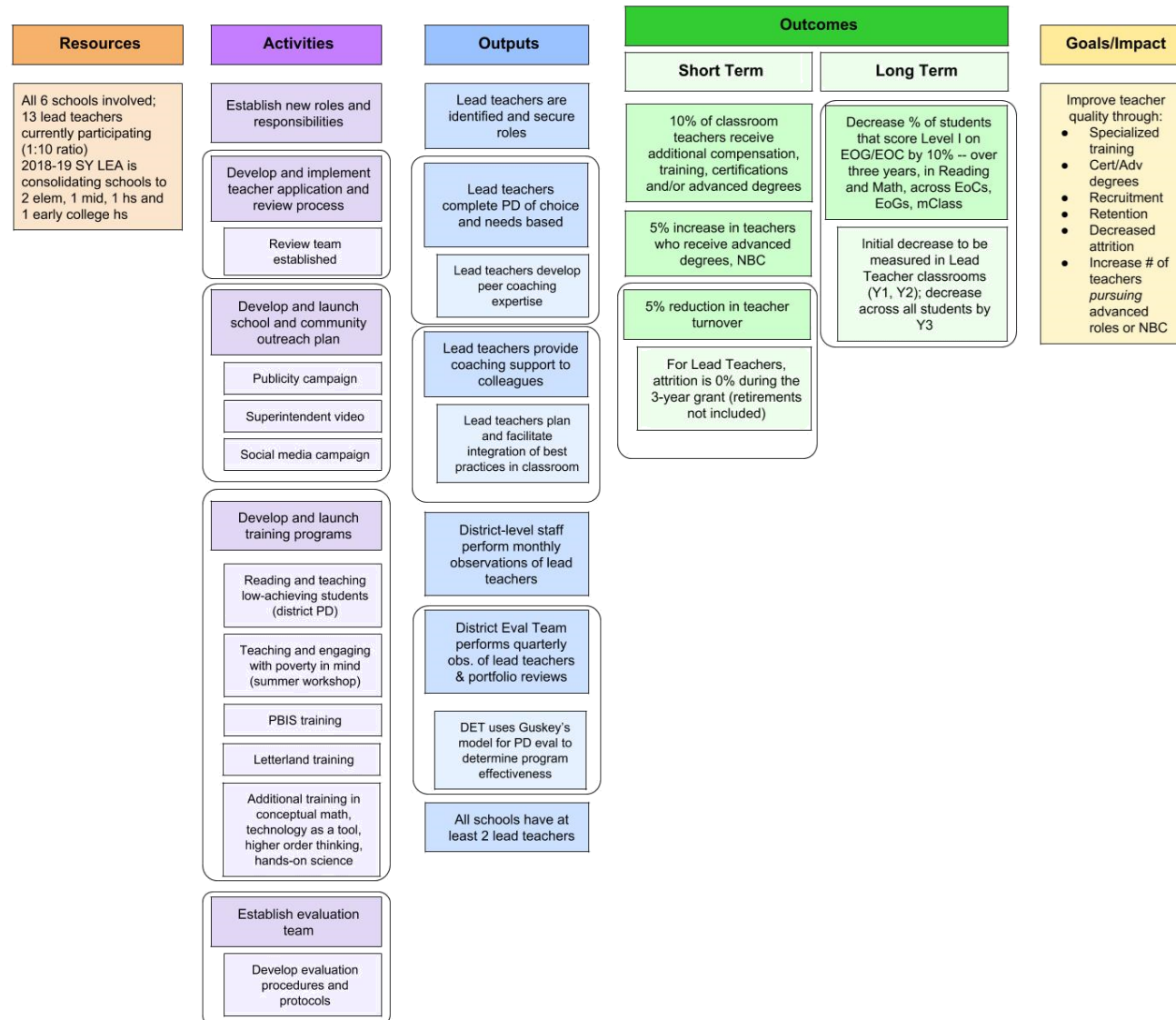
Chapel Hill-Carrboro

Charlotte-Mecklenburg

Edgecombe

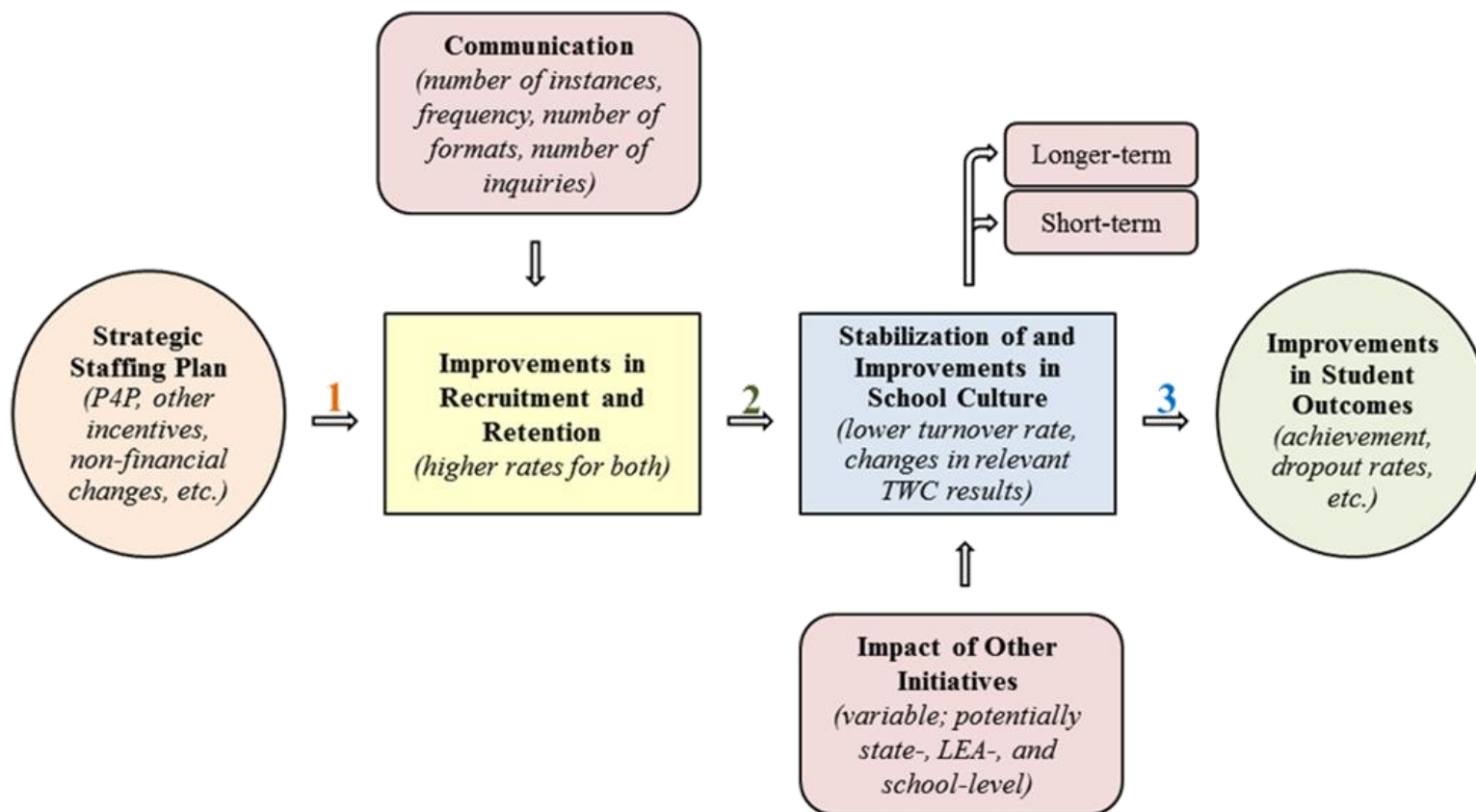
Pitt

Vance


Washington

Appendix B. Strategic Staffing Theory of Change

Evaluations of state- and local-level strategic staffing initiatives funded by North Carolina's Race to the Top grant (2010-2014)¹⁸ all suggested that several intermediate changes needed to occur in a school (e.g., better teacher recruitment and retention, lower teacher turnover rate, etc.) before the existence of a staffing plan would have a measurable impact on student outcomes:



¹⁸ <http://cerenc.org/rttt-evaluation/equitable-supply-and-distribution-of-teachers-and-leaders/>



Contact Information:
Please direct all inquiries to Trip Stallings
dtstalli@ncsu.edu