The Learning Partners
Urban Teacher Residency Project
SUMMATIVE REPORT

Teaching & Learning in Collaboration

HUNTER COLLEGE NEW VISIONS FOR PUBLIC SCHOOLS

THE NEW YORK CITY DEPARTMENT OF EDUCATION

Rockman et al
December 2019

ACKNOWLEDGEMENTS

This report would not have been possible without the help of the LP-UTR partners and participants, who have supported our evaluation efforts throughout the program. We would like to thank the program directors and staff, data teams, coaches, mentors, principals, and residents for their willingness to share their ideas and inform the program evaluation. We would especially like to thank those who helped us assemble the data necessary for our summative study—the New Visions data team, Hunter's Office of Institutional Research, the New York City Department of Education, and the Independent Budget Office. We also drew often on the expertise and experience of LP-UTR partners, including Deputy Directors of Residencies, Marisa Harford and Rachelle Verdier; Sherryl Brown Graves, Acting Senior Associate Dean of Hunter College's School of Education and Erin Shay, Director of the School of Education's Office of Partnership Programs; and Brett Cluff and Elise Corwin and their teams at the New York City Department of Education. Our work benefitted immeasurably from their insights and support.

TABLE OF CONTENTS

Introduction 6
Methods 8
- Project Records and Performance Data 9
- Teacher (Resident and Mentor) Surveys and Samples 9
- Focus Groups and Interviews 11
- Student Data Samples 11

Summary of Findings 15

Part 1: Recruitment & Residents' Performance 19
- LP-UTR Cohort Composition 20
  - Race and Ethnicity 20
  - Educational Background 21
  - Gender Balance 22
- Residents' Program Performance 22
  - Residency Completion Rates 23
  - Residents' Coursework Performance and Certifications 23
  - Job Placement and Retention 25
- Recruitment of LP-UTR Mentors and Schools 26
  - Mentors' Backgrounds and Experience 26
  - LP-UTR Host Schools 28

Part 2: Evolution of the Model & Partners' Roles 29
- The LP-UTR Model and Partners' Roles 30
  - The New York City Department of Education and LP-UTR Schools 31
  - New Visions for Public Schools 31
  - Hunter College 32
- The LP-UTR Coaching Model 34
  - New Visions Coaches 34
  - School-Based Coaches 34
  - District Coaches 34
  - Changes in the Coaching Model 34

Part 3: Growth in Teachers' Skills & Confidence 37
- Performance Assessments of Residents' Growth 38
  - The Suite of Process Monitoring Assessments 39
  - LP-UTR Danielson-Based Rubric 40
  - Defense of Learning 43
- Teachers' Self-Assessments of Confidence 45
  - Residents' and Mentors' Confidence 45
- Where Mentors and Residents Gain Skills and Confidence 49

Part 4: Support, Collaboration, & Collective Efficacy 52
- School Support for Collaborative Activities 53
  - School-Level Support 53
  - Opportunities for Collaboration 54
- Collaborative Activities Across Schools 55
  - Levels of Participation in LP-UTR Activities 55
  - Building Skills and Confidence within and across Schools 57
- Teachers' Collective Efficacy 59

Levels of Confidence 59
New vs. Returning Mentors 60

Part 5: LP-UTR's Impact on Student Achievement 64
Primary Student Achievement Analysis 66
Regents Exam Performances of Students Taught by LP-UTR Graduates 66
Course Credits Earned among Students Taught by LP-UTR Graduates 68
Regents Exam Performances of Students Taught by LP-UTR Mentors 70
Course Credits Earned by Students Taught by LP-UTR Mentors 72
School-wide Achievement Analysis for Students With Disabilities and English Learners 73
Regents Examination Performances of Students with Disabilities 73
Course Credits Earned by Students with Disabilities 74
Regents Examination Performance by English Language Learners 75
Course Credits Earned by English Language Learners 76

Conclusions 78

References 81

Appendix A: Process Monitoring 84

Appendix B: Data Needs & Logic Model 88

Tables and Figures

Table 1. Years 1–3 Survey Respondents by Subject, Baseline 10
Table 2. Years 1–3 Survey Respondents by Subject, End of Year 10
Table 3. Overall Student Achievement Data Samples for LP-UTR Graduates 13
Table 4. Overall Student Achievement Data Samples for Mentor Teachers 13
Table 5. Overall Student Achievement Data Samples for Mentor Teachers, by Subject and Years of Participation 14
Table 6. Number of Residents Enrolled by Subject Area 20
Table 7. Gender Representation, by Cohort 22
Table 8. Program Performance, Residency Enrollment and Completion Rates 23
Table 9. Average GPAs, by Residency Cohort 24
Table 10. Program Performance, Certifications 24
Table 11. Resident Teaching Status in NYC DOE or NV Charter School 25
Table 12. Number of Coaches, by Type 35
Table 13. Key Process Monitoring Assessments, by Cohort 39
Table 14. Attainment of Assessment Benchmarks by Cohort and Time 40
Table 15. LP-UTR Danielson-Based Rubric 41
Table 16. Score Alignment Between the Danielson Framework (2013) and NYS Department of ED 41
Table 17. Domains and Competencies of the Defense of Learning Rubric 44
Table 18. Where Mentors Develop Skills and Gain Confidence 49
Table 19. Where Residents Develop Skills 50
Table 20. Mentors' and Residents' Ratings of Support from Colleagues 53
Table 21. Reports of Participation in LP-UTR 56
Table 22. LP-UTR Mentors' Confidence in Collaborative Practices 60

Table 23. Where Mentors Developed Efficacy Skills 61
Table 24. Where Residents Developed Efficacy Skills 62
Table 25. Regents Exam Performances of Students Taught by Graduates, Cohort 1 67
Table 26. Regents Exam Performances of Students Taught by Graduates, Cohort 2 67
Table 27. Overall Regents Exam Performances, Cohorts 1 and 2 Combined 68
Table 28. Analysis of Graduates Course Credits, Cohort 1 69
Table 29. Analysis of Graduates Course Credits, Cohort 2 69
Table 30. Analysis of Graduates Course Credits, Cohorts 1 and 2 Combined 70
Table 31. Analysis of Regents Scores, Mentors One Year Only 71
Table 32. Analysis of Regents Scores, Mentors Multiple Years 71
Table 33. Analysis of Course Credits, Mentors One Year Only 72
Table 34. Analysis of Course Credits, Mentors Multiple Years 73
Table 35. Analysis of Regents Scores, Special Education Students Years 1-3 Schools 74
Table 36. Analysis of Course Credits, Special Education Students Years 1-3 Schools 75
Table 37. Analysis of Regents Scores, English Language Learners Years 1-3 Schools 76
Table 38. Analysis of Course Credits, Special Education Students Years 1-3 Schools 77
Table 39. Residency Assessment Timeline and Benchmark Targets, Cohort 1 84
Table 40. Residency Assessment Timeline and Benchmark Targets, Cohort 2 84
Table 41. Residency Assessment Timeline and Benchmark Targets, Cohort 3 85
Table 42. LP-UTR Danielson-Based Rubric Data, Cohort 2 85
Table 43. LP-UTR Danielson-Based Rubric Scores by Timepoint, Cohort 2 86
Table 44. LP-UTR Danielson-Based Rubric by Assessor Role, Cohort 2 86
Table 45. LP-UTR Danielson-Based Rubric Data, Cohort 3 86
Table 46. LP-UTR Danielson-Based Rubric Scores by Timepoint, Cohort 3 87
Table 47. LP-UTR Danielson-Based Rubric by Assessor Role, Cohort 3 87
Table 48. Rockman Data Needs, Years 1–3 88

Figure 1. Ethnic Representation of Residents 20
Figure 2. Institutions Attendance by Residents 21
Figure 3. Previous Degrees Held by Residents 21
Figure 4. Previous Work Experience by Residents 22
Figure 5. Mentor Teachers' Preparation Programs 26
Figure 6. Years of Overall Teaching Experience, Mentors 27
Figure 7. Years Teaching at Current School, Mentors 27
Figure 8. LP-UTR School Participation, by Year 28
Figure 9. Average Component-Level Scores on the LP-UTR Danielson-Based Rubric, by Time: Cohort 2 42
Figure 10. Average Component-Level Scores on the LP-UTR Danielson-Based Rubric, by Time: Cohort 3 43
Figure 11. Defense of Learning Scores by Cohort and Time 45
Figure 12. Instructional Practices Confidence 46
Figure 13. Assessment Practices Confidence 47
Figure 14. Instructional Practices Confidence, Changes from Fall to Spring 48
Figure 15. Assessment Practices Confidence, Changes from Fall to Spring 48
Figure 16. Residents: Frequency of Opportunities for Collaboration 54
Figure 17. Mentors: Frequency of Opportunities for Collaboration 54
Figure 18. Mentors: School Environment and Collaboration 57
Figure 19. Residents: School Environment and Collaboration 58
Figure 20. Confidence Related to Efficacy 59
Figure 21. Percentage of Gains in Confidence that Teachers Attribute to LP-UTR 63

Teaching & Learning in Collaboration

INTRODUCTION

Most teachers In the U.S. begin their professional learning continuum in a traditional teacher preparation program, with coursework in core content, theory, and pedagogy, followed by supervised student teaching, induction and mentoring support during the first years in the classroom, and ongoing professional development for the remainder of their careers. While traditional programs still train the vast majority of teacher candidates, alternative certification programs have, over the past 20 years, come to play a critical role in preparing new teachers. Begun as a way to stem teacher shortages and staff high-need urban and rural schools, alternative programs typically provide intensive summer training for teacher candidates, who then start their professional learning continuum in the classroom, as full-time teachers. For the most part college graduates with majors in fields other than education, they then gradually acquire their formal academic coursework and credentials at their own pace, through partner institutions.

Teacher residency programs (TRPs)—including the Learning Partners–Urban Teacher Residency, or LP-UTR, funded by a 2014 U.S. Department of Education Teacher Quality Partnership award—offer a third option, one that fills the space between traditional and alternative certification programs (Zeichner, 2010). Helping new teachers simultaneously develop the academic and practitioner knowledge they need to succeed, clinically-oriented TRPs immerse teacher candidates in the core pedagogical coursework in an initial year of teaching, during which they also have the expert guidance from experienced mentors and coaches. Modeled after the medical school approach of using designated schools as teaching "hospitals," the learning is designed to be more school-focused. By design, when candidates become new teachers, their professional learning continues, not only through ongoing professional development but also through formal and informal means as they continue to teach in communities of practice, in the same or similar schools as their initial placement.

TRPs still represent a small percentage of teacher preparation programs in the U.S., but they have attracted considerable interest and spread all across the country in recent years. Federal and foundation sources have provided funding to develop and scale residency programs; moreover, the larger field of teacher preparation has increasingly shifted its attention to expanding the clinical, on-the-job training of teachers to support their transition into the teaching profession (e.g., AACTE, 2018; NCATE, 2010).

Interestingly, despite this sustained interest in and continued expansion of clinical preparation for teachers, research on their effectiveness remains limited. For example, reported findings of retention outcomes from residency studies rely primarily on descriptive rather than causal methods. Many studies face issues due to small sample sizes, and very few have investigated effectiveness on student achievement outcomes. Research conducted by Rockman et al for the Urban Teacher Residency (UTR, Sloan & Blazevski, 2015; Sloan, Allen,

Blazevski, Carson, & Rockman, 2018), is one such exception, and the present evaluation study was aimed at building upon that earlier research. Together, the growing push for residency programs or expanded clinical preparation, and the need for improved research about TRPs, suggest there is a growing audience within the field that can benefit from research like the present study.

The study relies in part on metrics refined during UTR—like LP-UTR, a partnership between New Visions for Public Schools, Hunter College, and the New York City Department of Education (NYC DOE)—launched in 2009 and funded by an earlier five-year award from the U.S. Department of Education's Teacher Quality Partnership program (www2.ed.gov/programs/tqpartnership). UTR was designed to staff high-needs New York City schools and content areas with highly-qualified teachers, not only giving new teachers a way to enter the field under the guidance of a trained mentor, but also developing school capacity and support for experienced as well as novice teachers. Hunter, New Visions, and the NYC DOE further developed their residency model during the Math and Science Teacher Residency (MASTER), an NSF-funded initiative that, from 2013–2016, trained 47 new math and science teachers (see Sloan, Allen, Blazevski, & Carson, 2017).

Partners began their five-year LP-UTR project in 2014, with an initial planning year, followed by three years of program implementation and a final year for analysis and reporting. When the third cohort of LP-UTR teacher candidates entered the program for the 2017–2018 school year, it was the ninth consecutive cohort under the overarching UTR residency model, which continues to enroll aspiring teachers each year (advance.newvisions.org/utr).

LP-UTR employed the same core components as UTR and MASTER, including:

- a selective recruitment process,
- a year-long residency under the tutelage of trained mentors and instructional coaching staff,
- Hunter College coursework leading to a Master's degree and New York state certification,
- induction-year support, and
- a stipend in exchange for a four-year commitment to teaching in New York City's public schools.

LP-UTR was, however, a new venture for project partners. It combined the residency model with the NYC DOE's Learning Partners Program (LPP), introduced in 2014 under by then Chancellor Carmen Farina as a way to help teachers and administrators share their expertise and learn from one another across school boundaries (We Teach NYC, 2019).[1] In addition to offering new opportunities to foster professional growth Learning Partners also provided a set of structures for teacher collaboration intended to support schools that were interested in advancing their teacher induction and professional development work. LP-UTR partners selected schools which expressed interest and previous experience to serve as "host" schools to form triads with "partner" schools.

[1] For an overview of the Learning Partners program, see: http://schools.nyc.gov/Academics/InterschoolCollaboration/learningpartners/default.htm. For more detailed case studies about the program, see Wallenstein, J., Corwin, E., Duckett, W., Gowdey, L., Hamburg, D., Jordan, K. Leck, C., Peralta, R., & Vasquez, A. (2015). Learning Beyond Boundaries: Case Studies of Interschool Collaboration in the New York City Department of Education Learning Partners Program. New York City Department of Education.

Mentors and coaches in each triad worked together to identify a problem of practice that would provide a focus and guide to their work across schools. LP-UTR thus broadened the scope of teacher professional development associated with the program and further explored the best ways to prepare, support, and retain teachers both during their preparation and as they continue in their careers. It also expanded the areas in which residents could be certified, adding Teaching English as a Second Language (TESOL), an area increasingly considered high-needs, while continuing to prepare Special Education, English, science, and math teachers.

LP-UTR also offered new opportunities for research. Over the course of the five-year project, Rockman continued to work with partners to refine ways to explore and assess partners' efforts to develop teachers, integrate theory and practice in coursework and classrooms, build capacity—and sustain changes. The report focuses on five main research questions that reflect the goals and outcomes set forth by LP-UTR partners (see Appendix B for the project logic model) as they began their joint initiative:

- *Did LP-UTR partners meet their goals of recruiting a diverse corps of aspiring teachers who would succeed in the program—and in the classroom?*
- *How did the LP-UTR program evolve over the three project years?*
- *In what ways did both novice and experienced teachers grow by participating in LP-UTR?*
- *What effect did cross-school collaboration have on school support, teachers' growth, and collective efficacy?*
- *What impact did LP-UTR participation have on student achievement?*

Following an overview of methods, the report begins with a summary of key findings pertaining to each of these central questions. The subsequent sections further explore each one, generally following the broadening of goals from shorter- to longer-term outcomes: Part 1 shares findings on outcomes related to recruitment and residents' program performance. Part 2 describes the evolution of the LP-UTR model as partners defined their roles and honed features and processes needed to support teacher and school growth. Part 3 begins the discussion of teachers' growth, based on performance assessments designed to monitor residents' development, as well as and mentors' and residents' self-assessments of skills and confidence. Shifting from professional learning to the professional learning community, Part 4 discusses mentors' and residents' feedback on levels of support and collaboration, and their self-assessments of collective efficacy. The fifth section reports on results from analyses of LP-UTR's impact on student achievement. The report concludes by revisiting the central questions and suggesting areas for further investigation.

METHODS

The LP-UTR evaluation used a mixed-methods approach and metrics refined over multiple years of the UTR and MASTER evaluations. It also employed new strategies to gauge the impact of combining elements of the LP and UTR models. The evaluation drew on the theoretical framework of improvement science, which offers a problem-

centered approach to implementing and scaling innovative ideas and insights on how educators, in collaborative networks, might address the complex challenges of school change (see, for example, Bryk 2009; Bryk, Gomez, and Grunow, 2011; Bryk, Gomez, Grunow, & LeMahieu, 2015; Redding, Canata, & Taylor Haynes, 2017).

Our methods, designed to generate data that could be triangulated across multiple sources, included reviews of recent literature situating LP-UTR in the larger context of teacher residencies; analyses of resident demographic and program performance data from New Visions and Hunter, used to examine the composition and diversity of LP-UTR cohorts and program selectivity and success outcomes. The evaluation also relied on self-reported survey data from residents and mentors on efficacy, practice, and school climate; and qualitative data from partners and participants, gathered through focus groups, interviews, and observations. In addition, we analyzed school- and student-level academic performance using publicly available data and data assembled through requests from New Visions and the NYC DOE to compare performance of students taught by LP-UTR mentors and residents to matched groups of teachers not participating in either Learning Partners or LP-UTR.

Project Records and Performance Data

- Information about LP-UTR's design and documentation—e.g., the Cycles of Learning overview and template, the project framework, and descriptions of leadership roles—came from the NYC DOE forms and guidelines.
- Admissions, placement, and tracking information on LP-UTR residents came from the *New Visions' Teacher Certification Team Data Warehouse*; data on residents' performance on the tools in the New Visions assessment suite came from the *New Visions Comprehensive Process Monitoring System.*
- Hunter College's *Office of Partnerships* provided data on residents' program completion rates, GPAs, and edTPA (Education Teacher Performance Assessment, scale.stanford.edu/teaching/edtpa) certification performance.
- New Visions and Hunter partners provided updates on program data and helped the evaluation team reconcile any issues related to inconsistencies or missing data.
- The source of comparison data about ethnic diversity and gender balance was the publicly available reports from the *Independent Budget Office (IBO) of the City of New York Public Schools*, which tracks trends and indicators among NYC DOE teachers.

Teacher Surveys and Samples

Rockman administered baseline surveys each fall to mentors and end-of-year surveys to both mentors and residents. Surveys included items to measure classroom practices and self-efficacy, and school culture, collaboration, and support. Additional items asked teachers about the role that preparation, practice, and support activities and/or providers played in their growth and development. The baseline response rates were high, with 76% of Year 1 mentors and 92% of both Year 2 and 3 mentors completing baseline surveys (see Table 1). Response rates for end-of-year surveys varied widely between years (see Table 2). In Year 1, 71% of mentors completed the survey. Rates increased to 90% in Year 2 before a decline to 58% in Year 3. Resident response rates for the end of year survey also varied between years. In the first year only 46% of residents completed the survey. This increased to 86% for the second year and then declined slightly to 75% in the third year.

TABLE 1. YEARS 1-3 SURVEY MENTOR RESPONDENTS BY SUBJECT, BASELINE

	Subject	Mentor Respondents	% of Participating Mentors	Returning Mentors	New Mentors
Year 1 (2015–16)	ELA	11	79% (N=14)	4	7
	Special Ed	9	69% (N=13)	5	4
	TESOL	6	86% (N=7)	1	5
	TOTAL/RESPONSE RATE	**26/34**	**76% (N=34)**	**10**	**16**
Year 2 (2016–17)	ELA	15	94% (N=16)	5	10
	Special Ed	17	89% (N=19)	4	13
	TESOL	3	75% (N=4)	1	2
	Science	6	100% (N=6)	3	3
	Math	2	100% (N=2)	2	0
	TOTAL/RESPONSE RATE	**43/47**	**92%**	**15**	**28**
Year 3 (2017–18)	ELA	5	100% (n=5)	4	1
	Special Ed	18	86% (n=21)	12	6
	TESOL	7	87% (n=8)	1	6
	Science	6	100% (n=6)	3	3
	Math	6	75% (n=8)	0	6
	Unknown	2	n/a	1	1
	TOTAL/RESPONSE RATE	**44/48**	**92%**	**21**	**23**

TABLE 2. YEARS 1-3 SURVEY RESPONDENTS BY SUBJECT, END OF YEAR

	Subject	Mentor Respondents	% of Participating Mentors	Returning Mentors	New Mentors	Resident Respondents	% of Participating Residents
Year 1 (2015–16)	ELA	8	62% (N=13)	3	5	6	46% (N=13)
	Special Ed	7	78% (N=9)	3	4	4	44% (N=9)
	TESOL	2	100% (N=2)	0	2	1	50% (N=2)
	TOTAL/RESPONSE RATE	**17/24**	**71%**	**6**	**11**	**11/24**	**46%**
Year 2 (2016–17)	ELA	15	100% (N=15)	6	9	15	100% (N=15)
	Special Ed	15	94% (N=16)	2	13	10	63% (N=16)
	TESOL	2	50% (N=4)	1	1	4	100% (N=4)
	Science	5	83% (N=6)	2	3	6	100% (N=6)
	Math	1	100% (N=1)	1	0	1	100% (N=1)
	TOTAL/RESPONSE RATE	**38/42**	**90%**	**12**	**26**	**36/42**	**86%**
Year 3 (2017–18)	ELA	3	60% (n=5)	3	0	5	100% (n=5)
	Special Ed	13	62% (n=21)	6	7	15	71% (n=21)
	TESOL	4	50% (n=8)	1	3	8	100% (n=8)
	Science	4	67% (n=6)	2	2	5	83% (n=6)
	Math	4	50% (n=8)	0	4	3	38% (n=8)
	TOTAL/RESPONSE RATE	**28/48**	**58%**	**12**	**16**	**36/48**	**75%**

Focus Groups and Interviews

- *Mid- and End-of-Year Resident and Mentor Focus Groups.* In winter, and again in early summer of all three years, we conducted hour-long focus groups with LP-UTR teachers. Grouped by subject area or schools, residents and mentors responded to questions about preparation and teacher development; collaboration and support; and school integration and program coherence. In years 2 and 3, we conducted focus groups with coaches, and in year 4 we conducted individual interviews. The end of year (early summer) focus groups and interviews were designed to gather general feedback about residency and mentoring experiences and more specific feedback about triad work and collaboration.

- *Partner and Stakeholder Interviews.* In October 2016, February 2017, and May 2017, we conducted phone or in-person interviews with eight school administrators from the host and partner schools in the two case study triads (see below). We also interviewed Hunter partners (April–May 2017 and February 2019) and TESOL and Special Education faculty members to discuss coursework and programmatic changes. We also took part in ongoing conversations with the NYC DOE Learning Partners team and the LP-UTR Steering Committee, and we conducted interviews with NYC DOE and New Visions project leaders in February 2019. Structured interview protocols aligned to the constructs were used.

- *Case Study Site Visits.* In consultation with project partners, Rockman chose two Year 1 triads providing range in school size, focus, demographics, culture, and prior residency experience to serve as LP-UTR case study sites. Some triads remained intact for Year 2, but others were reconfigured. In 2016–17 and 2017–18, Rockman followed the original case-study schools to their new triads, joining LP-UTR teams on three cross-school visits in October, January, and May. During visits, evaluators observed team sessions and classrooms, and debriefed on goals, activities, and takeaways from the visits with mentors, coaches, and administrators.

- Additional demographic and performance data on NYC DOE schools came from the districts' publicly available *School Quality Reports* and *School Performance Dashboards*; the NYC DOE also administers annual surveys on school culture and support, and makes the data available to the public. We also discussed evaluation activities in monthly calls with the NYC DOE internal research team, to coordinate activities, share findings, and avoid redundancies and reduce burdens on participants.

Student Data and Samples

For the primary analyses, we acquired 2017–18 achievement data for the schools of Cohort 1 (LP-UTR n=10, comparison schools n=29) and Cohort 2 (LP-UTR n=17, comparison schools n=50). No prior student achievement data was available in the records acquired from the NYC DOE. For the schoolwide analysis, we included all three LP-UTR Cohorts (17 LP-UTR Cohort 3 schools and 51 comparison schools (if still open in 2017–18). We matched LP-UTR schools to a set of comparison public schools in NYC using Mahalanobis distance (MD) matching, which involves using a single multivariate metric to optimize the covariate balance between schools. MD scores are based on school demographics (% economically disadvantaged, % Hispanic, % African American, % students with disabilities, % English learners, % female, and total enrollment). For each LP-UTR

school, we selected three comparison schools with the closest matching MD scores.[2] Each comparison school could only be matched to one LP-UTR school. The school sample from which we drew matched schools included all schools not participating in the LP-UTR program or the connected UTR or MASTER programs. We then submitted the list of LP-UTR and comparison schools to the NYC DOE, requesting de-identified achievement data (Regents scores and course credits) for all students in these schools. Achievement data linked to courses and de-identified/scrambled teacher IDs allowed us to connect student data to either the LP-UTR or comparison group. We matched teachers based on years of teaching, teaching assignment, and similar courses. No information was available about teacher certification programs that may have been associated with the comparison schools.

The residents of Cohort 1 were ELA, special education, or TESOL teachers. In Cohorts 2 and 3, there were ELA, math, science, special education, and TESOL teachers. In previous reports of the LP-UTR evaluation, the primary achievement analysis of linked teacher-student was limited to ELA, math, and science teachers, because NYC DOE did not typically identify special education, TESOL or other support teachers as "teachers of record" in their record-keeping. Beginning with the 2017–18 data records, some TESOL and Special Education teachers appeared as the "teacher of record" for a designated course and were therefore included in the linked teacher-student achievement analysis. Largely for the purposes of project stakeholders being able to compare Year 3 results to the way results of the prior two years were presented, we also examined achievement outcomes for all students with disabilities and for all EL students at the participating schools (i.e., unlinked to teachers).

Outcomes for the teacher-linked achievement analyses included 2017–18 Regents scores and percentage of credits earned of credits attempted only for students specifically linked to LP-UTR graduates or mentor teachers and matched comparison teachers. See Tables 3 and 4 for a breakdown of the numbers of teachers and students in the main achievement samples. School requirements and a student's grade level course progression can influence the number of credits earned, therefore the number of credits earned divided by the number of credits possible provided a more meaningful metric for outcome comparisons. We used students only once within a subject area, but they could appear in multiple analyses across the full set of analyses. Outcomes for the school-wide analyses included all 2017–18 Regents scores and all credits for students with disabilities and ELs at each school. We compared outcomes using analysis of covariance (ANCOVA), controlling for gender, economic disadvantage, EL status, special education status, and ethnicity (Hispanic and African American). Again, due to the unavailability of data, we were unable to account for prior achievement in our analyses. Throughout this evaluation, the LP-UTR team has focused on analyzing results by cohort, as their program tended to evolve over time and they wanted to be able to reflect on a cohort's results in the context of what transpired during the year.

[2] Upon receipt of the data from NYC DOE, it was determined that one LP-UTR school had only two acceptable matches; one was dropped due to difference in school type. In addition to MD scores, we used ANOVA to confirm that there were no significant differences in school-level demographics between LP-UTR and non-LP-UTR schools (none were found).

TABLE 3. OVERALL STUDENT ACHIEVEMENT DATA SAMPLES FOR LP-UTR GRADUATES

			LP-UTR		Non-LP-UTR	
		Content Domain	Teacher Sample	Student Sample	Teacher Sample	Student Sample
Regents	SY2016-17 (Year 1 graduates)	ELA	4	139	4	148
		Math	1	43	1	18
		Science	1	83	1	79
		Social Studies	2	142	2	180
		Subtotal	8	407	8	425
	SY2017-18 (Year 2 graduates)	ELA	10	340	11	299
		Math	5	179	5	158
		Science	9	344	9	229
		Social Studies	1	27	1	30
		Subtotal	25	890	26	716
		Combined Total	33	1,297	34	1,141
Course credits	SY2016-17 (Year 1 graduates)	ELA	4	205	4	230
		Math	1	51	1	28
		Science	1	90	1	106
		Social Studies	2	166	2	241
		Subtotal	8	512	8	605
	SY2017-18 (Year 2 graduates)	ELA	10	700	14	731
		Math	4	472	5	339
		Science	8	459	10	508
		Social Studies	1	35	1	39
		Subtotal	23	1,666	30	1,617
		Combined Total	31	2,178	38	2,222

TABLE 4. OVERALL STUDENT ACHIEVEMENT DATA SAMPLES FOR MENTOR TEACHERS

		LP-UTR		Non-LP-UTR	
		Teacher Sample	Student Sample	Teacher Sample	Student Sample
Regents	Mentors 1 Year Only	35	1,364	35	1,525
	Mentors Multiple Years	10	406	10	405
	TOTAL	**45**	**1,770**	**45**	**1,930**
Course Credits	Mentors 1 Year Only	32	1,878	36	2,426
	Mentors Multiple Years	10	600	11	814
	TOTAL	**42**	**2,478**	**47**	**3,240**

TABLE 5. OVERALL STUDENT ACHIEVEMENT DATA SAMPLES FOR MENTOR TEACHERS, BY SUBJECT AND YEARS OF PARTICIPATION

			LP-UTR		Non-LP-UTR	
		Content Domain	Teacher Sample	Student Sample	Teacher Sample	Student Sample
Regents	Mentor for 1 Year Only	ELA	9	292	9	327
		Math	9	253	9	392
		Science	8	464	8	427
		Social Studies	9	355	9	379
		Subtotal	35	1,364	35	1,525
	Mentor for Multiple Years	ELA	5	196	5	178
		Math	2	64	2	36
		Science	3	147	3	191
		Social Studies	-	-	-	-
		Subtotal	10	407	10	405
		Combined Total	45	1,771	45	1,930
Course credits	Mentor for 1 Year Only	ELA	10	446	9	509
		Math	6	298	9	630
		Science	9	687	9	722
		Social Studies	7	435	9	550
		Subtotal	32	1,866	36	2,411
	Mentor for Multiple Years	ELA	6	355	6	480
		Math	2	78	2	54
		Science	2	164	3	279
		Social Studies	-	-	-	-
		Subtotal	10	597	11	813
		Combined Total	42	2,463	47	3,224

SUMMARY OF FINDINGS

Throughout the project, LP-UTR partners engaged in collaborative, strategic activities designed to train both novice and experienced teachers, encourage shared practice within and across schools, build leadership capacity, and improve student achievement. Residents worked closely with mentor teachers to develop skills necessary to be successful in and meet the challenges of urban classrooms. Mentors and coaches participated in cross-school collaborations, bringing lessons learned from other teachers back to their schools and residents to move learning forward. Supporting learning within and across boundaries, schools assumed more responsibility for teachers' development. During the program's planning year and over three cohorts of residents, LP-UTR partners also worked to continuously improve the residency model and the structures needed to support a reflective, collaborative school culture.

Five years of external evaluation and summative analyses, the results of which are summarized here, help to answer the five central research questions, and, more broadly, indicate ways in which project partners laid the groundwork for sustained collaborative growth.

Did LP-UTR partners meet their goals of recruiting a diverse corps of aspiring teachers who would succeed in the program, and in the classroom?

- **Selective recruitment and screening resulted in three cohorts of residents who were diverse and gender balanced, bringing a wide range of experience to the program and classroom.** In all three cohorts, over half of the residents self-identified as non-white or people of color. By Cohort 3, gender balance improved, with 60% of residents identifying as female and 40% male. Data also showed an increase in work experience with each successive cohort—and thus the likelihood of increasing numbers of career changers.

- **Residency completion rates, which rose over three cohorts, were very close to original project goals.** Over LP-UTR's three years of implementation, 83% (108 of 130 enrolled) residents completed their residency, just three percentage below goals. Cohort 1 attrition was higher, due in part to the departure of TESOL residents, but program adjustments and added supports led to substantially less attrition and completion rates of 88% in Cohorts 2 and 3.

- **Residents' performance on degree and credentialing requirements were consistently high across cohorts.** Residents maintained high GPAs, and an average of 94% pass rates on the edTPA credentialing exam. Numbers of residents earning their degrees and obtaining NYS Initial Certification were also high across cohorts.

- **Program records provide strong evidence that LP-UTR was successful in placing and retaining its graduates in local classrooms.** All of the successful residency completers (n=108) were hired in high-need

NYC schools. By the end of the project, every residency completer (100%), regardless of cohort, had taught in a NYC public school the year immediately following their residency. All but two residents in Cohorts 1 and 2 taught in a NYC public school for a second consecutive year, and all but one Cohort 1 graduate (92%) taught a third consecutive year.

- **Numbers of LP-UTR mentors entering the profession through alternative programs and New Visions-Hunter College residencies increased each year.** The percentage of mentors trained in an alternative certification program grew from 35% in Cohort 1 to 66% and 80% in Cohorts 2 and 3. Twenty-five of the Cohort 3 mentors had taken part in a New Visions-Hunter residency—as residents, mentors, or both.

How did the LP-UTR model evolve over the three project years?

- **LP-UTR program changes included an increase in the number of school-based coaches.** In contrast to UTR and MASTER, mentors and residents at host schools had access to school-based coaches, increasing the capacity of the residency model in the school. By the third year of implementation, the number of school-based coaches had tripled from three in Year 1 to nine by Year 3. By adding additional coaches in LP-UTR schools, the program continued to develop teacher leadership capacity in teachers and schools.

- **Over the course of LP-UTR, New Visions partners continued to modify their suite of process monitoring assessments in an effort to better support residents' learning.** Changes involved modifying the assessment instruments and progress benchmarks as well as the use of performance data in cycles of inquiry for continued teacher development.

- **Refinement of data systems, norming, and training strengthened the data-driven aspects of the LP-UTR model.** Using data to make continuous, strategic adjustments was a key part of the LP-UTR model. New Visions improved its online platform and visualization tools and provided targeted training. Through multiple norming sessions, mentors and coaches achieved strong interrater reliability, which helped to focus and align support for each resident. Time was also set aside for stakeholders to access process assessment data for reflection and action planning.

In what ways did both novice and experienced teachers grow by participating in LP-UTR?

- **Capstone assessment data provided evidence of residents' growth.** In each successive cohort, a greater percentage of residents met or exceeded the established benchmarks on the Defense of Learning.

- **Danielson ratings of classroom performance showed increasing trajectories of growth in residents' ability to effectively plan instruction, manage their classroom, and use best practices.** Data gathered by mentors and coaches with the Danielson-based observation rubric, administered four times during the residency year, showed that residents in LP-UTR developed as teachers over the course of their residency, with gains in lesson-planning and classroom managements skills and use of best practices.

- **End-of-year survey data indicated growing confidence among residents in a range of instructional skills**. Self-reported data gathered as residents completed their clinical year indicated that residents were most confident in their abilities to motivate and engage students (Cohort 1 *M*=3.1, Cohort 2 *M*=3.4, Cohort 3 *M*=3.1) and develop effective lesson plans (Cohort 1 *M*=3.0, Cohort 2 *M*=3.4, Cohort 3 *M*=3.1), and less confident in classroom management skills. Confidence levels in assessment and data-use practices increased for Cohorts 2 and 3. In all three cohorts, mentors' self-assessments were higher, but margins narrowed for the second and third cohorts.

- **In each cohort, and across almost all categories, mentors reported high levels of confidence as well as increases over the course of the year.** Mentors generally indicated that they were very confident in their practice. Pre-post data from all three cohorts still showed growth in confidence in all but one instructional category, which remained constant. Cohort 2 mentors' self-reports showed the largest gains in confidence fall to spring.

What effect did cross-school collaboration have on teachers' growth and collective efficacy?

- **Mentors from all three cohorts reported observing classrooms more frequently across than within schools as part of triad work.** Mentors and school-based coaches shared skills and practices with colleagues in their triad to continue to develop as mentors and coaches and to help better support their residents.

- **Residents reported being more engaged in school activities than in triad activities**. There was a slight increase in the frequency of school-based activities between Cohorts 1 and 2, and a substantial increase for Cohort 3 residents.

- **Resident participation in triad visits hosted by their own school increased in each year of the program.** Self-reported survey feedback indicated increased participation by residents when their schools hosted triad teachers from other schools. Residents' participation in visits to other schools increased in the second year of the program, but it involved only a few schools and less than half of residents.

- **Compared to mentors, who credited LP-UTR cross-school sharing with growth, residents attributed their growth to a wide range of project partners and activities.** About a quarter of mentors attributed growth in their ability to share skills and practices with teachers in other schools to LP-UTR cross-school sharing. Residents credited Hunter coursework, LP-UTR coaching, and New Visions PD and support with helping them to develop specific subsets of skills.

What impact did participation in LP-UTR have on student achievement?

- **Overall, across core high school subjects, student achievement outcomes in LP-UTR graduate teachers' classrooms tended to be comparable to or better than those in non-LP-UTR classrooms.** Mirroring both past evaluation results about LP-UTR mentor teachers as well as the pattern of results for LP-

UTR graduates, students taught by the LP-UTR mentor teachers attained similar or greater levels of achievement than their peers taught by comparison mentor teachers.

- **Students taught by mentors who participated in LP-UTR multiple years performed better on exams compared to the non-LP-UTR group on four out of five contrasts.** The more experienced LP-UTR mentor teachers thus exhibited a pattern of achievement results that contrasted with the novice mentor teachers and more closely resembled the patterns observed among the LP-UTR graduate teachers.

- **Comparative analyses of schoolwide achievement data revealed that students with disabilities fared consistently better in core courses at LP-UTR schools than at non-LP-UTR schools.** Results for all core high school subjects—ELA, math, science, social studies—showed higher credit earning rates, at statistically significant levels, at LP-UTR schools.

- **For English language learners, there was consistent evidence of a positive LP-UTR impact on course credit earning.** Examination results were mixed, but ELs still tended to fare the same or better than peers at similar schools. Although average exam performances were at comparable levels, English learners at LP-UTR schools earned a higher percentage of course credits in ELA, math, science, and social studies than students in non-LP-UTR schools.

PART 1: RECRUITMENT & RESIDENTS' PERFORMANCE

Describing the theory of action behind the LP-UTR model, Marisa Harford, New Visions' Director of Teacher Residencies, explained that it "involves structures and systems around every program element—starting with recruitment."

- Selective LP-UTR recruitment helped diversify the new teacher corps entering NYC classrooms:
  - In all three cohorts, over half of the residents identified as non-white or people of color.
  - A bachelor's degree was the highest degree earned by 70% of residents in each cohort, which also included residents with advanced degrees.
  - Residents, including increasing numbers of career changers, brought a wide range of previous work experience, some up to 10 years.
  - Gender balance improved, with a 60/40 balance between females and males in Cohort 3.
- Residency completion rates were lower in Cohort 1, but rose from 71% to 88% for Cohorts 2 and 3. Over LP-UTR's 3 years, 83% of the residents completed the year-long residency, just 2 residents shy of partners' goal of 85%.
- High GPAs and edTPA pass rates provided other indicators of solid performance by LP-UTR residents.
- Program records convincingly show LP-UTR has been successful in retaining its graduates in local classrooms in the first years after program completion.
- Besides the strong record of retaining LP-UTR graduates in NYC public schools, records show that many graduates are hired by LP-UTR schools.
- Increasing numbers of mentors trained through residencies suggest familiarity with residency goals and processes. High numbers of returning mentors and host schools in Cohort 3 further indicate school support for those goals and processes.

This report also begins with recruitment, describing the composition and program performance of the three cohorts of LP-UTR residents, selected through a rigorous, multi-step process designed to diversify the new teacher workforce and select teachers most likely to succeed in the program and in the classroom. In addition to providing data on the background characteristics considered in LP-UTR admission—race and ethnicity, educational background, gender, work experience—this part of the report also shares performance data that confirms their program success, including residency completion figures, Hunter GPAs, edTPA passage rates, along with job placement and retention figures.

LP-UTR mentors and host schools, also screened through a rigorous recruitment process, played a critical role in that success, and Part 1 concludes with mentor profiles and a discussion of how multi-year participation of mentors and schools in LP-UTR might have affected program sustainability.

LP-UTR COHORT COMPOSITION

Selecting candidates whose diverse backgrounds made them a good fit for urban classrooms was a focus throughout the development and scaling of the residency model. Our analysis of demographic background characteristics includes factors LP-UTR partners considered during a lengthy screening and admissions process, including race and ethnicity, educational background, prior work experience, and gender. The figures reported here are based on the number of residents enrolled in the program at the start of the cohort. (See Table 6.)

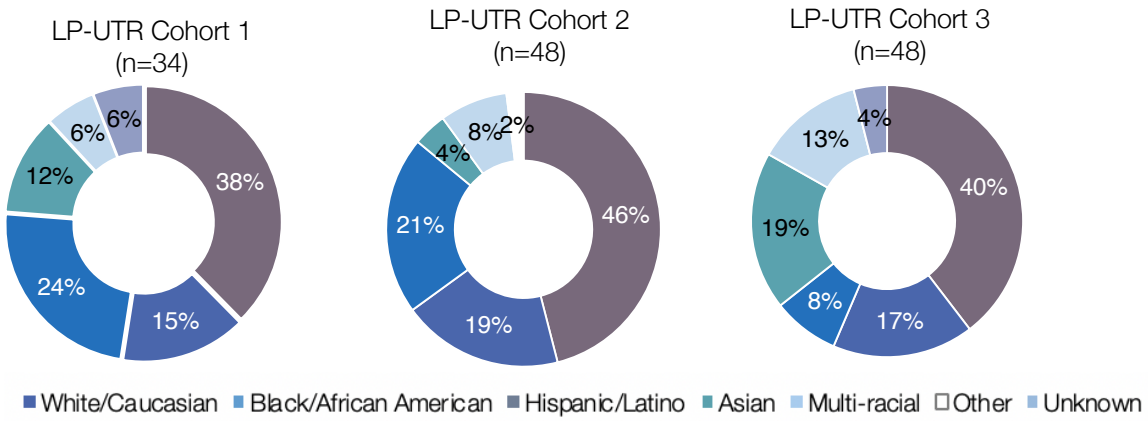
TABLE 6. NUMBER OF RESIDENTS ENROLLED BY SUBJECT AREA

	Cohort 1	Cohort 2	Cohort 3
	N	N	N
Biology		2	2
Chemistry		1	1
Earth Science		3	3
ELA	14	16	5
Math		2	8
Special Education	13	19	21
TESOL	7	5	8
Cohort Totals	**34**	**48**	**48**

Race and Ethnicity

One of the goals in resident recruitment was selecting candidates who brought diversity to each cohort. In all three LP-UTR cohorts, over half of the residents identified as non-white or people of color. The percentage of residents who self-identified as Black/African American remained relatively constant across the three years of the program, increasing from 15% in Cohort 1 to 19% in Cohort 2, then decreasing slightly to 17% in Cohort 3. The percentage of Hispanic/Latino residents was similar in Cohorts 1 and 2, at 24% and 21% respectively, but decreased significantly to 8% in Cohort 3. This decrease in Hispanic/Latino residents was in part offset by an increase in the percentage of Asian residents from 4% in Cohort 2 to 19% in Cohort 3. (See Figure 1.)

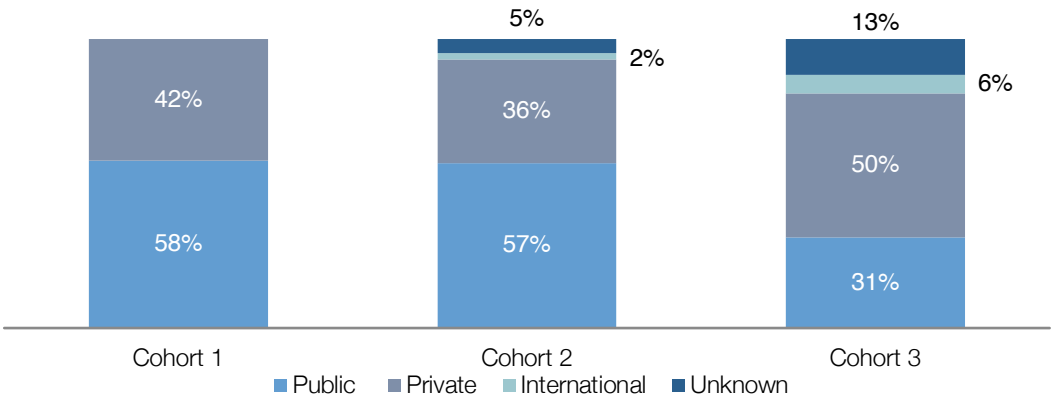
FIGURE 1. ETHNIC REPRESENTATION OF RESIDENTS



Educational Background

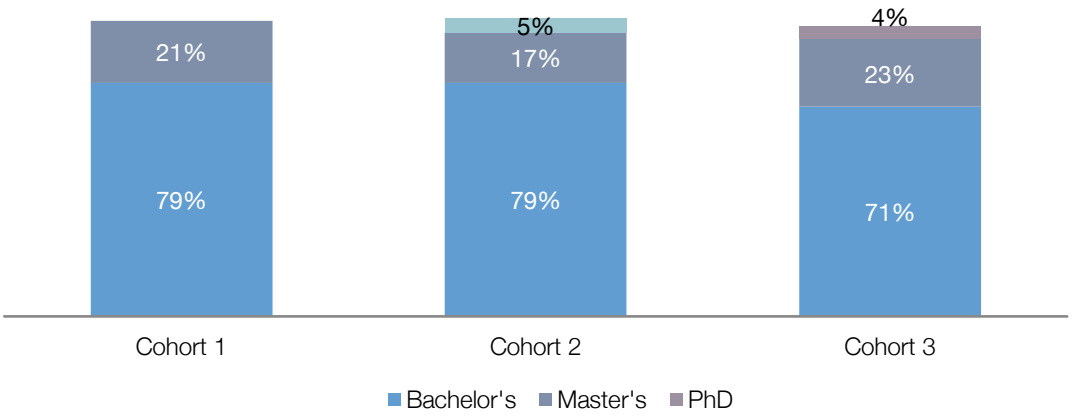
Partners also looked for a range in educational backgrounds, including the kinds of degree-granting institutions residents attended and the degrees they held, as well as the experience they had acquired since graduation. LP-UTR Cohorts 1 and 2 had similar percentages of residents who completed their post-secondary schooling at public universities, with 58% of Cohort 1 residents and 57% of Cohort 2 residents having done so. Less than a third, 31%, of Cohort 3 residents attended a public university. (See Figure 2.)

FIGURE 2. INSTITUTIONS ATTENDED BY RESIDENTS



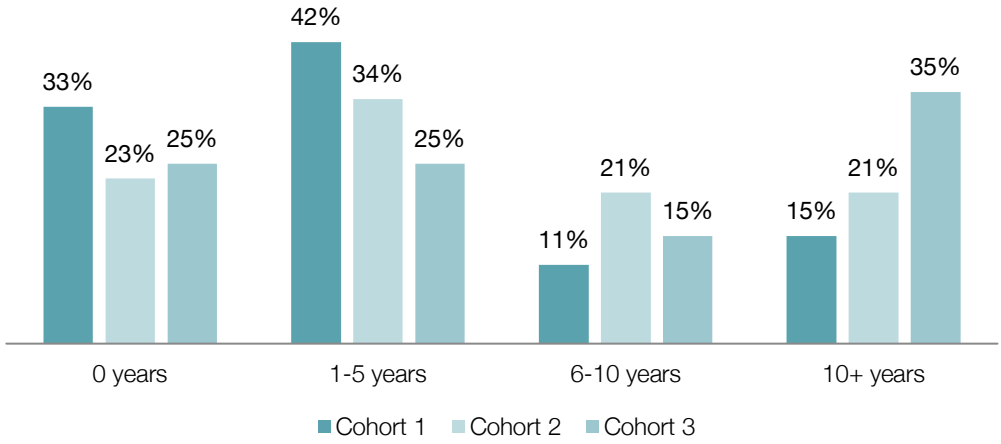
The previous degrees held by residents were very similar across all three LP-UTR resident cohorts. Seventy-nine percent of residents in Cohorts 1 and 2 held a Bachelor's degree prior to enrolling in LP-UTR. The remaining Cohort 1 residents (21%) held Master's degrees. In Cohort 2, 17% of residents held Master's and 5% Doctoral Degrees. Cohort 3 had slightly fewer residents with Bachelor's degrees (71%), and more residents with Master's (23%) and Doctorate (4%) degrees. (See Figure 3.)

FIGURE 3. PREVIOUS DEGREES HELD BY RESIDENTS



A wide range of work experience could also help diversify cohorts, and data showed an increase in work experience with each successive LP-UTR Cohort—and thus the likelihood of increasing numbers of career changers. In Cohort 1, 26% of residents had six or more years of work experience. This increased to 42% of residents in Cohort 2, and again to 50% by Cohort 3. (See Figure 4.)

FIGURE 4. PREVIOUS WORK EXPERIENCE BY RESIDENTS



Gender Balance

There are typically more females than males in school faculties, and females outnumbered males in all three LP-UTR cohorts. Cohort 2 was the most imbalanced cohort, with females comprising 75% of the cohort. Cohort 3 was the most gender-balanced cohort, with 60% of the residents identifying as female and 40% as male. (See Table 7.)

TABLE 7. GENDER REPRESENTATION, BY COHORT

	Cohort 1		Cohort 2		Cohort 3	
	N	%	N	%	N	%
Female	23	68%	36	75%	29	60%
Male	11	32%	12	25%	19	40%

Source: New Visions Residency Program Overview File, April 26, 2019

RESIDENTS' PROGRAM PERFORMANCE

Another important goal of the screening and admissions process was to select residents who seemed most likely to succeed, both in the program and in the classroom. Program success was in part measured by teacher candidates completing the year-long clinical residency and Hunter coursework, meeting the New York State certification requirements, and obtaining a position in a New York City school, preferably a high-needs school. The discussion below shares results of these metrics for all three LP-UTR cohorts. A later discussion focuses on measures of residents' classroom success, including a set of performance assessments developed by New Visions for the residency model (see p. 38).

Residency Completion Rates

Over the three years of the program, a total of 83% of the residents (108 of 130 enrolled) completed the year-long residency, which was just two residents shy of the goal of 85% (see Logic Model, Appendix B). The remaining 17% withdrew or were counseled out of the program (see Table 8). While Cohort 1 had a 71% completion rate, there was considerable improvement in subsequent years, with Cohorts 2 and 3 exceeding the goal, with 88% completion rates. Across subjects, there was variability in program completion rates: in Cohort 1, only 29% (n=2) of TESOL residents completed the program—due in part to the addition of TESOL to the program and a lack of uniformity and certainty about school TESOL policies—and 69% (n=9) of Special Education residents, compared to 93% (n=13) of ELA residents. In Cohort 3, math had only a 50% (n=4) completion rate.

TABLE 8. PROGRAM PERFORMANCE, RESIDENCY ENROLLMENT AND COMPETION RATES

	SUBJECT	Residents Enrolled	Withdrew / Out Counseled		Completed Residency	
		N	N	%	N	%
YEAR 1	ELA	14	1	7%	13	93%
	TESOL	7	5	71%	2	29%
	Special Education	13	4	31%	9	69%
	Subtotal	**34**	**10**	**29%**	**24**	**71%**
YEAR 2	ELA	16	1	6%	15	94%
	TESOL	5	1	20%	4	80%
	Special Education	19	3*	16%	16	84%
	Math	2	1	50%	1	50%
	Biology	2	0	0%	2	100%
	Chemistry	1	0	0%	1	100%
	Earth Science	3	0	0%	3	100%
	Subtotal	**48**	**6**	**13%**	**42**	**88%**
YEAR 3	ELA	5	0	0%	5	100%
	TESOL	8	0	0%	8	100%
	Special Education	21	1	5%	20	95%
	Math	8	4	50%	4	50%
	Biology	2	1	50%	1	50%
	Chemistry	1	0	0%	1	100%
	Earth Science	3	0	0%	3	100%
	Subtotal	**48**	**6**	**13%**	**42**	**88%**
	Total	**130**	**22**	**17%**	**108**	**83%**

*One resident took a leave of absence.
Source: New Visions Residency Program Overview File, April 26, 2019

Residents' Coursework Performance and Certifications

The LP-UTR program monitored the residents' GPAs for their Hunter College coursework. Cohort 1 GPAs for ELA, Special Education, and TESOL were slightly greater than those of Cohort 2 and very similar to those of Cohort 3. For every subject area except math, the GPAs of Cohort 3 residents exceeded those of Cohort 2. TESOL residents had a difference of 0.08 in GPA, the smallest of any subject area, between Cohorts 2 and Cohort 3. Chemistry, which only had one resident in each cohort year, had a difference of 0.56, the largest, between Cohorts 2 and 3. (See Table 9.)

TABLE 9. AVERAGE GPAS, BY RESIDENCY COHORT

Subject	Cohort 1 2015–16	Cohort 2 2016–17	Cohort 3 2017–18
Biology		3.67 (n=2)	3.92 (n=1)
Chemistry		3.43 (n=1)	3.96 (n=1)
Earth Science		3.67 (n=3)	3.75 (n=3)
ELA	3.79 (n=13)	3.71 (n=15)	3.83 (n=5)
Math		4.00 (n=1)	3.70 (n=4)
Special Education	3.88 (n=9)	3.74 (n=17)	3.85 (n=19)
TESOL	3.79 (n=2)	3.61 (n=4)	3.69 (n=8)
Total	3.82 (n=24)	3.71 (n=43)	3.80 (n=41)

Source: Hunter College of Office Partnerships, March 18, 2019.

Of the 83% (n=108) of enrollees who completed the residency, 77% (n=83) earned degrees, 68% (n=74) obtained NYS Initial Certification, and 72% (n=78) submitted edTPA as of March 2019. edTPA pass rates were high, with 94% (n=73) of those who submitted portfolios passing, 36% (n=26) with mastery; Cohort 3 had a 100% edTPA pass rate across subjects. (See Table 10.) According to program leaders, everyone who has not yet earned their MA or initial certification is still "persisting" in the program—they have not dropped out, but just require more time.

TABLE 10. PROGRAM PERFORMANCE, CERTIFICATIONS

	SUBJECT	Degree Conferred (to residency completers)		NYS Initial Certification		edTPA		
		N	%	N	%	Submitted	Passed	Passed with Mastery
YEAR 1 (n=24)	ELA (n=13)	13	100%	12	92%	12	12 (100%)	9 (75%)
	TESOL (n=2)	2	100%	2	100%	2	2 (100%)	0 (0%)
	Special Education (n=9)	8	89%	8	89%	8	7 (88%)	4 (57%)
	Subtotal	**23**	**96%**	**22**	**92%**	**22**	**21 (95%)**	**13 (62%)**
YEAR 2 (n=43)	ELA (n=15)	13	87%	13	87%	13	12 (92%)	2 (17%)
	TESOL (n=4)	4	100%	3	75%	3	3 (100%)	2 (67%)
	Special Education (n=17)	14	82%	11	79%	13	10 (77%)	5 (50%)
	Math (n=1)	1	100%	1	100%	1	1 (100%)	0 (0%)
	Biology (n=2)	1	50%	0	0%	0	0 (0%)	0 (0%)
	Chemistry (n=1)	0	0%	1	100%	0	0 (0%)	0 (0%)
	Earth Science (n=3)	2	67%	2	67%	2	2 (100%)	1 (50%)
	Subtotal	**35**	**81%**	**31**	**72%**	**32**	**28 (88%)**	**10 (36%)**
YEAR 3 (n=41)	ELA (n=5)	4	80%	4	80%	4	4 (100%)	0 (0%)
	TESOL (n=8)	0	0%	0	0%	1	1 (100%)	0 (0%)
	Special Education (n=19)	17	89%	12	63%	13	13 (100%)	3 (23%)
	Math (n=4)	0	0%	1	25%	1	1 (100%)	0 (0%)
	Biology (n=1)	1	100%	1	100%	1	1 (100%)	0 (0%)
	Chemistry (n=1)	1	100%	1	100%	1	1 (100%)	0 (0%)
	Earth Science (n=3)	2	67%	2	67%	3	3 (100%)	0 (0%)
	Subtotal	**25**	**61%**	**21**	**51%**	**24**	**24 (100%)**	**3 (13%)**
	Total	**83**	**77%**	**74**	**69%**	**78**	**73 (94%)**	**26 (33%)**

Source: Hunter College Office of Partnerships, March 18, 2019

Job Placement and Retention

A key outcome goal for LP-UTR job placement was that at least 92% of the residents would teach in NYC DOE public schools during their first year of teaching post-residency (see logic model, Appendix B). By the end of the project, every resident, regardless of cohort, had taught in a NYC DOE public school the year immediately following their residency (see Table 11). In addition, all but two of 66 residents in Cohorts 1 and 2 taught in a NYC DOE public school for a second consecutive year (and one of the two returned in the following year). Moreover, all but one Cohort 1 graduate was still teaching in NYC in their third year after residency, far exceeding the three-year goal of 80%.

Program records convincingly show LP-UTR is successful in retaining its graduates in local classrooms in the initial years after program completion. The program includes an incentive for staying, in the form of a stipend in exchange for a four-year commitment to teaching in New York City's public schools, but positive findings comport not only with evidence of success in early retention efforts reported by Rockman in the evaluation of the original UTR project, and longer-term retention rates reported in an extension study (Sloan, et al., 2015, 2018).

TABLE 11. RESIDENT TEACHING STATUS IN NYC DOE OR NV CHARTER SCHOOL

	SY2016–17		SY2017–18		SY2018–19	
	Count	Percent	Count	Percent	Count	Percent
Cohort 1 (n=24)	24	100%	23*	96%	23*	96%
Cohort 2 (n=42)	N/A	N/A	42	100%	42	100%
Cohort 3 (n=42)	N/A	N/A	N/A	N/A	42	100%

Source: New Visions Residency Program Overview File, June 10, 2019
* The resident who was not teaching in 2017–18 returned to teaching in 2018-19 and a different resident who was previously teaching did not teach in 2018-19.

Besides demonstrating a strong record of retaining the LP-UTR graduates within New York City public schools, the program has also shown that many of its graduates are hired by the LP-UTR schools, in essence developing a corps of new teachers trained in and committed to inquiry and shared practice. Overall, 46 of 108 total residents (43%) were hired at schools that hosted LP-UTR residents during at least one of the three years of the residency program. About half of the LP-UTR schools (13 of 25) hired at least one of the residents they had hosted during the project, and six of those schools also hired a resident that had been hosted at another LP-UTR school. Altogether, 40% of all 25 schools had hired an LP-UTR graduate that had been hosted at another LP-UTR school. Resident placements by cohort were as follows:

- Five of the 24 Cohort 1 residents (21%) took up full-time positions as teachers of record in their LP-UTR host Schools, and five other Cohort 1 residents were hired by two other LP-UTR host schools.
- 10 of the 42 Cohort 2 residents (24%) became teachers of record in their LP-UTR host school in their first year of teaching post residency, and nine residents (21%) were hired by other LP-UTR host schools including five at one those schools.
- Nine Cohort 3 residents (21%) were hired by their LP-UTR host Schools, and eight (19%) were hired by other LP-UTR host Schools.

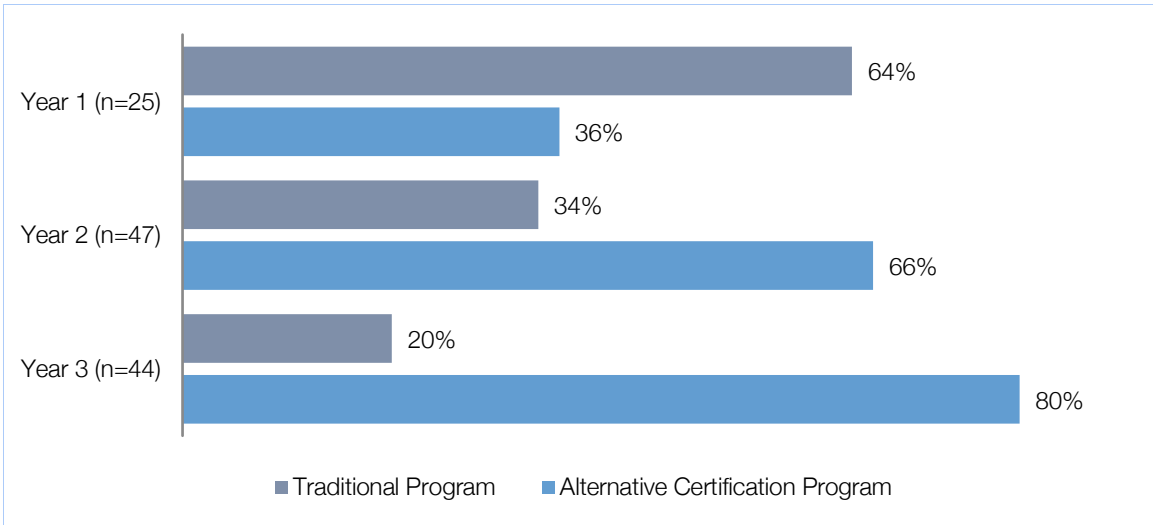
RECRUITMENT OF LP-UTR MENTORS AND SCHOOLS

LP-UTR's selective recruitment included not just residents but also mentors and host schools. Over the course of UTR and MASTER, New Visions developed screening procedures for selecting mentors who would be committed to supporting residents' needs and growth, taking part in school-based program activities, and taking on new leadership roles in those activities. School leaders made a commitment to systematically support both residents and mentors. For LP-UTR, the emphasis on building a supportive culture of reflective inquiry and shared practice increased as schools assumed more responsibility for developing teachers and teacher leaders.

Mentors' Backgrounds and Experience

The number of LP-UTR mentors increased at pace with the number of residents between Years 1 and 2 of the program. What also increased was the percentage of mentors themselves who entered teaching through an alternative certification program, from 36% in Year 1 to 66% in Year 2, and again to 80% in Year 3 (see Figure 5). A significant number of mentors across all three cohorts also brought familiarity with a Hunter–New Visions residency (UTR or MASTER) to LP-UTR. In Year 1, six of the 17 mentors completing the end-of-year survey had taken part in one of the residency programs, all as mentors. In Year 2, 15 out of 42 respondents had previously participated in a residency program—12 as mentors, and two as residents. In Year 3, 25 of the 44 mentors who responded to the survey had been involved previously, 16 as mentors, four as residents, and five as both residents and mentors. As discussed later in this report, the backgrounds and development opportunities during LP-UTR positioned mentors as advocates for both expanded clinical preparation and shared practice, which could help meet the longer-term goals of program scalability and sustainability (see Part 4 p. 53, and Conclusions, p. 81).

FIGURE 5. MENTOR TEACHERS' PREPARATION PROGRAMS



Source: REA Annual Surveys

Mentors across the three cohorts had very similar levels of teaching experience. The largest percentage of mentors across cohorts had between six and 10 years of teaching experience (see Figure 6). While mentors across cohorts had similar overall years of teaching experience, the number of years spent teaching *at their*

current schools varied widely. Forty percent of Cohort 1 mentors had only been at their current school for two to three years, 50% of Cohort 2 teachers had between four and five years of teaching experience at their school, while 51% of Cohort 3 mentors had over 11 years of experience at their current schools. (See Figure 7.)

FIGURE 6. YEARS OF OVERALL TEACHING EXPERIENCE, MENTORS

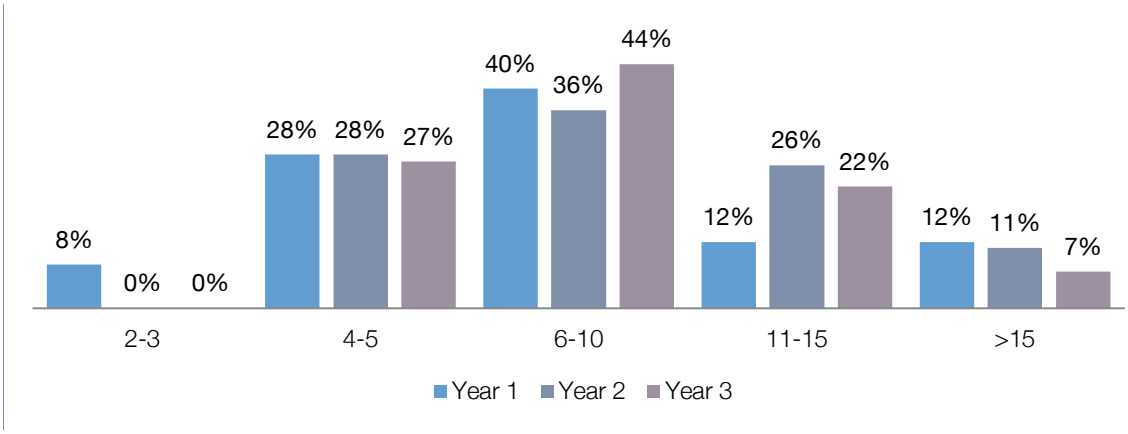
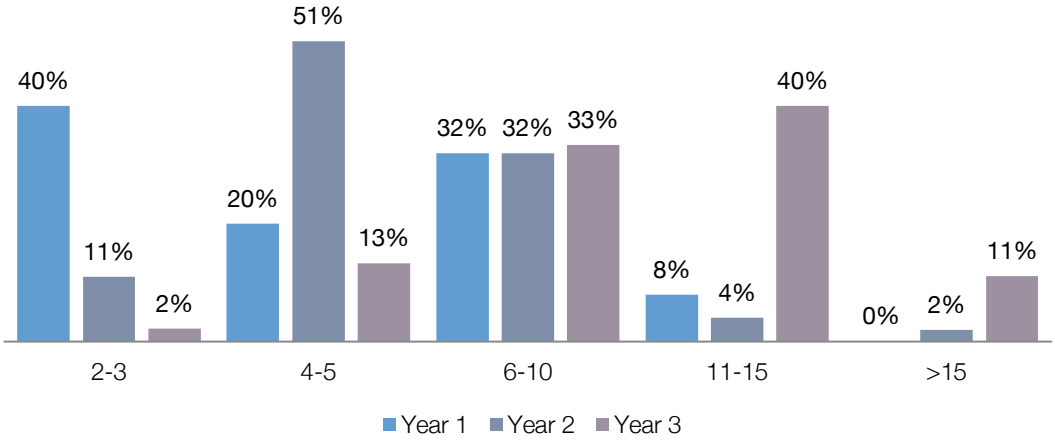


FIGURE 7. YEARS TEACHINGT AT CURRENT SCHOOL, MENTORS



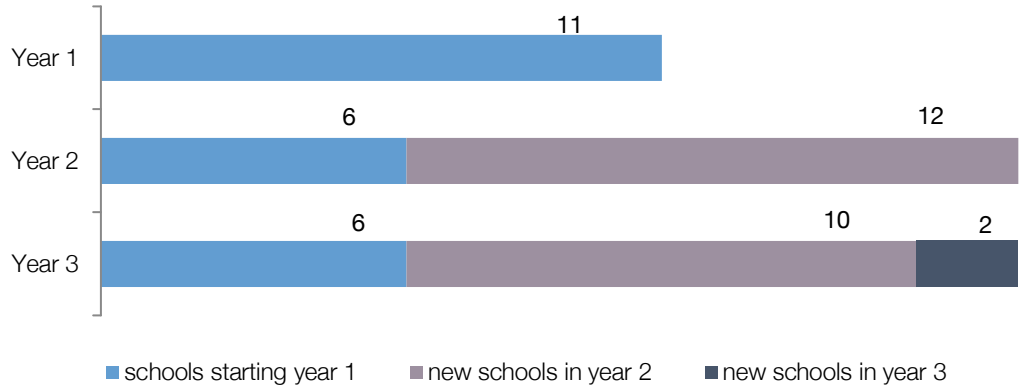
LP-UTR participation not only provided mentors new opportunities to collaborate and learn from each other, but also an opportunity and invitation to develop leadership skills. Mentor selection for LP-UTR included having the designation of being a "model teacher," which the Unified Federation of Teachers in New York City defines as "highly-skilled" and having a "drive to improve the instructional quality of their schools." The model teacher designation, granted by school administrators, required that teachers make tenure prior to the start of the school year, receive advanced ratings on the prior year's review, and apply to and be interviewed for the position, which also came with extra responsibilities related to leadership roles. This represented another building block for sustaining a culture of collaboration and shared practice.

LP-UTR Host Schools

For LP-UTR, schools with previous experience were selected as "host" schools and formed triads with "partner" schools. Beyond the guidance that was an important part of the residency model, Learning Partners also provided a set of structures for teacher collaboration intended to support schools interested in advancing their teacher induction and professional development work. Mentors and coaches in each triad, for example, worked together to identify a problem of practice that would provide a focus and guide for their work across schools. LP-UTR broadened the scope of teacher professional development associated with the program to include experienced mentor teachers and further explored the best ways to prepare, support, and retain teachers both during their preparation and as they continue in their careers.

The development of LP-UTR participating schools was, for several, a process of progressive commitment and growth. Many of the LP-UTR schools that started in Year 1 or Year 2 participated in the project for more than one year (see Figure 8). Specifically, six of the 11 schools participating in the first year of LP-UTR continued for all three years of the program; 10 of the 12 schools that started in Year 2 went on to participate in Year 3. In Year 3, two new schools joined, for a total of 18. As with returning mentor teachers, the repeat participation of schools signaled a commitment to developing teachers that could sustain project goals and practices beyond the official grant period.

FIGURE 8. LP-UTR SCHOOL PARTICIPATION, BY YEAR



Source: New Visions Data

PART 2: EVOUTION OF THE MODEL & PARTNERS' ROLES

Although partners had developed a residency program, the theory of action guiding their model reached beyond producing effective new teachers. As Marisa Harford noted, the goal was also "to build capacity in schools, in mentors, and coaches." Combining the residency model with Learning Partners expanded that scope even further, beyond school boundaries. The challenge for LP-UTR partners was to ensure that the combined model included the structures necessary to support so broad a level of skill- and capacity-building and collaboration. The numbers of returning mentors and schools discussed in Part 1 suggested a shared commitment to collaboration, but there were other practical questions: What constituted a common goal, one that could be shared across schools but still locally defined? What instructional needs were broad enough to be cooperatively addressed but specific enough to lead to actual changes in practice? What was the balance between respecting schools' and triads' autonomy and working toward project-wide goals?

- Collaborations between partners continued to move and shift with school and program needs.
- Learning Partners structures evolved to meet school and triad needs:
  - In Year 2, the Learning Focus Area (LFA) evolved from "teacher development" to "strengthening the instructional and leadership development of resident and mentor teachers."
  - In Year 3, some triads added an instructional focus tying school to triad goals and ensuring that the shared goal was still locally defined.
- Hunter College continued to play the role of connecting coursework to classroom practice.
  - In addition to adding TESOL to the program, in Cohort 3, TESOL as well as Special Ed residents were given a faculty point person to help them apply coursework to practice.
- New Visions continued to evolve data-driven, reflective inquiry and support for residents, mentors, and coaches, refining performance assessments and norming activities, building leadership capacity with mentors and coaches, and advancing data systems.
- The coaching model also saw continuous improvement, with the expansion of school-based coaches and support for their roles and growth.

This part of the summative report describes the combined model's evolution, and the decisions and changes partners made as the project moved forward. Our analysis investigates the following questions: What is the LP-UTR residency model? What prompted improvements and what changes were made over the three cohorts? How did partner roles evolve? The focus here is on the changes and improvements, and stakeholders' responses, based on feedback shared during focus groups and interviews conducted at the program's end. Parts 3 and 4 examine the impact of the gradually refined LP-UTR model on teachers' growth and school capacity and culture.

THE LP-UTR MODEL AND PARTNERS' ROLES

In a newly published book, New Visions Deputy Directors of Residencies, Marisa Harford and Rachelle Verdier lay out the expanded components if the LP-UTR model (Harford & Verdier, 2019). These components closely align with the eight features of high-quality teaching residencies outlined by Guha, Hyler, and Darling-Hammond (2016), but, according to Harford and Verdier, the more comprehensive features of LP-UTR differentiate it from other residency models. The LP-UTR model includes:

- A strategic partnership between New Visions, Hunter College, and the NYC DOE, who collaborate in an ongoing process of examining program data and making strategic adjustments.
- A rigorous admissions process designed to diversify the new teacher workforce, which means attracting a wide range of teachers, including career changers who wish to enter the teaching profession, and selecting candidates whose ethnicity better reflects student enrollments and thus research showing that having teachers of color and teachers who students can identify with improves outcomes (Gershensen, 2014; Egalite, et al., 2015); a stipend to offset costs in exchange for a four-year teaching commitment.
- At host schools, residents teach at least one course for a full-year alongside mentors.
- A commitment from school leaders to systematically support the mentors in their development as teacher leaders and support the residents' needs.
- Close attention to the alignment of the graduate coursework and the immediate needs of the classroom faced by the residents in their clinical experiences.
- Ongoing professional development for the highly-qualified mentors who provide one-on-one support to residents, and themselves have the support of school-based and program coaches and are immersed within a culture of reflective inquiry.
- Inclusion of cross-school partnerships that enable the LP-UTR stakeholders to engage in professional learning communities with similar schools; cross-school visitations that expand the opportunities for mentors to observe modelled instruction and use it to reflect on their planning and practice.
- Work centered on collaborative reflection through inquiry learning and continuous improvement efforts; a focus on data analysis and established systems for gathering and monitoring key data—including data from performance assessments designed to support residents as they develop their practice during the full-year residency—to inform practice and programmatic adjustments.
- Parallel learning structures—that is, use of the same decision-making processes whether they are making programmatic level decisions or student-level decisions. At each of the multiple levels of learning, the program stakeholders use strategic data to cyclically reflect and make decisions.

The features of the LP-UTR model that, according to the research, characterize successful clinical programs—strong partnerships, recruitment mechanisms, clinical preparation, mentoring, coursework linked to classroom practice, financial incentives—had been refined throughout UTR and MASTER and continued to be refined during LP-UTR. The features related to cross-school sharing and building stronger, more reflective communities on a broader scale required adjustments. Some of the key ones involved leveraging the structures already built into LP and UTR, and at the same time, according to Verdier, "maintaining key elements of their identity." The discussion below describes what each partner brought to the combined model and some of the changes made to create a coherent combined model.

The New York City Department of Education and LP-UTR Schools

The Learning Partners' Framework included a focus on cycles of collaborative learning across school boundaries. The shared emphasis on a cyclical inquiry process was one of the elements that enabled the LP and UTR models to dovetail so easily. LP triads began their collaboration by defining a Learning Focus Area (LFA) to guide their partnership work. Each school team, made up of principals, assistant principals, and teachers, then identified a local Problem of Practice (PoP) linked to the overarching LFA, which guided their team activities in-between cross-school visits. The inquiry work that teams engaged in was documented in a Cycles of Learning (CoL) template, designed to foster continuous improvement as school teams describe how they planned, implemented, reviewed, and shared their inquiry work.

For LP-UTR, the shared LFA in Year 1 was "teacher development," which shifted to "strengthening the instructional and leadership development of resident and mentor teachers" in Year 2. Year 3 continued to build on the shared LFA from Year 2 with some LP-UTR triads adding an instructional focus, such as writing or literacy, that further tied school to triad goals. Teachers and administrators generally liked the idea of a shared goal that was still locally defined, and discussions framed around an instructional need or problem of practice they could realistically address. A similar challenge, at the project level, was how partners could make the combined work more uniform, while at the same time respecting triad autonomy and school variations.

Other changes in existing program structures, especially between Cohort 1 to Cohort 2, not only gave schools and triads more autonomy but also led to more school investment in triad work. The operational change of sharing the month-by-month syllabus and having host schools lead professional development (PD) sessions, for example, gave schools both responsibility and ownership.

New Visions for Public Schools

New Visions had, over the course of UTR and MASTER, refined program recruitment and admissions procedures, as well as a number of activities and structures around inquiry and formative assessments, including monthly reflective seminars and coaching sessions focused on inquiry. New Visions had also developed a suite of assessment tools to track residents' progress in teaching practice (using the Danielson Framework), on lesson and unit planning and two Defense of Learning presentations.

In support of formative assessment practices and performance assessment tools there was also a strong emphasis on data-driven practice. The robust use of the process monitoring assessments for observing resident practice continued during LP-UTR, but the assessment data was not always entered into the online platform due to technical issues or timing challenges (rates of data availability for the various assessments are presented in Part 3, p. 37). This was unfortunate for the lost opportunities among stakeholders to avail themselves of the dashboard features when analyzing teacher practice assessment data (Harford & Verdier, 2019). The LP-UTR program periodically heard from various stakeholders that they had experienced technical issues when they attempted to submit assessment data, which was particularly true during the first year of LP-UTR implementation, but then greatly improved during subsequent iterations through the efforts of the New Visions Systems, Data Analytics, and Research team. With each successive cohort of residents, LP-UTR observed some reoccurring

patterns in the levels of submitted assessment data that were largely attributed to waves of increased demands on their time that align with the NYD DOE calendar. The amount of time to complete all the assessments was increased by the fact that, when resident benchmarks were unmet, the residents were required to do revisions (to lesson plans, teaching episodes, etc.) to meet the standards. These revisions were viewed as important and powerful learning opportunities, and the program anticipated the need for revisions because they understood that there is a diverse range of developmental trajectories among novice teachers. Several unsuccessful attempts to meet standards triggered the development of an intervention plan.

Using stakeholder feedback, the LP-UTR program periodically reflected on the particulars about the assessments—the components of practice, how they are scored, the frequency of use, how to support their use—and made modifications as needed. Collectively, these changes had a positive influence on the submission of assessment data given that the overall availability of process monitoring data improved with each cohort.

The program also observed gains in assessment completion through other strategies they employed, such as sending frequent electronic reminders, reminding stakeholders at in-person events, adding submission rate information to the data visualizations within the process monitoring system, periodically presenting aggregated data to the stakeholders, and providing targeted training. For example, the LP-UTR team provided training about the data system to administrators in Cohort 3, and there was a subsequent increase in data gathering by administrators. LP-UTR had provided a summary of data submission to coaches and mentors, and then later witnessed a rise in submissions.

During the project, the program came to understand the importance of setting aside structured time where LP-UTR stakeholders would access their process assessment data in the online system for reflection and action planning (Harford & Verdier, 2019). On these structured occasions, there was an added opportunity for individuals to note when particular assessment data was missing. More frequent feedback loops about data submissions could yield gains in response rates through simply raising awareness of what data is missing. As the project progressed, there was a shift in the coach seminars that had the participating coaches facilitating the learning, developing ownership of the process data by highlighting certain data and engaging their professional learning community in reflecting on it. The program also made iterative improvements to the data visualizations in Tableau during the project to facilitate understanding of the data.

During a focus group at the end of LP-UTR's third year of implementation, a project team member saw the improved "data systems and more rigorous research" as one of the most valuable features of LP-UTR.

Hunter College

Hunter College continued to play the role of connecting coursework to classroom practice through the LP-UTR program. In an interview with three Hunter College team members, they described the supports Hunter College faculty provide the resident. In all Hunter College alternative pathways to teaching programs, which includes LP-UTR, residents are provided dedicated staff resources specifically assigned to support alternative certification students from application to graduation that are not there for students going through a traditional pathway. While

they agreed, “there’s no one best pathway” to become a teacher, there is a benefit to the level of support that is provided through the LP-UTR program and the opportunities residents have to reflect and connect what they are doing in the classroom with what they are learning through their coursework.

Although Hunter College already had a TESOL program, a second change affecting Hunter and LP-UTR was the addition of TESOL as a subject area for LP-UTR. The first LP-UTR cohort of residents in the TESOL program had a high attrition rate with five of the seven residents exiting the program early. Hunter staff indicated, “whenever we add a new subject area, there are some wrinkles that develop and particularly with TESOL there were some growing pains around how it is a K-12 certification which is different from all of the other subjects.” For Cohort 3, additional supports were made for both TESOL and Special Education residents to better “support their coursework and placement experience.”

While not a lot changed in terms of coursework and preparation for LP-UTR residents at Hunter over the nine years of UTR, there were some small “incremental changes,” that helped shape the program. One of the incremental changes for Special Education residents was the timeframe for the Learning Lab course. The coursework itself did not change, but instead of being offered during the fall and spring of the residency year, it was split between the spring semester of the residency year and the fall semester of the first year as a teacher of record. The Hunter team noted the feedback from residents was that the Learning Lab was “overwhelming” while finishing their residency year and preparing for the edTPA assessment. The only challenge to shifting the timing of the course was that “residents could potentially be in two schools if they weren’t hired by their host school.”

Program partners saw a need for increased collaboration between Hunter College faculty and mentors. In response, faculty members were invited and a few attended a mentor professional development session hosted by New Visions. This was a time where faculty and mentors were able to hear from each other and start conversations about how to bridge the two components of the program. An outcome of this collaboration was that Hunter College provided a dedicated faculty point person to Special Education and TESOL residents to help the residents better navigate the application of their coursework into their work in the classroom.

Again, responses from LP-UTR participants confirmed the positive impact of changes. In focus groups, residents were positive about the supports provided by Hunter College, in particular the support provided by the Assistant Director in the Office of Partnership Programs. Residents found it helpful to have a “point person” who was able to help find answers when it came to Hunter College structures and systems. Special education and TESOL residents also noted the usefulness of the faculty supports in their content areas.

Residents’ also attributed much of their growth as teachers to their Hunter coursework, which, based on their ratings, contributed to around 20% to 30% in most skill development categories; residents rated help in developing effective lesson plans as Hunter’s most significant contribution. (See Table 20, p. 51)

THE LP-UTR COACHING MODEL

LP-UTR included three types of coaches, all of which provided support for the mentor-resident pairs, which included observing and providing feedback to residents in partnership with the mentor; assisting with monthly resident seminars; engaging in mentor support; facilitating triad work; and acting as a liaison between the program and school administration.

New Visions Coaches

From the start of UTR, New Visions coaches have pushed into schools, providing support for both residents and mentors, along with guidance for the school. Sharing that role with school-based and district coaches, the New Visions coaches, while also supporting various schools, took on additional responsibilities that included "aspects of program development and implementation." The New Visions coaches provides support to all residents around the resident hiring process. Working up to the spring, the New Visions coaches hosted workshops on building a resume and provide support with mock interviews and demo lessons. The New Visions coaches were a liaison between hiring schools and residents which included "maintaining a bank of resumes to be shared with our prospective schools and schools who were interested in specific content areas to hire." In addition, the New Visions coaches planned and facilitated monthly seminars for residents, mentors and coaches.

School-Based Coaches

School-based coaches worked closely with residents and mentors at the school where they were teaching. While school-based coaches agreed their role "was not as hands-on from the day-to-day work of a mentor," they were always available when residents and mentors needed some guidance. Coaches conducted bi-weekly to monthly observations and assessed the residents using the Danielson rubric. In addition, school-based coaches helped facilitate the monthly resident seminars at New Visions which helped support the residents through the inquiry projects that they were doing for the Defense of Learning and guided the Learning Partners inter-school collaborations, including the triad visits, work around the Learning Focus Area (LFA) and the Cycles of Learning (COL). School-based coaches participated in mentor seminars and provided guidance and support to mentors throughout the year.

District Coaches

District coaches, similar to New Visions coaches, bring an outside perspective to the LP-UTR program. District coaches also have a deep understanding of the work of the Learning Partners program that they contribute to school-team and triad work. However, similar to all other coaches, district coaches observe residents and support resident and mentor development.

Changes in the Coaching Model

Developing the model, developing the coaches. Over the three years of the LP-UTR program, coaching continued to draw from the expertise of the Department of Education's Learning Partners staff, New Visions staff experienced with the residency, and school-based coaches.

Through the three cohorts of residents, some of the elements of coaching have remained consistent, and others evolved with the needs of the program. By Cohort 3, coaches had more structures around their responsibilities with residents and mentors. During an interview a coach noted, "there were models and phone calls that helped the coaches prepare for meetings, there was a syllabus for meetings when there hadn't been before." Most of the coaches saw these new structures as a way to take on more of the "logistical stuff" which gives more time for mentors to focus on supporting the residents.

Across the three cohorts some of the coaching activities varied to better support teachers and schools. One change was in the number of school-based coaches. By Cohort 3, the number of school-based coaches tripled compared to the first year of LP-UTR when school-based coaches were only in schools that were identified as a host school. (See Table 12.) By having school-based coaches in partner schools, it makes the "model more sustainable" and gives more leadership opportunities to the partner schools.

TABLE 12. NUMBER OF COACHES, BY TYPE

Type of Coach	Year 1 Coaches	Year 2 Coaches	Year 3 Coaches
New Visions	2	2	2
School-based	3	6	9
District	1	1	1
Total	**6**	**9**	**12**

In focus groups with residents, one of the key discussion topics addressed concerned the advantages and disadvantages of the different types of coaching—New Visions coach, school-based coach, or district coach.[3] Residents who had a school-based coach valued the coach's accessibility and ability to watch them teach. One resident explained, "my coach also taught the same content as me so I saw [the coach] every day. [The coach] was always available to answer any questions and I got to see my coach teach." The residents who had a New Visions or district coach appreciated being able to communicate via email. One resident said, "I actually in a way liked it because I didn't have the extra observation every time this person came into the classroom."

Coaches themselves found the advantages of residents having school-based coaches as "having someone at the school as a first point of contact to resolve issues." Their opinion was that school-based coaches have a greater idea of the context of what's happening at the school on any given day, with a better understanding of the day-to-day operations of the students, faculty, and administration. They also expressed that the advantage of having a New Visions or district coach was that those coaches might tend to have a greater degree of objectivity and the ability to have "a dedicated role with time to really look for additional resources." Coaches reported that the disadvantage for a school-based coach is not being able to "escape" or finding that they have a "lack of privacy" because residents and mentors sometimes forget that the school-based coaches are juggling responsibilities such as their own teaching, students, etc., whereas the New Visions and district coaches are only in the schools at select times, which makes it harder for residents and mentors to reach out when a problem arises during the school day.

[3] For the purpose of the focus groups and interviews, the New Visions and district-based coaches were lumped into one group, as they both were seen by residents and mentors as coaches not based at the school.

Another change in the coaching model was to focus on coaching support. Between Cohorts 1 and 2 “instead of just saying okay we’re just going to hold coach seminars, we also implemented coach individual check-ins,” said Rachelle Verdier, Deputy Director, UTR. Coach seminars continued to be an opportunity for all coaches to gather and share feedback on what was happening at each school, while individual coach check-ins provided an opportunity for coaches to grow individually where each school-based coach was assigned a “buddy” in order to help build the coach’s skills. Verdier went on to say, “I thought it was really effective because it gave us an opportunity to support the coaches, but also a moment in time where we can collect real time information about how the program is running on the ground.” The coaching seminars also focused on Eliciting and Interpreting Student Thinking (EIST) one of the Teaching-Works High-Leverage Practices. The UTR project team selected EIST as a focus in Professional Learning Communities through the National Center for Teacher Residencies (NCTR), designed to support network programs in improving coaching around high-leverage focus practices in the residency model.

PART 3: GROWTH IN TEACHERS' SKILLS & CONFIDENCE

As noted throughout this report, the overarching goal of LP-UTR was to ensure the growth of both novice and experienced teachers. That had been the goal of the previous residency programs, and structures were in place to ensure and examine residents' and mentors' growth—but adding the cross-school component presented a new set of challenges and prompted changes.

During the first year of LP-UTR implementation, there was a clarification of the roles and expectations of mentors and residents. It was decided that specific program aspects such as the cross-school visits would focus on mentor development. Mentors would, in turn, use key takeaways from their visits to inform their work with their residents. Other aspects of the program, including Hunter coursework and residents' clinical training, helped ensure their growth.

This section of the report shares evidence of how both mentors and residents grew through their LP-UTR participation. Our analyses addressed the following questions:

- How did LP-UTR use performance-based assessments to monitor and support resident growth?
- How confident were mentors and residents around instructional and assessment practices?
- Did confidence levels change over the course of each year or from cohort to cohort?
- To what did residents and mentors attribute confidence?

- Classroom performance and capstone Defense of Learning (DoL) assessment data provided evidence of residents' growth trajectories.
    - Average scores on the Danielson-based rubric, administered 4 times during the year, tended to show increasing trajectories in residents' ability to plan instruction, manage classrooms, and use best practices, especially evident when comparing baseline and end of year scores.
    - Residents tended to score highest in creating an environment of respect and rapport; they also did well in classroom management relative to other component averages. They showed the most improvement in using questioning or prompts and discussion and in using assessment in instruction.
    - The overall growth trends were true for both mentor- and coach-reported scores.
    - DoL performances provided further evidence residents' development and program progress: in each successive cohort a greater percentage of residents met or exceeded the benchmarks.
- Mentors developed confidence in their instructional practice and assessment skills through LP-UTR.
- **Mentors indicated they developed skills through teaching and interacting with school colleagues, while residents attributed skill development across program components—New Visions professional development, coaching, Hunter coursework, and teaching and interacting with colleagues.**

PERFORMANCE ASSESSMENTS OF RESIDENTS' GROWTH

During the year of clinical experience, the LP-UTR program—like many teacher residencies (Guha, Hyler, & Darling-Hammond, 2016)—utilized multiple assessment instruments focused on the residents' practice. These performance-based assessments include a rubric consisting of six components from the Danielson Framework (2013), the Defense of Learning assessment, the Lesson Design Rubric, the Unit Design Rubric, and the Professionalism Rubric. Together, these assessments are used to formatively monitor and support the learning of the residents (Wiliam, 2011) through clarifying expectations for good teaching practice, defining how effectiveness is evaluated, and providing a shared focus for frequent feedback.[4] Within the program, the information that these assessments provided for instructional development was commonly known as process monitoring data. To capture this data in order to provide stakeholders a process monitoring system, LP-UTR uses an online platform, Tableau, which includes data visualizations and spreadsheets. The process monitoring system is intended to support residents, mentors, coaches, and administrators in reflecting on the process data to determine learning needs, provide and discuss feedback (also captured in Tableau), and make plans of action. Because both LP-UTR mentors and coaches assessed and provided direct support to the residents, they worked together to align their feedback to residents, and this was less of a challenge when the coaches were school-based rather than from New Visions or the district.

Although the chief purpose of the process monitoring assessments was to guide and evaluate residents, residents, mentors, and coaches noted that they were often focused on compliance, or completing the assessments at the designated times, rather than the utility of the instruments for supporting resident development. This feeling among some residents and other educators may have stemmed in part from actual or self-perceived shortfalls in data literacy and data-driven decision-making competencies, or on frustrations with problems in the data system itself. It is also important to acknowledge that the school system of course places additional measurement-related demands on the educators and students, which can influence how the program assessments are experienced.

There were instances when coaches' and mentors' collection of formative assessment data led to changes in program structures. For example, the project team recognized that some residents struggled with lesson planning throughout the year. In response, the program team in conjunction with Hunter College faculty, created a lesson design template for Cohort 3 residents to use. This template enabled Hunter Faculty and program staff to better assess residents' preparedness to design lessons. While residents' scores on the lesson design rubric did not change from Cohort 2 to 3 (see Table 15 on page 41), the template helped to clarify expectations for lesson planning across the cohort and provided coherence across program stakeholders.

[4] In addition, LP-UTR used a rubric for monitoring mentor competencies and another rubric for coach competencies. Target scores were identified for the end-of-the-year use of each rubric. These provided the program additional feedback loops to support stakeholders and track program implementation.

The Suite of Process Monitoring Assessments

A listing of the primary assessments of resident practice is shown in Table 13. All five of the listed assessments were used with each cohort of residents. During the project, the LP-UTR team periodically made some changes to the assessments, which largely involved adjustments to some of the benchmarks. Overall, most modifications were made after Cohort 1, and the suite was almost entirely uniform across Cohorts 2 and 3[5]. For cohort-specific details of the assessment timelines and corresponding benchmarks for these assessments, see Tables 40–50 in Appendix A.

TABLE 13. KEY PROCESS MONITORING ASSESSMENTS, BY COHORT

ASSESSMENT	ASSESSEE	ASSESSOR
Danielson-Based Assessment	Resident	Mentor, Coach, + Administrator
Defense of Learning Assessment	Resident	Coach (w/ Mentor input)
Professionalism Rubric	Resident	Mentor + Coach
Lesson Design Rubric	Resident	Mentor + Coach
Unit Design Rubric	Resident	Coach

The program goal was to attain a 90% rate of submission for each of the process monitoring assessments; but as stated above, the program stakeholders were challenged in submitting all their assessment data. Table 14 provides a breakdown of the data availability in the process monitoring system for the various assessments by cohort and time. The overall trend in data availability from cohort to cohort improved from year to year. Generally, data availability was highest during some of the earliest administrations of a particular assessment and often decreased towards the end of the year.

For an overview of resident performance data on the key process monitoring assessments (see Table 14), which lists the numbers of residents meeting the benchmarks for each assessment by cohort and time (based on submitted scores). In the following subsections, we discuss the data for the two process monitoring assessments deemed most important by LP-UTR leaders, the Danielson-based assessment and the Defense of Learning.

[5] One of the larger shifts made to the process monitoring data in between cohorts was to begin collecting observation data during the summer preceding the residency school year to provide an earlier set of baseline data for the stakeholders than was initially available. Program leaders would monitor the submission rates and scores to ensure the resident-mentor dyads set off on the right foot. The program found this assessment shift to improve their support for the residents' learning.

TABLE 14. ATTAINMENT OF ASSESSMENT BENCHMARKS BY COHORT AND TIME

Assessment	C2 (48 enrollees 41 residency completers)		C3 (48 enrollees, 42 residency completers)	
	Number of Residents that Met Benchmark	Percent of Residents with Data that Met Benchmark	Number of Residents that Met Benchmark	Percent of Residents with Data that Met Benchmark
Danielson-Based Assessment	T1 N/A	N/A	T1 N/A	N/A
	T2 43 of 43	100%	T2 39 of 44	89%
	T3* 21 of 21	100%	TA* 17 of 17	100%
	T4 34 of 40	85%	T3 19 of 22	86%
	T5 29 of 35	83%	T4 29 of 40	73%
Defense of Learning Assessment	T1 28 of 42	67%	T1 34 of 44	77%
	T2 30 of 41	73%	T2 29 of 34	85%
Professionalism Rubric	T1 41 of 48	85%	T1 45 of 47	96%
	T2 38 of 45	84%	T2 43 of 45	96%
	T3 N/A	N/A	T3 32 of 35	91%
	T4 32 of 32	100%*		
Lesson Design Rubric	T1 N/A	N/A	T1 N/A	
	T2 30 of 41	73%	T2 33 of 45	73%
	T3 19 of 33	58%	T3 16 of 29	55%
	T4 0 of 4	0%		
Unit Design Rubric	T1 24 of 37	65%	T1 30 of 35	86%

N/A = not applicable, no benchmark established
* Scores from school administrators only.
Note: The number of residents with available scores was partially affected by program withdrawals in each cohort.

LP-UTR Danielson-Based Rubric

The Danielson 2013 Framework for Teaching (Danielson, 2013) is a rubric used for supporting teacher development and for the evaluation of teacher practice. The LP-UTR program used a rubric that included six of the 22 components from the Danielson 2013 Framework, spanning three of the four domains in the Framework. The LP-UTR Danielson-Based Rubric also contains several additional items (see Table 15). The limited number of components, or competencies, are driven by how teacher practice is measured in New York City. The NYC DOE found over time that given just several observation sessions, it was not effective or practical for teachers and administrators to be focused on a wide array of competencies. The LP-UTR partners elected to use a Professionalism Rubric that was developed during earlier iterations of the UTR program (see description in Appendix A) in lieu of the Domain 4 of the Danielson Framework because the Professionalism Rubric was designed to provide more detailed and actionable feedback about elements of professionalism that were more germane to the LP-UTR model.

TABLE 15. LP-UTR DANIELSON-BASED RUBRIC

Domain*	Component**
1. Planning & Preparation	D1e: Designing Coherent Instruction
2. Classroom Environment	D2a: Creating an Environment of Respect and Rapport
	D2d: Managing student behavior
3. Instruction	D3b: Using questioning / prompts and discussion
	D3c: Engaging students in learning
	D3d: Using assessment in instruction

* The fourth domain in the Danielson Framework, Professionalism, is not included in this rubric.
** The numbering of the components aligns with the numbering/lettering of the component as it is represented within the 2013 Danielson Framework, which has 22 components: Components 1a–1f, components 2a–2e, components 2a–3e, and components 4a–4f.

The rating system in the LP-UTR Danielson-Based Rubric for the six component items of the Danielson Framework is aligned to the Measures of Teacher Practice Rating system (see Table 16).[6] For each rating item, observers record comments. In addition, there is one other rating item, "overall outcomes, based on student data," which is rated using a unique response scale: 1 (fewer than 50% of students met objective), 2 (50-75% met objective), 3 (75-90% met objective), 4 (90-100% met objective), and N/A because the assessor was not able to engage in analysis of the student work. Three additional open-ended items provide opportunity for observer to summarize overall strengths, overall areas for growth, and major next steps to improve practice. Mentors and coaches engaged in norming exercises at each quarterly mentor PD session and at coach seminars.

One of the short-term outcomes identified in the LP-UTR logic model is that residents will meet benchmarks on the Danielson-based rubric. The LP-UTR team established benchmark scores for each administration of the Danielson-based rubric. The benchmarks were slightly modified beginning with Cohort 2.

TABLE 16. SCORE ALIGNMENT BETWEEN THE DANIELSON FRAMEWORK (2013) AND NYS DEPARTMENT OF ED

Score	Danielson	NYS
1	Unsatisfactory	Ineffective
2	Basic	Developing
3	Proficient	Effective
4	Distinguished	Highly effective

The broad trends observed across cohorts, in resident performances on the Danielson-based rubric showed how residents developed as teachers over the course of their residency.

- The average scores on the Danielson-based rubric, which was administered by mentors and coaches four times during the residency year, tended to show positive growth in residents' ability to effectively plan instruction, manage their classroom, and use best practices. This was especially evident when comparing the baseline and end of year scores.

[6] The evaluation of the original UTR project (Sloan & Blazevski, 2015) noted that in the early phase of the project mentors had felt that there was a lack of specificity in the simple 4-point scale and that residents should not be receiving 1s. In response, Rockman, in collaboration with project team, developed a response scale that enabled some more nuanced scoring within each of the four levels. By adding a low, middle and high score to each level (i.e., low 1, 1, high 1, low 2, 2, high 2, etc.), they maintained the ability to readily convert scores into the standard 4-point system used in district evaluations.

- Percentages of residents meeting progress benchmarks ticked down as thresholds increased. Cohorts 2 and 3 showed a decline of roughly 15 percentage points from the first to final benchmark.
- The overall trend in teaching practice growth was true for both mentor- and coach-reported scores on the Danielson-based rubric.
- Generally, scores from both program coaches and mentors were in close alignment, supported by the norming activities they periodically engaged in over the year.
- The average scores for each of the individual components of the Danielson-based rubric increased over the course of each residency (see Figures 9 and 10).
- In comparing average performances for individual components, at both the start and end of the residency, residents tended to score the highest in creating an environment of respect and rapport; they also did well in classroom management relative to the other component averages.
- Residents tended to show the most improvement in using questioning or prompts and discussion while teaching and also in using assessment in instruction. On average, they had scored lower on these two components at baseline relative to the other components.

FIGURE 9. AVERAGE COMPONENT-LEVEL SCORES ON THE LP-UTR DANIELSON-BASED RUBRIC, BY TIME: COHORT 2

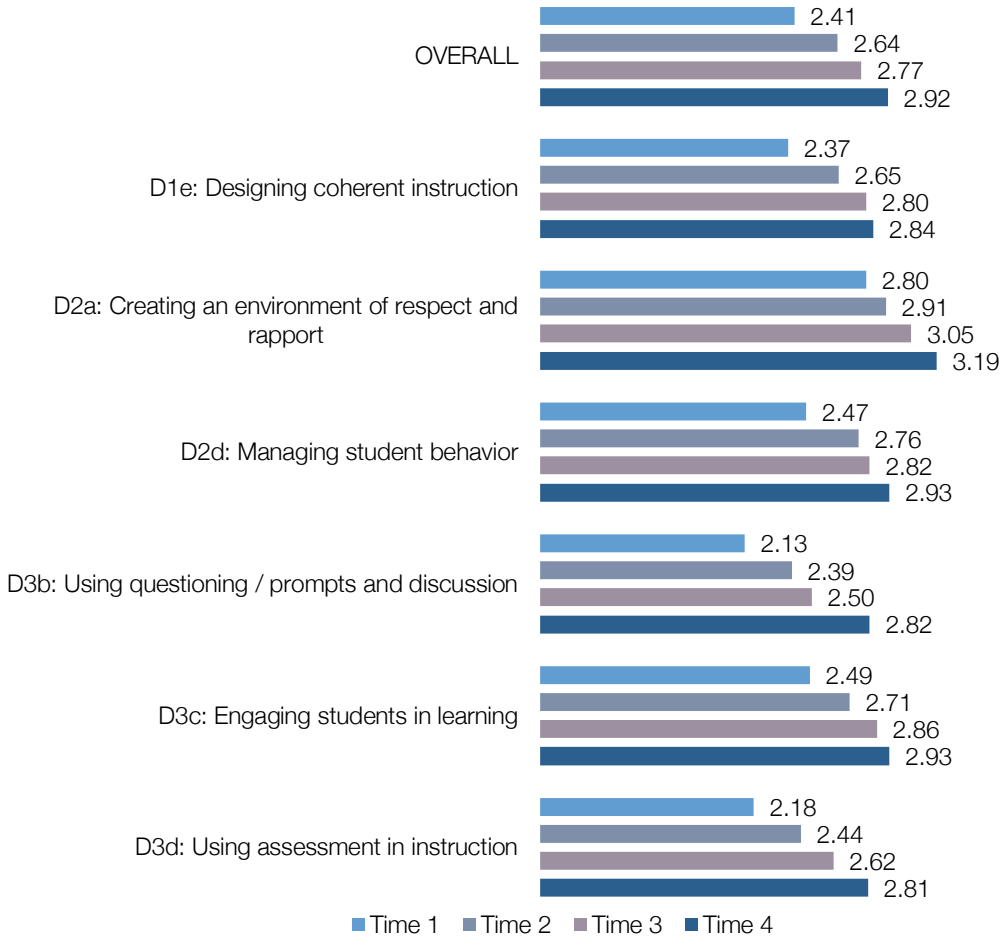
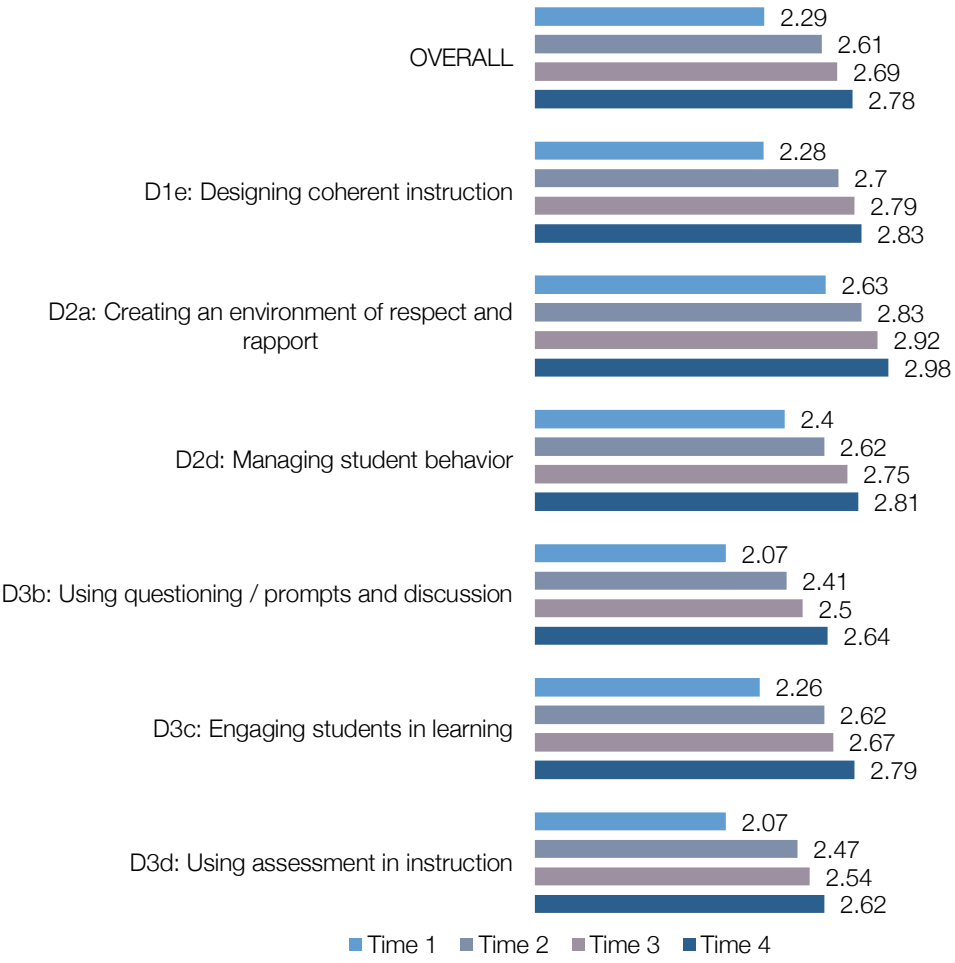


FIGURE 10. AVERAGE COMPONENT-LEVEL SCORES ON THE LP-UTR DANIELSON-BASED RUBRIC BY TIME: COHORT 3



Defense of Learning

With the Defense of Learning (DoL) assessment, residents explain and demonstrate to an audience of observers the application of theory-based interventions using the inquiry cycle "with a specific set of students, in a specific context, with specific instructional goals" (Harford & Verdier, 2019, p.26). This assessment is conducted at two time points. The first DoL is a midyear practice run intended to familiarize the resident with the experience in front of a smaller team of observers and a smaller group of students. The program also valued this first DoL because it was a predictor of program success and could be used to signal the need for an intervention plan for the second half of the residency. The second DoL is the actual capstone assessment, conducted near the end of the school year, with the full team of program officers, mentors, administrators, colleagues, and faculty.

TABLE 17. DOMAINS AND COMPETENCIES OF THE DEFENSE OF LEARNING RUBRIC

Domain	Competency
Student Learning	Initial Student Mastery Level (Baseline)
	Summative Assessment Results/ Student Learning Outcomes
	Artifacts of Student Learning
Theory of action	Selection of focus skill/ content
	Establishing Goals and Objectives
	Initial plans for instruction
	Formative assessments
Using data and resources to drive practice	Revisions to Instruction
	Utilizing resources to improve practice
	Reflections on learning and professional growth
Preparation and defense	Materials and preparation
	Oral defense

For the Defense of Learning, there were 12 variables or competencies each rated on the same 4-point scale as the other assessments. Weighting was used to emphasize three different competencies. Specifically, summative assessment results/student learning outcomes, formative assessments, and revisions to instruction were each doubled in score. The maximum total score was thus 60.

The Defense of Learning performances of Cohorts 2 and 3 provided further evidence of how residents developed during the residency. In addition, in each successive cohort a greater percentage of residents met or exceeded the established benchmarks on the Defense of Learning.

Cohort 2

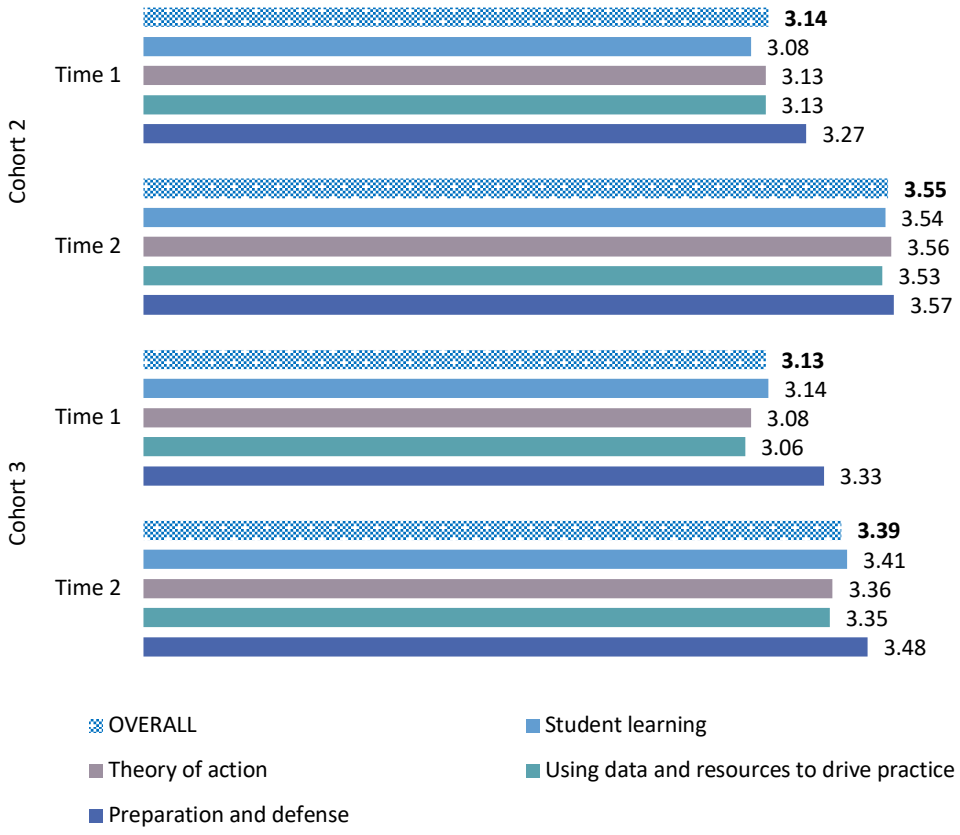
- The availability of DoL scores in Cohort 2 was high at both timepoints (88% and 85%). While the instrument remained unchanged for Cohort 2, the benchmarks increased slightly from 42 to 45 for Time 1 and from 45 to 48 for Time 2. These higher benchmarks reflected the new structure of the clinical residency seminar. The revised structure required residents and coaches to participate in reflective seminars that involved program directors and coaches facilitating small groups of residents in learning the content of the session.
- Similar to Cohort 1, residents overall showed improvement in their DoL scores and subscores from the midyear and end of year administrations. However, over one-quarter of the residents with submitted scores did not attain the final benchmark for the year.

Cohort 3

- For Cohort 3, the availability of data at Time 1 for the DoL assessment was strong (92%), but less data was available at Time 2 (71%).

- The percentage of residents who met the established benchmark improved from Time 1 (77%) to Time 2 (85%).

FIGURE 11. DEFENSE OF LEARNING SCORES BY COHORT AND TIME



TEACHERS' SELF-ASSESSMENTS OF CONFIDENCE

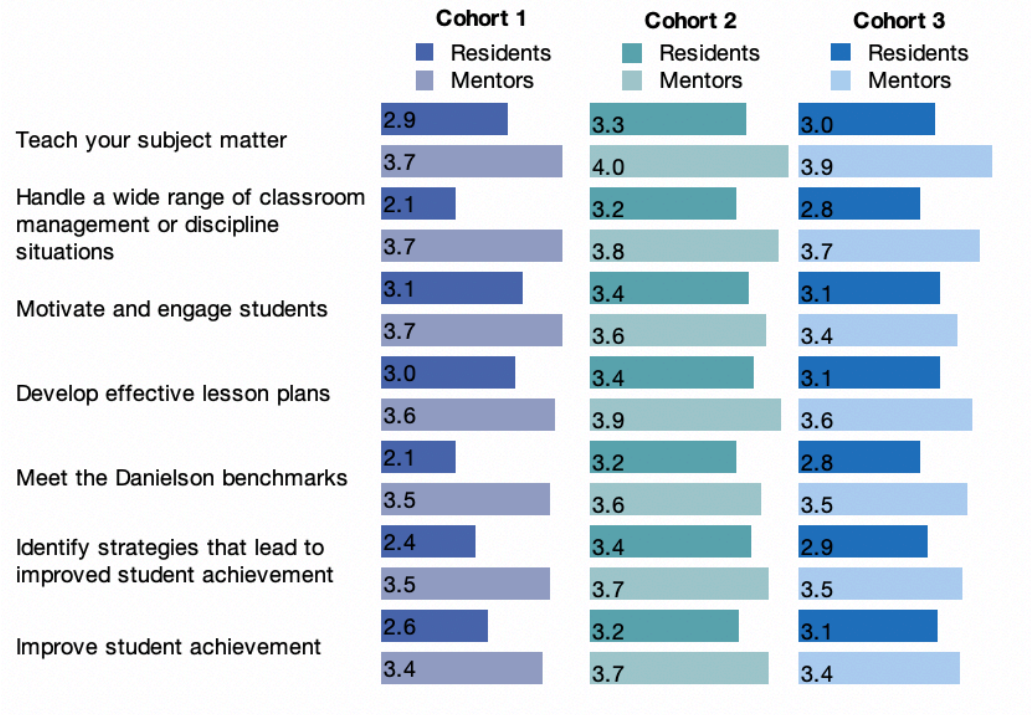
Residents' and Mentors' Confidence

An important outcome goal of the LP-UTR program was to help residents and their mentors increase their confidence in instructional and assessment practices. **End-of-year survey data from mentors and residents in all three cohorts indicated that both groups of teachers were confident in a range of instructional skills. Although, not surprisingly, mentors reported significantly higher levels of confidence than residents, the margins narrowed with Cohorts 2 and 3, and residents rated themselves as "confident" in most skills and practices (see Figure 12).**

- On average, mentors' self-assessments trended toward "very confident" in most skills, closer to a four than a three on a four-point scale (1=not at all confident and 4=very confident).
- Residents across all three cohorts expressed the greatest confidence in their abilities to motivate and engage students (Cohort 1 *M*=3.1, Cohort 2 *M*=3.4, Cohort 3 *M*=3.1) and develop effective lesson plans (Cohort 1 *M*=3.0, Cohort 2 *M*=3.4, Cohort 3 *M*=3.1).

- In each cohort, residents were least confident in their ability to handle a range of classroom management situations (Cohort 1 *M*=2.1, Cohort 2 *M*=3.2, Cohort 3 *M*=2.8) and to meet the Danielson benchmarks (Cohort 1 *M*=2.1, Cohort 2 *M*=3.2, Cohort 3 *M*=2.8). Resident confidence in these areas increased a full point between Cohorts 1 and 2 but then decreased slightly from Cohort 2 to Cohort 3.
- Compared to their peers in Cohorts 1 and 3, Cohort 2 residents had the highest levels of confidence in each of the seven instructional practices categories. The scale average for Cohort 2 residents was 3.3 compared to an average of 2.6 for Cohort 1 and 3.0 for Cohort 3.

FIGURE 12. INSTRUCTIONAL PRACTICES CONFIDENCE



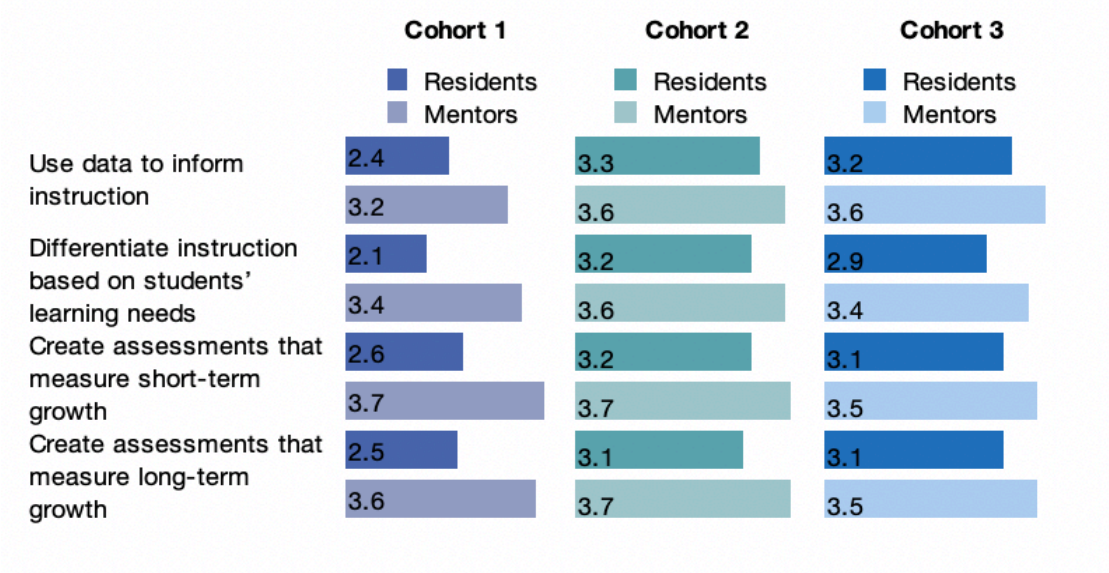
Residents of Cohorts 2 and 3 were confident in their assessment and data use practices, while Cohort 1 was markedly less confident. Mentors in all three cohorts reported higher levels of confidence in their assessment skills than residents.

- Cohort 2 and 3 residents reported higher confidence levels in their assessment skills than Cohort 1 residents. Across the four assessment skills, the average confidence rating for Cohort 1 residents was 2.4, while for both Cohorts 2 and 3 the average was 3.1.
- Cohort 2 mentors were the most confident in their assessment skills with an average scale score of 3.7, followed by Cohort 1 with a score of 3.5, and Cohort 3 with a score of 3.4. Mentors and residents in Cohort 3 had the smallest margin of difference in their confidence ratings of their assessment skills.

During focus groups, residents also indicated they have grown as teachers through their LP-UTR experience. Residents noted how they have "become more efficient," and been able to provide instruction that is more "student centered." While residents and mentors noted areas of growth, both agreed that residents still need to work on classroom management, a skill that takes time and practice to build upon.

Together, this data supports expectations that the novice teachers (residents) would feel less efficacious than the more experienced (mentor) teachers, and it also suggests that residents (particularly in the latter cohorts) developed confidence in a range of teaching practices through their LP-UTR participation.

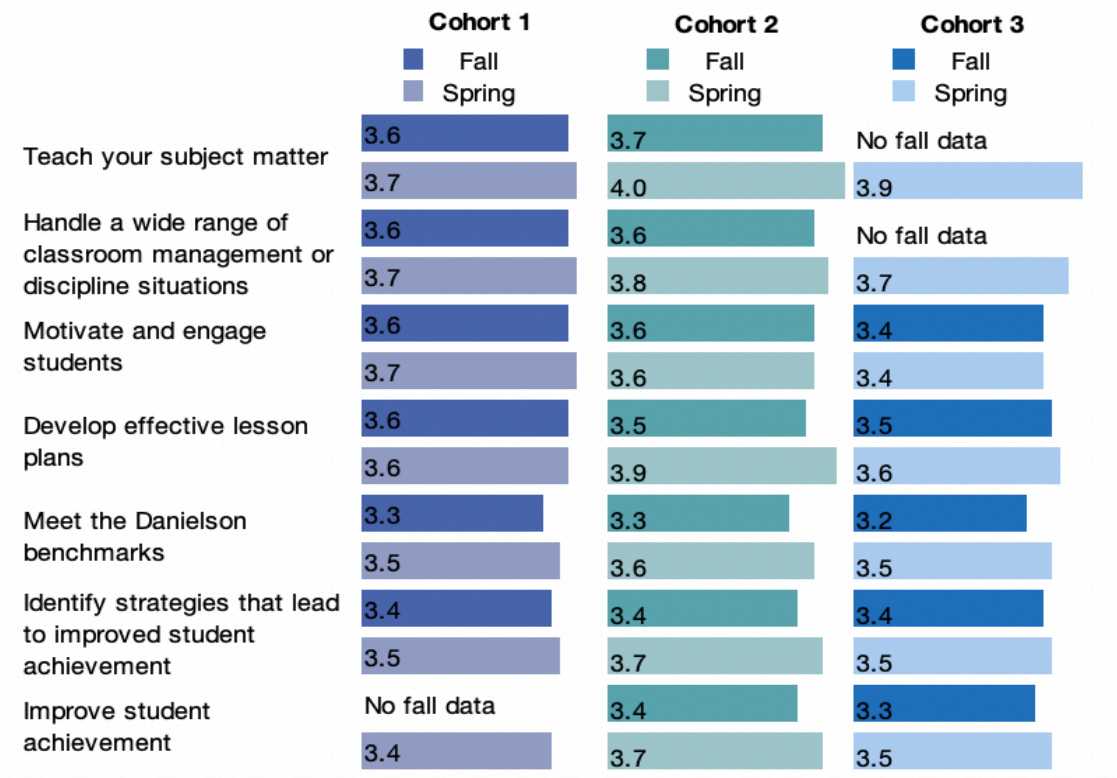
FIGURE 13. ASSESSMENT PRACTICES CONFIDENCE



Changes from Fall to Spring in Mentors' Confidence

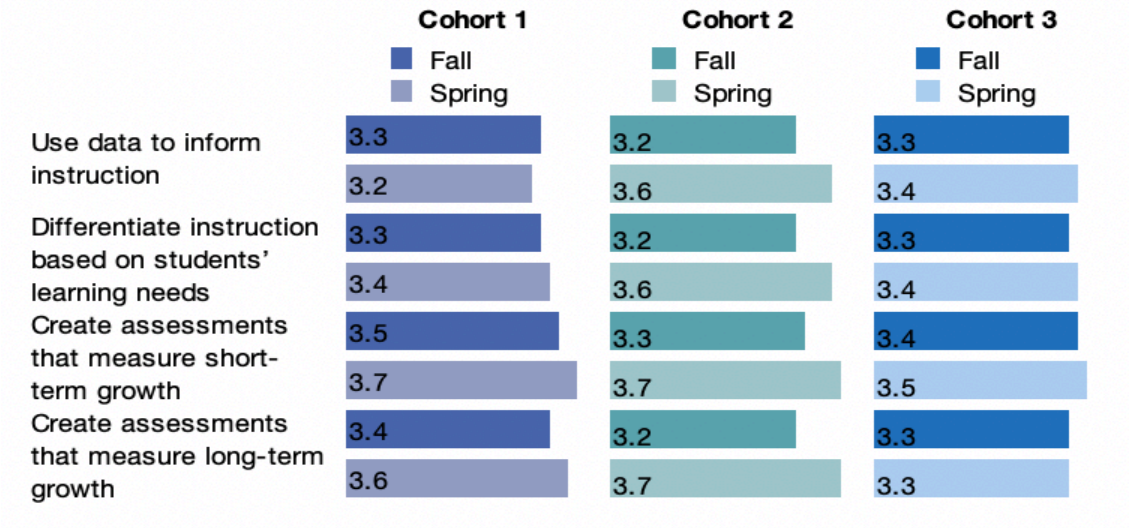
To assess evidence of learning growth during the school year, mentors provided responses on both baseline and end of year surveys. **In each cohort, the pre-post results showed a slight increase in mentors' confidence in all but one of the instructional categories.** Cohort 2 mentors demonstrated the largest gains in confidence in their ability to develop effective lesson plans with an increase of 0.4 points from fall to spring. (See Figure 14.)

FIGURE 14. INSTRUCTIONAL PRACTICES CONFIDENCE, CHANGES FROM FALL TO SPRING



Likewise, in each cohort and across almost all categories, mentors' confidence in their assessment practices increased slightly from fall to spring. Cohort 2 mentors showed the greatest gains in confidence with an increase of at least four tenths of a point in each of the four categories (see Figure 15).

FIGURE 15. ASSESSMENT PRACTICES CONFIDENCE, CHANGES FROM FALL TO SPRING



WHERE MENTORS AND RESIDENTS GAIN SKILLS AND CONFIDENCE

Across the instructional practice and assessment skills categories, mentors from all three cohorts consistently indicated that they primarily developed their skills through teaching and interacting with school colleagues. For all skill categories mentors attributed at least, and in some cases far greater than, 50% of their skill development to teaching and interacting with school colleagues. In the first year of the program mentors expressed that the New Visions PD and support contributed little to their skill development with four categories contributing zero percent. By Year 3 there was a significant increase in the percent of skill development across all categories attributed to New Visions. Skill development from LP-UTR cross-school sharing was similar across cohorts and ranged from 0% to 17% across categories. The contributions of LP-UTR coaches to mentor skill development ranged from 2% to 17% and were greatest among mentors in Cohort 3 for four of the seven categories. (See Table 18.)

TABLE 18. WHERE MENTORS DEVELOP SKILLS AND GAIN CONFIDENCE

		New Visions PD, support	LP-UTR coach	Teaching, School colleagues	LP-UTR cross-school sharing
Meet the Danielson benchmarks	Year 1	15%	15%	65%	5%
	Year 2	21%	17%	58%	4%
	Year 3	22%	12%	59%	7%
Teach your subject matter	Year 1	0%	6%	78%	17%
	Year 2	13%	5%	73%	10%
	Year 3	17%	7%	62%	14%
Develop effective lesson plans	Year 1	0%	7%	87%	7%
	Year 2	17%	7%	67%	10%
	Year 3	20%	13%	60%	8%
Motivate and engage students	Year 1	0%	6%	82%	12%
	Year 2	19%	2%	69%	10%
	Year 3	11%	14%	65%	11%
Handle a range of classroom management or discipline situations	Year 1	0%	7%	93%	0%
	Year 2	3%	6%	91%	0%
	Year 3	13%	13%	64%	10%
Identify strategies that lead to improved student achievement	Year 1	14%	14%	67%	5%
	Year 2	19%	15%	52%	15%
	Year 3	27%	9%	55%	9%
Improve student achievement	Year 1	17%	11%	72%	0%
	Year 2	15%	9%	62%	15%
	Year 3	23%	9%	55%	14%

Source: REA Annual Surveys

While mentors primarily attributed their skill development to teaching and interacting with school colleagues, residents attributed their skill development more broadly across program components. Residents self-reported gaining skills from teaching and interacting with their school colleagues which ranged from 28% to 56%. Working with their LP-UTR coach contributed to between 11% and 35% of residents' skill development with the contributions being most pronounced for helping residents meet the Danielson benchmarks. Similarly, New Visions Professional Development and support, which ranged from 0% to 26%, was highest for contributing to residents' ability to meet the Danielson benchmarks. Hunter Coursework contributed to around 20% to 30% in most skill development categories; residents rated help in developing effective lesson plans as Hunter's most significant contribution. LP-UTR cross-school sharing had very little effect on resident skill development and confidence. Residents in Cohort 1 indicated that LP-UTR cross-school sharing did not contribute anything to their skill development with ratings of 0% in each category. Attribution to cross-school sharing ticks up in Cohorts 2 and 3, with the most pronounced shift being that 6% of Cohort 2 residents identified LP-UTR cross-school sharing as contributing to their ability and confidence to teach their subject matter. (See Table 19.)

TABLE 19. WHERE RESIDENTS DEVELOP SKILLS

		Hunter Coursework	New Visions PD, support	LP-UTR coach	Teaching, School colleagues	LP-UTR cross-school sharing
Meet the Danielson benchmarks	Year 1	13%	20%	27%	40%	0%
	Year 2	13%	19%	35%	33%	0%
	Year 3	21%	26%	22%	28%	3%
Teach your subject matter	Year 1	25%	6%	25%	44%	0%
	Year 2	26%	9%	20%	39%	6%
	Year 3	24%	12%	17%	42%	5%
Develop effective lesson plans	Year 1	31%	6%	19%	44%	0%
	Year 2	34%	17%	28%	51%	2%
	Year 3	29%	15%	21%	34%	1%
Motivate and engage students	Year 1	29%	7%	14%	50%	0%
	Year 2	22%	8%	17%	47%	5%
	Year 3	25%	13%	21%	38%	4%
Handle a range of classroom management or discipline situations	Year 1	27%	0%	18%	55%	0%
	Year 2	15%	16%	21%	45%	3%
	Year 3	20%	13%	27%	37%	2%
Identify strategies that lead to improved student achievement	Year 1	22%	11%	11%	56%	0%
	Year 2	25%	13%	22%	40%	0%
	Year 3	27%	19%	19%	33%	3%
Improve student achievement	Year 1	22%	11%	11%	56%	0%
	Year 2	23%	14%	22%	41%	0%
	Year 3	23%	14%	23%	36%	3%

Source: REA Annual Surveys

During focus groups conducted at the end of Cohort 3, mentors shared insights not captured in survey ratings. Several indicated that they had grown both as mentors and as teachers, through conducting observations and having conversations with colleagues. Working with a novice teacher has reminded some mentors to reflect on the skills they teach and think about how and why they are teaching the skill in a particular way. One mentor explained, "I like that the residents have focus skills that they're working on…it's just interesting to take an in-depth look at some of the skills I've been teaching and getting a fresh perspective on it." Another mentor added that, "in order to mentor, you have to understand what you're doing well and what you're not doing well," and then be able to describe what you are doing to the resident so they better understand the teacher moves that are being made in the classroom.

Mentors also indicated one area of growth was in building the skills needed to have difficult conversations with colleagues. As one mentor noted, "I've learned to have more difficult conversations. It's one thing because we work with adolescents so we have to have difficult conversations, but it's different when you're working with a grown person."

PART 4: SUPPORT, COLLABORATION, & COLLECTIVE EFFICACY

Parts 1 and 3 of this report focused largely on professional learning. Part 4 shifts to the professional learning community. Based on teachers' ratings of the support they received and levels of participation in collaborative activities, the discussion is in some ways an account of the extent to which LP-UTR dissolved the boundaries and barriers that have long given the impression that education is a "lone wolf" profession, with teachers largely isolated in their work. Those perceptions were altered somewhat by the advent of professional learning communities (PLC), which turned schools into collaborative workplaces (Ronfeldt, et al., 2015), but barriers still existed—even among UTR and MASTER teachers, whose survey responses revealed less collaboration across than within departments and few opportunities for teachers to collaborate *across* schools. (Rockman et al., 2015, 2017).

There was never any doubt among LP-UTR mentors about the general benefits of visiting other schools, seeing how other teachers approached instruction. Drawing on focus group and survey data, this section explores what teachers gained from that experience—how they rated levels of support and opportunities for collaboration, how often they participated in cross-school visits, and whether LP-UTR created a broader school community and sense of collective efficacy. The discussion also shares evidence of how changes made after LP-UTR's first year built both.

- Both residents and mentors felt supported by school colleagues. Residents across cohorts felt supported by their mentors; ratings of support from administrators rose in Cohorts 2 and 3.
- Mentors across cohorts assigned high ratings to administrators' support. Residents had mixed reactions to levels of support, and to integration in their school community.
- Survey feedback from mentors and residents about opportunities for support provided evidence of program improvements made after the initial year of LP-UTR implementation.
- Mentors and residents frequently collaborated with colleagues in their schools to build confidence and skills.
- Mentors reported high levels of participation across program years with 76% of mentors in Cohort 1, 97% in Cohort 2 and 92% in Cohort 3 indicating that they always participate in project activities. Residents were more likely to participate if visits occurred at their school and in their classroom.
- Mentors indicated with higher frequency than residents they built confidence and skills with colleagues in the triad.
- Cohort 2 mentors reported taking on a leadership role, observing classrooms, analyzing student work, using strategies, and collaborating within subject areas more often as part of triad work than within their own schools.
- Cohort 3 mentors reported modeling practices, observing classrooms, and collaborating across subject areas, and using a strategy shared by a colleague more frequently across schools than residents.

SCHOOL SUPPORT FOR COLLABORATIVE ACTIVITIES

LP-UTR, expanding the reach of UTR outside of the New Visions network of schools, allowed residents, mentors, and school administrators to collaborate across schools around the common goal of teacher development. Schools worked in dyad or triad partnerships with one host school and multiple partner schools that collectively found common shortcomings and visited each other's schools in hopes to better understand how to find solutions to bring back to their own school. In the first year of LP-UTR, 11 schools participated in the program; that number grew in Years 2 and 3 with 18 schools. Triads met on a regular basis, based on the triads needs and schedules, and rotated between schools and learning activities. The discussion below shares survey and focus group feedback from LP-UTR teachers about the school-level support and opportunities for collaboration made available to them, noting changes over the course of the project related to program growth and changes in structures.

School-Level Support

There was general consensus that LP-UTR teachers felt supported within their schools, especially by other teachers. There were some differences between how mentors and residents rated support. Survey data across all three cohorts indicated mentors generally perceived higher levels of coaching and school-based support than residents (see Table 20). This difference was especially pronounced in ratings of support from principals and assistant principals. The average mentor principal rating for the three cohorts was *M*=3.65 compared to *M*=2.57 for residents. Mentors rated support from their assistant principals as *M*=3.66 compared to *M*=2.98 for residents.

Cohort 1 mentors felt the most supported by their LP-UTR coach compared to their peers in Cohorts 2 and 3. Cohort 1 residents had the lowest rating of support for their LP-UTR coach, with an average rating of *M*=2.70. This increased to *M*=3.69 for Cohort 2 residents, which was higher than Cohort 2 mentors' ratings for their LP-UTR coach. By Cohort 3, mentors *(M*=3.50) and residents (*M*=3.51) had very similar perceptions of support from their LP-UTR coaches. These increases likely reflect the changes from Cohort 1 through Cohorts 2 and 3 in the coaching model and numbers of school coaches, support for coaches—including individual check-ins—and other efforts to coordinate and structure their activities (see Part 2, p. 29).

TABLE 20. MENTORS' AND RESIDENTS' RATINGS OF SUPPORT FROM COLLEAGUES

	YEAR 1, SPRING 2016				YEAR 2, SPRING 2017				YEAR 3, SPRING 2018			
	Mentor		Resident		Mentor		Resident		Mentor		Resident	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Your principal	3.58	0.78	2.20	0.98	3.76	0.63	2.81	0.91	3.62	0.49	2.71	1.08
Your assistant principal(s)	3.58	0.85	2.60	1.02	3.79	0.57	3.08	0.92	3.62	0.56	3.26	0.77
Other teachers	3.71	0.46	2.90	0.70	3.53	0.64	3.44	0.76	3.50	0.57	3.20	0.82
Your LP-UTR coach	3.82	0.38	2.70	1.00	3.50	0.64	3.69	0.46	3.50	0.75	3.51	0.69
Your mentor	N/A	N/A	2.80	1.17	N/A	N/A	3.39	1.03	N/A	N/A	3.57	0.80

Scale: 1-not at all, 2=to a small extent, 3=to some extent, 4=to a great extent
Source: REA annual teacher survey

Opportunities for Collaboration

Throughout the resident focus groups, residents echoed the word "support." For the most part, residents felt supported by their mentors, coaches, and schools. One resident stated, "some of the individuals at the school had already done this program and were able to give support in addition to my mentor and coach." Another resident said, "I feel like I've been really supported through the program, in my school too."

FIGURE 16. RESIDENTS: FREQUENCY OF OPPORTUNITIES FOR COLLABORATION

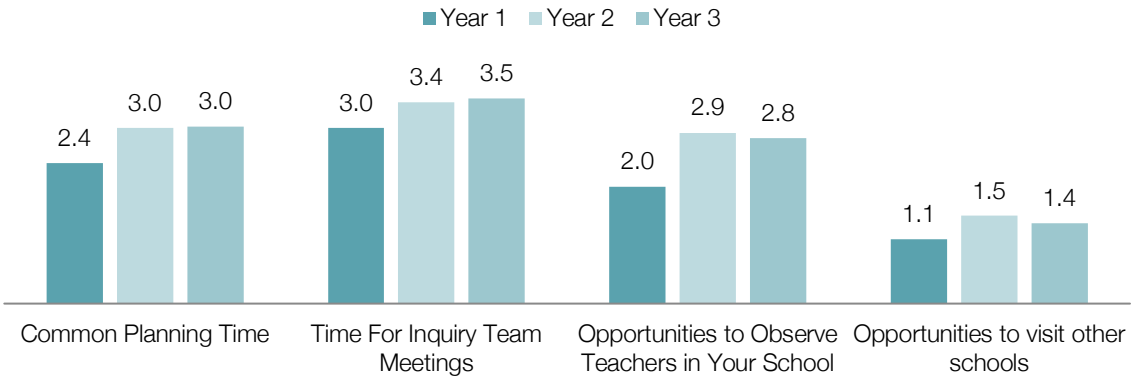
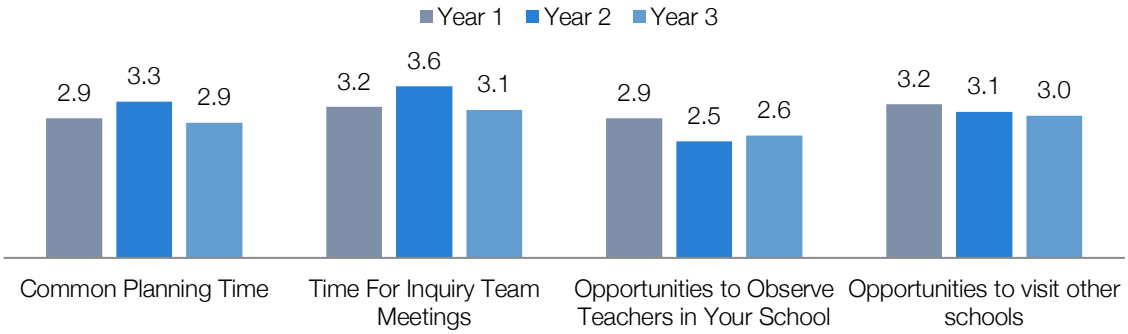


FIGURE 17. MENTORS: FREQUENCY OF OPPORTUNITIES FOR COLLABORATION



Survey feedback from mentors and residents about how often various opportunities for support were made available to them provided evidence of the types of program improvements made after the initial year of LP-UTR implementation. Residents' ratings indicated a perceived increase from the first year of the program to the second year in opportunities for common planning time, time for inquiry team meetings, opportunities to observe teachers in their school, and opportunities to visit other schools (see Figure 16). Mentors ratings showed an increase in perceived opportunities for common planning time and time for inquiry team meetings in subsequent years of the program. Interestingly, mentors indicated a decrease between Years 1 and 2 in opportunities to observe teachers in their school and opportunities to visit other schools (see Figure 17).

COLLABORATIVE ACTIVITIES ACROSS SCHOOLS

One of the challenges, and eventual changes that grew out the combination of the LP and UTR models was how residents could most benefit from the cross-school activities. During the first years of the project, there was a clarification of roles and expectations of mentors and residents, and it was decided that specific aspects of the program, such as the cross-school visits, would focus on aspects of mentor development. Mentors would in turn, use key takeaways from the cross-school visits to inform their work with their residents.

Levels of Participation in LP-UTR Activities

During the first years of the project, there was a clarification of roles and expectations of mentors and residents, and it was decided that specific aspects of the program, such as the cross-school visits, would focus on aspects of mentor development. Mentors would in turn, use key takeaways from the cross-school visits to inform their work with their residents. Other aspects of the program, including coursework at Hunter College and their in-school teaching experience, were tailored specifically to helping residents develop skills. As the number of school-based coaches increased from three in Cohort 1 to nine by Cohort 3, school-based coaches were another subset of teachers who experienced opportunities for growth and development through their participation in the program. In addition, other teachers at participating schools, including early career teachers, were primarily exposed to the program indirectly through their interactions with participating peers or through the effect on the overall school culture. Furthermore, the involvement of other non-mentor or resident teachers in program activities varied by school and triad with some triads during the first year of the program including non-mentor or resident teachers in the cross-school visits, others including them when they hosted visits, and others not integrating those teachers in any official capacity.

Participation in LP-UTR in each Cohort varied based on school culture, program experience, and the roles of administrators, coaches, and mentors. End-of-year survey responses from residents and mentors provided a composite, three-year picture of LP-UTR participation. (See Table 21.)

- Resident participation in triad visits hosted by their own school increased in each year of the program. Forty percent of Cohort 1 residents indicated they had taken part in visits at their school. This increased to 60% of residents in Cohort 2 and 66% in Cohort 3.
- Resident participation in visits to other schools increased in the second year of the program but involved less than half of residents. Only one resident in Cohort 1 indicated they had taken part in a visit hosted by another triad school. Thirty-nine percent of Cohort 2 and 37% of Cohort 3 residents had participated in visits to other schools.
- Perceived participation of administrators decreased in each year of the program for both mentors and residents (Mentors-Cohort 1 M=3.41, Cohort 2 M=3.30, Cohort 3 M=2.96; Residents- Cohort 1 M=3.40, Cohort 2 M=3.00, Cohort 3 M=2.21).
- In Cohorts 1 and 2 mentors perceived residents as participating more than residents viewed themselves as participating (Cohort 1 Mentor M=2.00, Resident M=1.38; Cohort 2 Mentor M=2.32, Resident

M=1.90). In Cohort 3, residents perceived themselves as participating to a greater extent than their mentors perceived them as being involved (Resident *M*=2.71, Mentor *M*=2.15).

- Mentors reported high levels of participation across program years with 76% of mentors in Cohort 1, 97% in Cohort 2 and 92% in Cohort 3 indicating that they always participate in project activities.

TABLE 21. REPORTS OF PARTICIPATION IN LP-UTR

			Never or rarely	Only occasionally	Fairly often	Always	Not sure	Mean	SD
COHORT 1	Administrators	Mentor		6%	47%	47%		3.41	0.62
		Resident			30%	20%	50%	3.40	0.55
	Mentors	Mentor		6%	12%	76%	6%	3.75	0.58
		Resident			30%	40%	30%	3.57	0.53
	Residents	Mentor	29%	53%	6%	12%		2.00	0.94
		Resident	70%			10%	20%	1.38	1
	Other teachers in your school	Mentor	29%	47%	12%	12%		2.06	0.41
		Resident	50%				50%	1.00	0
COHORT 2	Administrators	Mentor		19%	32%	49%		3.30	0.78
		Resident	7%	17%	30%	30%	17%	3.00	0.96
	Mentors	Mentor				97%		4.00	0
		Resident		20%	10%	60%	10%	3.44	0.85
	Residents	Mentor	19%	43%	24%	14%		2.32	0.94
		Resident	38%	38%	9%	9%	6%	1.90	0.96
	Other teachers in your school	Mentor	38%	46%	8%	5%	3%	1.81	0.82
		Resident	45%	28%	7%	3%	17%	1.63	0.82
COHORT 3	Administrators	Mentor	8%	27%	27%	38%		2.96	0.98
		Resident	9%	9%	26%	21%	35%	2.21	0.53
	Mentors	Mentor	0%	4%	4%	92%		3.88	0.42
		Resident	6%	3%	6%	53%	32%	2.15	0.49
	Residents	Mentor	35%	31%	19%	15%		2.15	1.06
		Resident	38%	18%	15%	12%	18%	2.71	0.52
	Other teachers in your school	Mentor	46%	23%	19%	8%		2.00	1.14
		Resident	24%	21%	12%	6%	38%	2.32	0.53

Source: REA Annual Surveys

As the number of school-based coaches increased from three in Cohort 1 to nine by Cohort 3, they were another subset of teachers who experienced opportunities for growth and development through their LP-UTR participation. In addition, other teachers at participating schools, including early career teachers, were exposed to the program indirectly through their interactions with participating peers or through the effect on the overall school culture. The involvement of other non-mentor or resident teachers in program activities varied by school and triad with some triads during the first year of the program including non-mentor or resident teachers in cross-school visits, others when they hosted visits, and others not integrating those teachers in any official capacity.

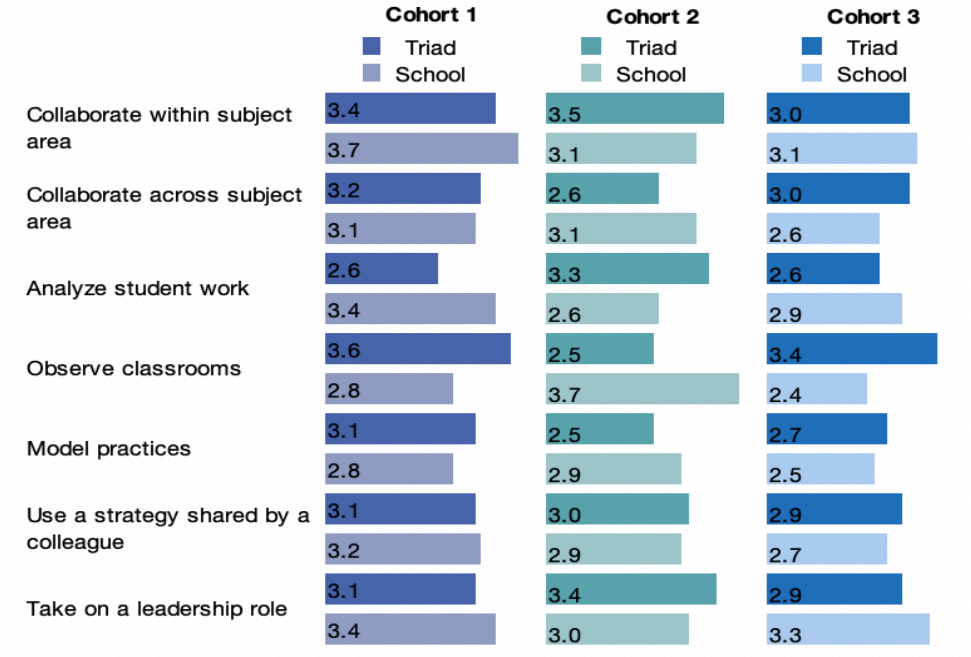
Building Skills and Confidence within and across Schools

Mentors

To get a sense of the extent that mentors and residents participated in activities that might have contributed to building their confidence and skills, we asked teachers how frequently they engaged in collaborative activities within their schools and across schools, or in their triad community.

- Cohort 1 mentors reported modeling practices, observing classrooms, and collaborating across subject areas more frequently across schools—or as part of triad work—than within schools. Activities they were involved in more frequently within schools included taking on a leadership role, using a strategy shared by a colleague, analyzing student work, and collaborating with teachers in their subject area.
- Cohort 2 mentors reported taking on a leadership role, observing classrooms, analyzing student work, using strategies, and collaborating within subject areas more often as part of triad work than within their schools.
- Cohort 3 mentors reported modeling practices, observing classrooms, and collaborating across subject areas, and using a strategy shared by a colleague more frequently across schools.

FIGURE 18. MENTORS: SCHOOL ENVIRONMENT AND COLLABORATION

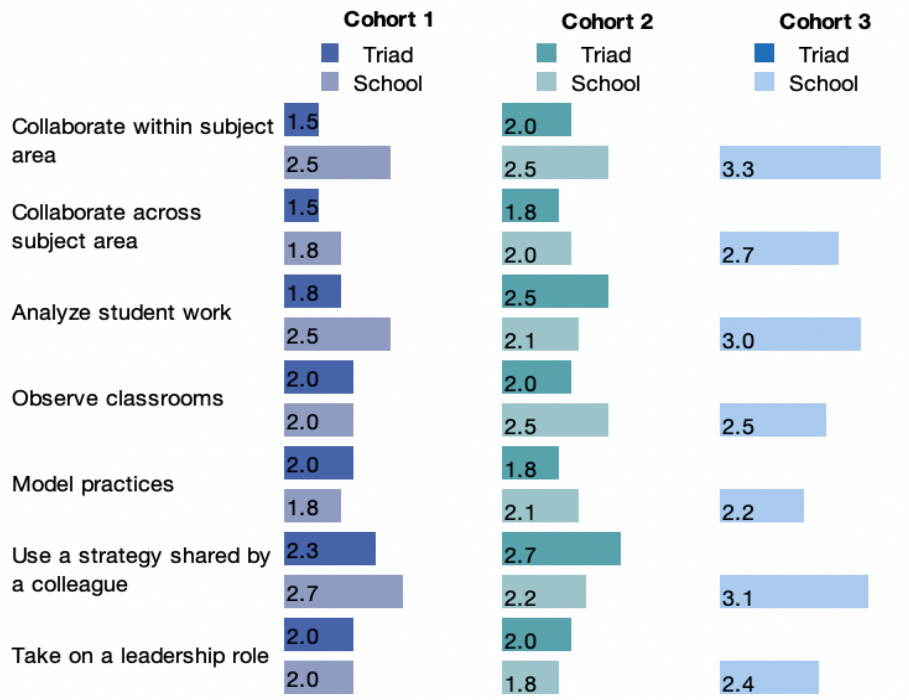


4-point Likert scale: 1=never or rarely happens, 2=happens only occasionally, 3=happens fairly often, 4=is a regular part of what we do
Source: REA annual surveys

Residents

- Residents reported being more engaged in school activities than in triad activities. There was a slight increase in the frequency of school-based activities between Cohorts 1 and 2, and a more a substantial increase for Cohort 3 residents.
- Cohort 1 and 2 residents indicated that they were more engaged in activities modeling practices in their triads than in their school. Cohort 2 residents also reported a higher level of frequency of collaborating across subject areas and observing classrooms as part of their triad work than in their schools.

FIGURE 19. RESIDENTS SCHOOL ENVIRONMENT AND COLLABORATION



4-point Likert scale: 1=never or rarely happens, 2=happens only occasionally, 3=happens fairly often, 4=is a regular part of what we do
Source: REA annual surveys
Note: Triad data wasn't collected for Cohort 3.

Cohort 3 residents had mixed reactions to school experiences and integration into the school community. During Cohort 3 end-of-year focus groups, residents noted how their experiences differed based on their schools' and mentors' level of involvement. Not all Cohort 3 residents were included in school activities, such as department or inquiry meetings; others residents were included—beyond the school day. For residents, the school experience also differed in the timeframe in which mentors would "hand over" a classroom. Some residents said they were expected to be teaching from Day 1; others voiced frustration over never feeling they owned a classroom.

Residents did note the program was different depending on the school and which content area they were placed. For some residents, the expectation was to be teaching from Day 1, while others were still trying to take on full responsibility for their focus class at the end of the school year. One resident said, "my class was never released to me. So, I've left not feeling like I have an idea how I'll be in terms of how I'm going to teach a class by myself."

The resident continued, "I think it's fine if it takes some residents a month, some three, some six months to get to that point, but you should get to that point." Another expectation residents felt schools were not always clear on was the responsibilities of Special Education and TESOL teachers compared to other content-area residents. One resident said, "my feeling is that my school needed extra help with ESL people and I was there to help them."

Residents liked that LP-UTR gave them classroom experience different from other teaching programs. Being immersed in the teaching profession was the perfect middle ground between traditional preparation programs and those like the Teaching Fellows. Residents also appreciated not going through the program alone—having other residents in the school and mentors and coaches—helped "build an extended support system."

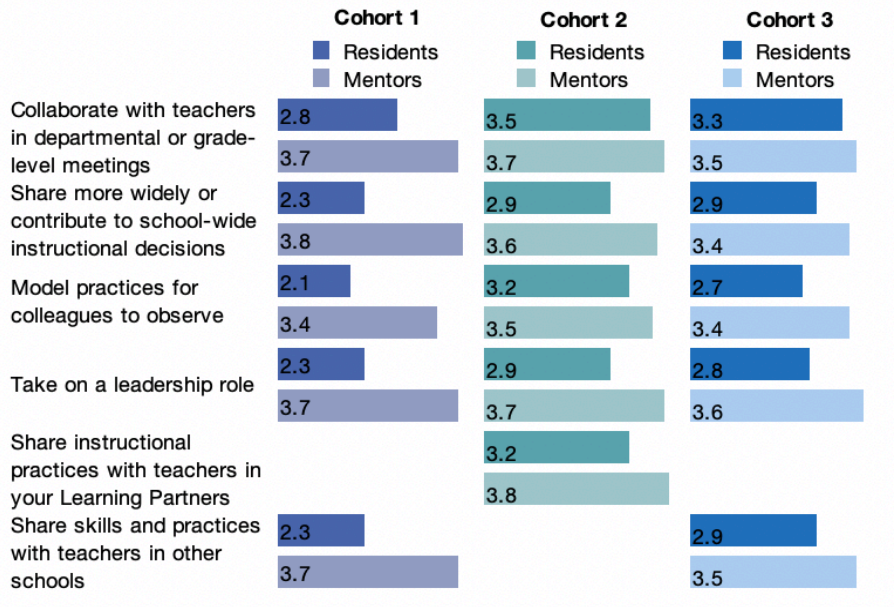
TEACHERS' COLLECTIVE EFFICACY

Levels of Confidence

Mentors' confidence levels related to efficacy items on the teacher survey were fairly high and consistent across categories and cohorts. Cohort 1 residents' confidence levels were noticeably lower than their peers in Cohorts 2 and 3 across all five efficacy categories.

- Residents' confidence in their ability to model practices for colleagues to observe increased from *M*=2.1 in Cohort 1 to *M*=3.2 in Cohort 2 and *M*=2.7 in Cohort 3.
- Cohort 1 residents also lagged significantly behind their peers in Cohorts 2 and 3 in their confidence regarding their ability to share skills or instructional practices with an *M*=2.3 compared to 3.2 and 2.9 for Cohorts 2 and 3. Note: the wording on this survey item changed between Cohort 1 and Cohort 2.
- Cohort 2 and Cohort 3 residents' confidence in their ability to collaborate with teachers in departmental or grade-level meetings was within two-tenths of a point of the confidence of their mentors.

FIGURE 20. CONFIDENCE RELATED TO EFFICACY



Source: REA Annual Surveys

New vs. Returning Mentors

To further explore changes in mentors' efficacy, we examined Cohort 2 and Cohort 3 fall and spring (baseline and end of year) survey responses, and disaggregated data by whether mentors were new or returning LP-UTR mentors. Overall mentor confidence levels for both Cohorts were high in the fall and increased in the spring. In the fall overall means were highest for collaborating with teacher in department or grade-level teams (Cohort 2 *M*=3.63, Cohort 3 *M*=3.62). In spring Cohort 2 ratings have high means in collaborating with teachers in department or grade-level teams (*M*=3.71), taking on a leadership role (*M*=3.73), and sharing instructional practices (*M*=3.78). For Cohort 3 the highest overall mean was taking on a leadership role (*M*=3.62). Overall means across all categories both in the fall and spring were lower for Cohort 3 than for Cohort 2. While in the fall some category means were higher for Cohort 2 returning mentors and others were higher for new mentors, by spring, new mentors had higher confidence ratings across all efficacy categories compared to those of returning mentors. For Cohort 3, returning mentors had higher fall and spring means than new mentors, although for some categories including collaborating with teachers in department or grade-level teams and sharing more widely or contributing to school-wide instructional decisions returning mentors confidence levels declined from fall to spring. (See Table 22.)

TABLE 22. LP-UTR MENTORS' CONFIDNECE IN COLLABORATIVE PRACTICES

		BASELINE			END OF YEAR		
		Overall Mean	Returning	New	Overall Mean	Returning	New
Collaborate with teachers in department or grade-level teams	Cohort 2	3.63	3.57	3.67	3.71	3.5	3.81
	Cohort 3	3.62	3.76	3.50	3.50	3.64	3.40
Share more widely or contribute to school-wide instructional decisions	Cohort 2	3.36	3.64	3.21	3.61	3.33	3.73
	Cohort 3	3.33	3.52	3.17	3.35	3.36	3.33
Model practices for colleagues to observe	Cohort 2	3.31	3.29	3.32	3.49	3.36	3.54
	Cohort 3	3.31	3.43	3.21	3.42	3.45	3.40
Take on a leadership role	Cohort 2	3.50	3.71	3.39	3.73	3.64	3.77
	Cohort 3	3.44	3.48	3.42	3.62	3.73	3.53
Share instructional practices with teachers in your Learning Partner triad*	Cohort 2	3.43	3.57	3.36	3.78	3.55	3.88
Share skills and practices with teachers in other schools*	Cohort 3	3.40	3.57	3.25	3.46	3.55	3.40

4-point Likert scale: (1=not at all confident, 2=somewhat confident, 3=generally confident, and 4=very confident)
Source: REA annual surveys
*Note: Question was worded differently in the Cohort 2 and Cohort 3 versions of the survey.

The end of year survey asked mentors and residents where they acquired confidence in collaboration and leadership skills. Tables 23 and 24 below show that both mentors and residents gained confidence in collaboration and leadership skills through teaching and interacting with their school colleagues. About a quarter

of mentors attributed growth in their ability to share skills and practices with teachers in other schools to LP-UTR cross-school sharing.

TABLE 23. WHERE MENTORS DEVELOPED EFFICACY SKILLS

		New Visions PD, support	LP-UTR coach	Teaching, School colleagues	LP-UTR cross-school sharing
Collaborate with teachers in departmental or grade-level teams	Cohort 1	15%	5%	70%	10%
	Cohort 2	15%	13%	56%	16%
	Cohort 3	14%	5%	62%	19%
Contribute to school-wide instructional decisions	Cohort 1	18%	9%	64%	9%
	Cohort 2	11%	11%	64%	13%
	Cohort 3	14%	9%	60%	16%
Model practices for colleagues to observe	Cohort 1	9%	13%	61%	17%
	Cohort 2	16%	13%	52%	20%
	Cohort 3	20%	14%	49%	16%
Take on a leadership role	Cohort 1	17%	13%	58%	13%
	Cohort 2	24%	14%	46%	16%
	Cohort 3	17%	13%	55%	15%
Share skills and practices with teachers in other schools	Cohort 1	17%	14%	45%	24%
	Cohort 2	21%	14%	41%	24%
	Cohort 3	16%	16%	40%	28%

Source: REA Annual Surveys

Like mentors, residents attributed the largest area of skill development to teaching and interacting with school colleagues, but residents also credited Hunter coursework, LP-UTR coaching, and New Visions PD and support with helping them to develop specific subsets of skills. Over 22% of residents in Cohorts 1, 2, and 3 indicated that Hunter coursework helped them to model practices for colleagues to observe, and over 19% attributed their ability to share skills and practices with teachers in other schools to their Hunter Coursework. Cohort 2 residents found their LP-UTR coach to be slightly more helpful in their skill development across categories, compared to their Cohort 1 and 3 peers.

TABLE 24. WHERE RESIDENTS DEVELOPED EFFICACY SKILLS

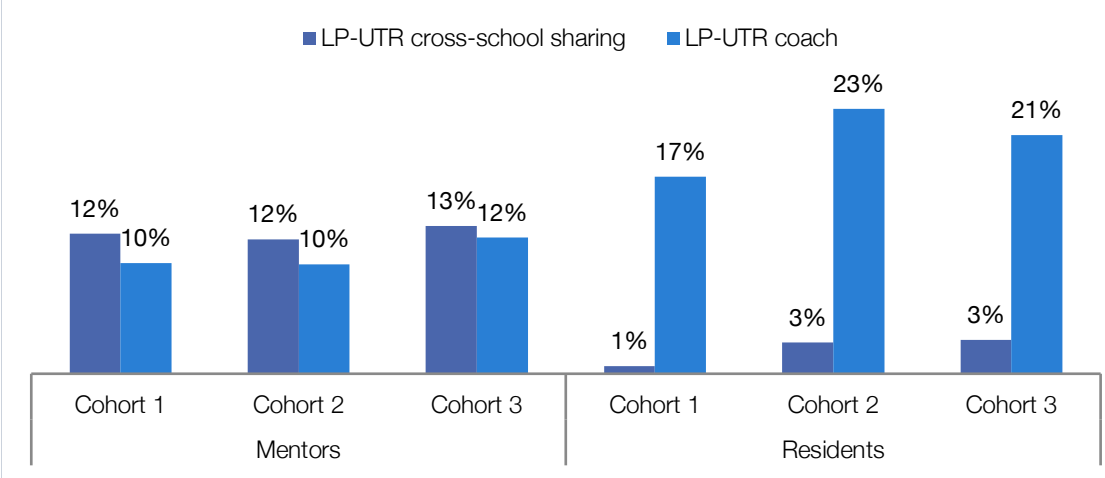
		Hunter Coursework	New Visions PD, support	LP-UTR coach	Teaching, School colleagues	LP-UTR cross-school sharing
Collaborate with teachers in departmental or grade-level teams	Cohort 1	8%	8%	25%	58%	0%
	Cohort 2	13%	13%	29%	41%	5%
	Cohort 3	17%	11%	20%	52%	2%
Contribute to school-wide instructional decisions	Cohort 1	14%	14%	14%	57%	0%
	Cohort 2	12%	16%	22%	42%	8%
	Cohort 3	18%	16%	23%	42%	2%
Model practices for colleagues to observe	Cohort 1	25%	13%	13%	57%	0%
	Cohort 2	22%	14%	22%	38%	3%
	Cohort 3	26%	15%	20%	35%	5%
Take on a leadership role	Cohort 1	14%	14%	14%	57%	0%
	Cohort 2	18%	18%	25%	38%	2%
	Cohort 3	18%	22%	19%	37%	3%
Share skills and practices with teachers in other schools	Cohort 1	22%	22%	11%	44%	0%
	Cohort 2	21%	15%	23%	37%	5%
	Cohort 3	19%	18%	18%	36%	9%

Source: REA Annual Surveys

Figure 21 below shows what percentages of confidence gained across all skill sets- practice, assessment, and collaboration- mentors and residents attribute to LP-UTR activities. Across all three cohorts, mentors consistently attributed slightly higher confidence gains to LP-UTR cross-school sharing than to their LP-UTR coach. Residents attributed between one and three percent of their confidence gains to LP-UTR cross-school sharing and between 17% and 23% to their LP-UTR coach.

As the number of school-based coaches increased from three in Cohort 1 to nine by Cohort 3, school-based coaches were another subset of teachers who experienced opportunities for growth and development through their participation in the program. In addition, other teachers at participating schools, including early career teachers, were primarily exposed to the program indirectly through their interactions with participating peers or through the effect on the overall school culture. Furthermore, the involvement of other non-mentor or resident teachers in program activities varied by school and triad with some triads during the first year of the program including non-mentor or resident teachers in the cross-school visits, others including them when they hosted visits, and others not integrating those teachers in any official capacity.

FIGURE 21. PERCENTAGE OF GAINS IN CONFIDNECE THAT TEACHERS ATTRIBUTE TO LP-UTR



PART 5: LP-UTR'S IMPACT ON STUDENT ACHIEVEMENT

This portion of the report presents results pertaining to the fourth and final major research question, *How does participating in LP-UTR impact student achievement?* To gauge the effectiveness of LP-UTR on students, we conducted a set of analyses using outcome data gathered at the end of the third and final year of the project for all LP-UTR schools and matched non-LP-UTR schools. The data included New York Regents examination performances and course credit accumulation during the 2017–18 school year. Achievement data for all courses taught by graduate and mentor teachers was included.

- In ELA, math, science and social studies, student achievement outcomes in LP-UTR graduate teachers' classrooms tended to be comparable to, or better than, those taught by other graduate teachers at non-LP-UTR schools. This was true of both the standardized Regents exam performances and the percentages of course credits earned.
- Students taught by the LP-UTR mentor teachers generally attained similar or greater levels of achievement than their peers taught by comparison mentor teachers at non-LP-UTR schools.
- Additional analyses of achievement among all students with disabilities at either LP-UTR or at matched comparison schools revealed that the students on average fared consistently better in core courses at LP-UTR schools.
- Analyses of achievement among all English learners (ELs) showed that ELs at LP-UTR schools consistently earned a higher percentage of course credits across core subjects than ELs at non-LP-UTR schools, and they generally performed at least at similar levels as their peers on Regents exams.
- Students of multiyear LP-UTR mentors demonstrated greater exam performance on 4 of the 5 contrasts compared to the non-LP-UTR group, and similar performance on the remaining contrast. The more experienced LP-UTR mentor teachers thus exhibited a pattern of achievement results that contrasted with the novice mentor teachers and more closely resembled the patterns observed among the LP-UTR graduate teachers.

Because of the critical role of the mentor teachers in LP-UTR and the support for professional development that they received, the achievement of students taught by those mentor teachers' (and not just by LP-UTR graduates) was another key long-term indicator of program success. The Regents examinations, which are statewide standardized examinations in core high school subjects, spanned the domains of ELA, math (Algebra, Algebra II/Trigonometry, Geometry), science (Earth Science, Living Environment), social studies/history (Global History, U.S. History). The LP-UTR evaluation used a quasi-experimental design to help address the need for causal research about the impacts of teacher residency programs on student outcomes.

- The primary analyses of student achievement sought to test whether students of LP-UTR teachers (both graduate and mentor teachers) outperformed the students of non-LP-UTR teachers, as measured by course examination scores and

course credits earned. One subset of these analyses examined the student outcomes associated with the graduates of LP-UTR, and the other set of analyses examined the student outcomes associated with the LP-UTR mentors. In addition, within each graduate and mentor teacher group, the analyses were conducted separately by amount of teaching experience.

- The secondary analyses of achievement examined the outcomes of all students with disabilities and of all English learners at the LP-UTR and matched comparison schools. This research paralleled analysis conducted earlier in the evaluation study (i.e., Years 1 and 2), when the NYC DOE data records did not link student-level data to particular special education teachers or TESOL teachers. The inability to link teachers to students required us to (separately) analyze the achievement outcomes of all special education students (to investigate schoolwide effects of special education teachers) and of all English learners (to investigate schoolwide effects of TESOL teachers) at LP-UTR and non-LP-UTR schools.

During the first two years of the project evaluation, the analyses of student achievement outcomes were limited to examining student achievement associated with the mentor teachers and not also the graduate teachers of LP-UTR. It was also not possible to examine student outcomes assessed at the end of the residency year, because the residents were not identified by NYC DOE as the teacher of record. For the examination of student achievement associated with mentor teachers, student-level data was linked to mentor teachers of ELA, math, and science, but student-level data could not be linked to the special education or TESOL mentor teachers because they, too, were not identified as the teacher of record in NYC DOE data systems or they taught within integrated teams of teachers. Therefore, the investigation of student outcomes for special education and TESOL mentor teachers was limited to including all students receiving special education services or to all English learners (ELs) at each mentors' school. Note that members of both subgroups of students were included in the primary analyses of ELA, math, and science outcomes.

These earlier analyses of Year 1 and Year 2 achievement data for LP-UTR mentor teachers yielded mixed results. First, the investigation of the NYS Regents examination scores tended to show comparable levels of academic performance between LP-UTR mentor teachers and non-LP-UTR mentor teachers. With the analyses of course credit earning, however, the findings showed a more favorable, positive trend in the percentage of credits earned out of those attempted within the LP-UTR mentor group than the non-LP-UTR mentor group. On the whole, the student outcomes associated with LP-UTR mentors were at similar levels if not higher levels of performance than the student outcomes at matched schools.

The impact findings about student achievement assessed at the end of the 2017–18 school year involved a series of analyses of Regents examination scores and course credit earning. The results are organized as follows:

1. Primary Achievement Analyses:
   Student data linked to LP-UTR teachers or matched comparison teachers
   a. Graduates
      i. Cohort 1 graduates: Analyses of subject-specific, student-level outcomes assessed at the end of the second year of teaching post-graduation.

ii. Cohort 2 graduates: Analyses of subject-specific, student-level outcomes assessed at the end of the first year of teaching post-graduation.
iii. Cohort 1 and 2 graduates combined: Analyses of the combined, subject-specific one-year outcomes of Cohort 2 and two-year outcomes of Cohort 1.

b. Mentors
   i. Teachers who served as a mentor for 1 year.
   ii. Teachers who served as a mentor for multiple years.

2. Secondary Schoolwide Achievement Analyses:
   Student data of all English learners and of all students with disabilities at the LP-UTR and matched comparison schools—not linked to particular teachers
   a. Students with disabilities: Analyses of subject-specific, student-level outcomes assessed at the end of the 2017–18 school year in all LP-UTR and matched non-LP-UTR schools (Years 1–3).
   b. English learners: Analyses of subject-specific, student-level outcomes assessed at the end of the 2017-18 school year in all LP-UTR and matched non-LP-UTR schools (Years 1–3).

PRIMARY STUDENT ACHIEVEMENT ANALYSES

Regents Exam Performances of Students Taught by LP-UTR Graduates

Analysis of covariance (ANCOVA) was used to compare Regents examination performances of students taught by LP-UTR graduate teachers with performances by students of graduate teachers at matched non-LP-UTR schools, controlling for gender, economic disadvantage, special education status, EL status, and ethnicity (% Black, % Hispanic). Again, prior achievement data was unavailable for our analyses. Throughout these individual results of achievement contrasts, it is important to remain aware of the small numbers of teachers that taught individual subjects and therefore interpret individual findings with caution. Trends across sets of analyses can provide a more robust understanding of achievement outcomes associated with LP-UTR participation.

Do students taught be LP-UTR graduates during their **second year of teaching** *after graduation outperform the students taught by non-LP-UTR graduates during their second year of teaching post-graduation as measured by* **course examination scores***?*

- Students of Cohort 1 LP-UTR residents during the second year of teaching post-graduation had statistically higher summative examination scores than students of matched teachers for 2 of the 5 achievement exam contrasts, and there were no statistically significant differences for the other 3 contrasts. (See Table 25.) Note that in 4 of the contrasts each group is represented by a single teacher. **This set of results suggests that students taught by LP-UTR teachers during the second year after graduation performed either the same or better on Regents exams (ELA, math, science, and social studies/history) than their peers in similar matched schools.**

TABLE 25. REGENTS EXAM PERFORMANCES OF STUDENTS TAUGHT BY GRADUATES, COHORT 1

Outcome Area	LP-UTR			Non-LP-UTR			Sig
	N Student (Teacher)	Adjusted Mean	SE	N Student (Teacher)	Adjusted Mean	SE	
Common Core ELA Regents Score	139(4)	71.0	1.31	148(4)	70.3	1.28	p=.71
Common Core ALGEBRA Regents Score	43(1)	**65.5**	1.72	18(1)	**52.8**	2.71	p<.01
LIVING ENVIRONMENT Regents Score	83(1)	71.1	1.49	79(1)	68.0	1.53	p=.18
GLOBAL HISTORY Regents Score	114(1)	73.4	1.23	169(1)	70.3	1.01	p=.06
U.S. HISTORY Regents Score	28(1)	**89.9**	2.00	11(1)	**78.1**	3.47	p=.01

Spring 2018 NY Regents Examinations

Do students taught be LP-UTR graduates during their **first year of teaching** *post-graduation outperform the students taught by non-LP-UTR graduates during their first year of teaching post-graduation as measured by* **course examination scores***?*

- Students of Cohort 2 residents during the first year of teaching after LP-UTR graduation had statistically higher scores than students of matched teachers for 1 of the 7 achievement exam contrasts (1 ELA, 3 math, 2 science, 1 history), statistically lower scores for 1 contrast, and no statistically significant differences for the other 5. (See Table 26.) **Overall, students of LP-UTR graduates, this time during the first year of teaching after graduation, tended to perform the same as non-LP-UTR graduates.**

TABLE 26. REGENTS EXAM PERFORMANCES OF STUDENTS TAUGHT BY GRADUATES, COHORT 2

Outcome Area	LP-UTR			Non-LP-UTR			Sig
	N Student (Teacher)	Adjusted Mean	SE	N Student (Teacher)	Adjusted Mean	SE	
Common Core ELA Regents Score	340(10)	**63.5**	1.04	299(11)	**60.5**	1.11	p=.05
Common Core ALGEBRA Regents Score	129(2)	**65.2**	0.87	69(2)	**70.3**	1.20	p<.01
Common Core GEOMETRY Regents Score	36(2)	52.4	2.42	62(2)	50.9	1.79	p=.63
ALGEBRA II / TRIGONOMETRY Regents Score	14(1)	57.1	3.35	27(1)	61.0	2.40	p=.37
EARTH SCIENCE Regents Score	271(6)	63.7	.99	156(6)	64.2	1.30	p=.78
LIVING ENVIRONMENT Regents Score	73(3)	68.8	1.34	73(3)	66.5	1.34	p=.25
GLOBAL HISTORY Regents Score	27(1)	57.3	2.87	30(1)	56.7	1.27	p=.88

Spring 2018 NY Regents Examinations

Do students taught be LP-UTR graduates in the first or second year of teaching post-graduation outperform the students taught by non-LP-UTR graduates as measured by course examination scores and course credits earned?

- Using the same set of scores as the above contrasts for matching exam subjects, combined cohort analyses showed LP-UTR students statistically outperformed non-LP-UTR students on 2 of 4 contrasts and had similar performances for the other 2 contrasts. **Thus, the students of LP-UTR graduates had similar or better levels of exam achievement than their peers.**

In summary, based on the mixed results across 12 comparisons of achievement outcomes across 4 domains and 7 subject areas, the LP-UTR group did not demonstrate higher achievement than the non-LP-UTR group. Eight of the 12 contrasts showed no LP-UTR effect, 3 showed a positive effect, and 1 showed a negative effect. **The overall pattern of results, thus, showed students of LP-UTR graduate teachers on average tending to perform the same, and in some cases better than their peers.**

TABLE 27. OVERALL REGENTS EXAM PERFORMANCES, COHORTS 1 AND 2 COMBINED

Outcome Area	LP-UTR			Non-LP-UTR			Sig
	N Student (Teacher)	Adjusted Mean	SE	N Student (Teacher)	Adjusted Mean	SE	
Common Core ELA Regents Score	479(14)	**66.1**	0.85	447(15)	**63.3**	0.88	p=.02
Common Core ALGEBRA Regents Score	172(3)	65.3	0.86	87(3)	66.5	1.21	p=.43
LIVING ENVIRONMENT Regents Score	156(4)	**70.2**	.99	152(4)	**67.1**	1.00	p=.04
GLOBAL HISTORY Regents Score	141(2)	70.6	1.17	199(2)	68.0	.98	p=.09

Spring 2018 NY Regents Examinations

Course Credits Earned among Students Taught by LP-UTR Graduates

The approach to investigating outcomes in credit earning again relied on ANCOVA, controlling for gender, economic disadvantage, special education status, EL status, and ethnicity (% Black, % Hispanic).

Do students taught be LP-UTR graduates during their **second year of teaching** *post-graduation outperform the students taught by non-LP-UTR graduates during their second year of teaching post-graduation as measured by* **course credits earned***?*

- Students of Cohort 1 residents during the first year of teaching after LP-UTR graduation had statistically higher percentage of credits earned than students of matched teachers for 2 of the 4 contrasts, and no statistically significant differences for the other 2 contrasts. (See Table 28.) **The results overall suggested students of LP-UTR graduates in their second year of teaching tended to fare at least the same if not better than their peers.**

TABLE 28. ANALYSIS OF GRADUATES COURSE CREDITS, COHORT 1

Outcome Area	LP-UTR			Non-LP-UTR			Sig
	N Student (Teacher)	Adjusted Mean	SE	N Student (Teacher)	Adjusted Mean	SE	
ELA % Course Credits Earned	205(4)	.83	.02	230(4)	.78	.02	p=.16
MATH % Course Credits Earned	51(1)	**.84**	.04	28(1)	**.48**	.06	p<.01
SCIENCE % Course Credits Earned	90(1)	**.89**	.04	106(1)	**.68**	.04	p<.01
SOCIAL STUDIES % Course Credits Earned	166(2)	0.81	.03	241(2)	.76	.02	p=.15

2017–18 School Year

Do students taught be LP-UTR graduates during their **first year of teaching** *post-graduation outperform the students taught by non-LP-UTR graduates during their first year of teaching post-graduation as measured by* **course credits earned***?*

- Students of Cohort 2 residents during the first year of teaching after LP-UTR graduation had statistically higher percentage of credits earned than students of matched teachers for 1 of the 4 contrasts, and no statistically significant differences for the other 3 contrasts. (See Table 29.) **The trend in results indicated that students taught by LP-UTR graduates during the year after graduation tended to perform at least the same as peers of non-LP-UTR graduates.**

TABLE 29. ANALYSIS OF GRADUATES COURSE CREDITS, COHORT 2

Outcome Area	LP-UTR			Non-LP-UTR			Sig
	N Student (Teacher)	Adjusted Mean	SE	N Student (Teacher)	Adjusted Mean	SE	
ELA % Course Credits Earned	700(10)	**.85**	.01	731(14)	**.76**	.01	p<.01
MATH % Course Credits Earned	472(4)	.77	.02	339(5)	.79	.02	p=.57
SCIENCE % Course Credits Earned	459(8)	.79	.02	508(10)	.75	.02	p=.07
SOCIAL STUDIES % Course Credits Earned	35(1)	.75	.06	39(1)	.72	.06	p=.75

2017–18 School Year

Do students taught be LP-UTR graduates in the first or second year of teaching post-graduation outperform the students taught by non-LP-UTR graduates as measured by **course credits earned***?*

- The combined cohort analyses demonstrated that students taught by LP-UTR graduates had a statistically significant higher percentage of credits earned than students of matched teachers for 1 of

the 4 contrasts, and no statistically significant differences for the other 3 contrasts. (See Table 30.) **The alternative, cohort-neutral analyses of percentage of credits earned thus yielded mixed results, but showed the students of LP-UTR graduates tended to on average perform at similar levels as their peers.**

TABLE 30. ANALYSIS OF GRADUATES COURSE CREDITS, COHORTS 1 AND 2 COMBINED

Outcome Area	LP-UTR			Non-LP-UTR			Sig
	N Student (Teacher)	Adjusted Mean	SE	N Student (Teacher)	Adjusted Mean	SE	
ELA % Course Credits Earned	1005(14)	**.84**	.01	961(18)	**.76**	.01	p<.01
MATH % Course Credits Earned	523(5)	.78	.02	367(6)	.76	.02	p=.30
SCIENCE % Course Credits Earned	524(9)	.79	.02	369(11)	.79	.02	p=.97

2017–18 School Year

Based on 8 comparisons of the percentage of credits earned (spanning 4 domains), results were mixed, with the LP-UTR group demonstrating a higher percentage of credits earned for 3 of the contrasts, and no difference in credit earning for the other 5 contrasts. **The trend of results for percentage of course credits earned thus mirrored the trends for examination outcomes by showing LP-UTR students to on average earn a similar percentage of course credits if not more than their peers.**

Regents Exam Performances of Students Taught by LP-UTR Mentors

To align with the needs of the LP-UTR team, the evaluation analyzed first-year mentor teachers separately from the mentor teachers with multiple years of experience.

Do students taught by LP-UTR first-year mentors outperform the students taught by non-LP-UTR first-year mentors as measured by course examination scores and course credits earned?

- Students of first-year LP-UTR teacher mentors demonstrated higher exam performance on 3 of the 9 contrasts (4 domains, 9 subjects) than students of matched teacher mentors, lower performance on 4 of the contrasts, and similar performance on 2 contrasts. (See Table 31.) Thus, **compared to the patterns of findings reported above for LP-UTR graduate teachers, the trend of outcomes for the first-year mentor teachers was less positive, with just 5 of the 9 contrasts showing similar or better performances by the students of LP-UTR mentor teachers.**

TABLE 31. ANALYSIS OF REGENTS SCORES, MENTORS ONE YEAR ONLY

Outcome Area	LP-UTR			Non-LP-UTR			Sig
	N Student (Teacher)	Adjusted Mean	SE	N Student (Teacher)	Adjusted Mean	SE	
Common Core ELA Regents Score	292(9)	**66.5**	1.28	327(9)	**59.0**	1.22	p<.01
Common Core ALGEBRA Regents Score	104(4)	**74.4**	.98	147(4)	**64.8**	0.81	p<.01
Common Core GEOMETRY Regents Score	93(3)	**63.6**	1.38	134(3)	**59.4**	1.14	p=.02
ALGEBRA II / TRIGONOMETRY Regents Score	56(2)	62.8	1.75	111(2)	66.0	1.23	p=.15
EARTH SCIENCE Regents Score	395(6)	**71.7**	0.67	282(6)	**76.3**	0.80	p<.01
LIVING ENVIRONMENT Regents Score	30(1)	**63.8**	3.03	86(1)	**72.0**	1.54	p=.03
CHEMISTRY Regents Score	39(1)	**62.3**	1.63	59(1)	**67.4**	1.29	p=.02
GLOBAL HISTORY Regents Score	291(6)	**69.0**	0.84	259(6)	**78.0**	0.89	p<.01
U.S. HISTORY Regents Score	64(3)	64.8	1.43	120(3)	67.2	1.04	p=.17

Spring 2018 NY Regents Examinations

Do students taught by LP-UTR multi-year mentors outperform the students taught by non-LP-UTR multiyear mentors as measured by course examination scores and course credits earned?

- Students of multiyear LP-UTR teacher mentors demonstrated greater exam performance on 4 of the 5 contrasts compared to the non-LP-UTR group, and similar performance on the remaining contrast. (See Table 32.) **The more experienced LP-UTR mentors thus exhibited a pattern of achievement results that contrasted with the novice mentors and more closely resembled the patterns observed among the LP-UTR graduates.**

TABLE 32. ANALYSIS OF REGENTS SCORES, MENTORS MULTIPLE YEARS

Outcome Area	LP-UTR			Non-LP-UTR			Sig
	N Student (Teacher)	Adjusted Mean	SE	N Student (Teacher)	Adjusted Mean	SE	
Common Core ELA Regents Score	196(5)	**77.0**	1.17	178(5)	**63.9**	1.23	p<.01
Common Core ALGEBRA Regents Score	64(2)	69.8	1.37	36(2)	67.7	1.84	p=.35
EARTH SCIENCE Regents Score	76(1)	**89.4**	1.18	33(1)	**73.4**	1.86	p<.01
LIVING ENVIRONMENT Regents Score	43(1)	**78.2**	1.50	79(1)	**73.6**	1.10	p=.02
CHEMISTRY Regents Score	28(1)	**81.7**	1.55	79(1)	**67.7**	0.91	p<.01

Spring 2018 NY Regents Examinations

In summary, the individual contrasts of exam outcomes for students taught by mentor teachers were mixed, with students of LP-UTR mentors faring statistically better than the comparison group in 7 of the 14 total contrasts, worse in 4 of the contrasts, and the same in 3. **In 10 of the 14 total contrasts, LP-UTR students performed at higher or similar levels as their peers on Regents examinations, and the trend in exam outcomes was more positive among the multiyear LP-UTR mentors than it was among the first-year LP-UTR mentors.**

Course Credits Earned by Students Taught by LP-UTR Mentors

Paralleling the above set of analyses, we conducted separate analyses of novice and experienced mentor teachers to examine rates of course credit earning in their classrooms, comparing their outcomes to those of matched peers in non-LP-UTR schools.

Do students taught by LP-UTR first-year mentors outperform the students taught by non-LP-UTR first-year mentors as measured by course examination scores and course credits earned?

- Students of first-year LP-UTR teacher mentors earned significantly higher percentages of credits than the comparison group in 3 of the 4 contrasts, and a similar percentage for the remaining contrast. **The students taught by LP-UTR teacher mentors therefore tended to earn at a minimum comparable but largely higher percentage of course credits than their peers.**

TABLE 33. ANALYSIS OF COURSE CREDITS, MENTORS ONE YEAR ONLY

Outcome Area	LP-UTR			Non-LP-UTR			Sig
	N Student (Teacher)	Adjusted Mean	SE	N Student (Teacher)	Adjusted Mean	SE	
ELA % Course Credits Earned	446(10)	.85	.01	509(9)	.83	.01	p=.31
MATH % Course Credits Earned	298(6)	**.96**	.02	630(9)	**.82**	.01	p<.01
SCIENCE % Course Credits Earned	687(9)	**.88**	.01	722(9)	**.75**	.01	p<.01
SOCIAL STUDIES % Course Credits Earned	435(7)	**.77**	.02	550(9)	**.82**	.02	p=.03

2017–18 School Year

Do students taught by LP-UTR multiyear mentors outperform the students taught by non-LP-UTR multi-year mentors as measured by course examination scores and course credits earned?

- Students of multiyear LP-UTR teacher mentors earned significantly higher percentages of credits than the comparison group in 2 of the 3 contrasts, and a similar percentage for the remaining contrast. Thus, **the students in classrooms taught by multiyear mentors of LP-UTR earned similar or higher percentages of course credits than the comparison students.**

TABLE 34. ANALYSIS OF COURSE CREDITS, MENTORS MULTIPLE YEARS

Outcome Area	LP-UTR			Non-LP-UTR			Sig
	N Student (Teacher)	Adjusted Mean	SE	N Student (Teacher)	Adjusted Mean	SE	
ELA % Course Credits Earned	355(6)	**.93**	.02	480(6)	**.82**	.01	p<.01
MATH % Course Credits Earned	78(2)	.84	.04	54(2)	.84	.04	p=.89
SCIENCE % Course Credits Earned	164(2)	**.92**	.02	279(3)	**.83**	.02	p=.01

2017–18 School Year

To summarize, in this set of findings about credit earning, students of LP-UTR teacher mentors on average earned a significantly higher percentage of credits in 5 of the 7 contrasts, and a similar percentage in the remaining 2 contrasts. **The positive trends in credit earning outcomes were similar for both first-year and multiyear mentors, and they consistently revealed students of LP-UTR teacher mentors earned similar or higher percentages of course credits than their peers.**

SCHOOL-WIDE ACHIEVEMENT ANALYSES FOR STUDENTS WITH DISABILITIES AND ENGLISH LEARNERS

For the secondary series of achievement analyses, we examined 2017–18 achievement outcomes for all students with disabilities and all ELs at the LP-UTR and matched schools. For the special education analyses, we looked at scores for students with disabilities only, using analysis of covariance, controlling for gender, economic disadvantage, EL status, and ethnicity (Hispanic and African American). For the TESOL analyses, we looked at scores for English learners only, using analysis of covariance, controlling for gender, economic disadvantage, special education status, and ethnicity (Hispanic and African American). As noted previously, because teacher and student data were not linked, we were unable to disentangle outcomes for LP-UTR graduates and mentors.

Regents Examination Performances of Students with Disabilities

Do students with disabilities in LP-UTR schools exhibit higher achievement than the students in non-LP-UTR schools as measured by course examination scores and course credits earned?

- In examining the 2017–18 Regents exam outcomes for all 3 cohorts of LP-UTR schools and comparison schools, 4 of the 9 course-specific contrasts (ELA, Algebra, Earth Science, and U.S. History) showed students with disabilities significantly outperforming peers at comparison schools, performing similarly to their peers in 4 additional contrasts (Geometry, Algebra II/Trigonometry, Living Environment, and Global History), and significantly worse in one contrast (Chemistry; see Table 35).
- Results for all subjects combined are also provided in Table 35, which showed LP-UTR schools with a significantly higher average Regents score, but a similar overall pass rate.
- **The general trend, therefore, was that the students with disabilities on average tended to perform the same or better than their peers at matched non-LP-UTR schools.**

TABLE 35. ANALYSIS OF REGENTS SCORES, SPECIAL EDUCATION STUDENTS YEARS 1–3 SCHOOLS

Outcome Area	LP-UTR Years 1 – 3 Schools			Non-LP-UTR Years 1 – 3 Comparison Schools			Sig
	N Student	Adjusted Mean	SE	N Student	Adjusted Mean	SE	
Regents ALL SUBJECTS # Passed at 55 or Above	1988	1.58	.03	5938	1.56	.02	p=.42
Regents ALL SUBJECTS Average Score	1988	**59.2**	.32	5938	**58.1**	.19	p<.01
Common Core ELA Regents Score	700	**59.5**	.80	2382	**57.2**	.43	p=.01
Common Core ALGEBRA Regents Score	906	**59.0**	.44	3175	**56.9**	.23	p<.01
Common Core GEOMETRY Regents Score	362	51.3	.79	762	51.0	.54	p=.72
ALGEBRA II / TRIGONOMETRY Regents Score	88	57.5	1.54	294	58.1	.84	p=.77
EARTH SCIENCE Regents Score	393	**59.5**	.83	1009	**55.1**	.52	p<.01
LIVING ENVIRONMENT Regents Score	853	59.5	.47	3455	59.8	.28	p=.67
CHEMISTRY Regents Score	82	**51.5**	1.47	184	**58.8**	.98	p<.01
GLOBAL HISTORY Regents Score	873	59.5	.57	2603	58.7	.33	p=.22
U.S. HISTORY Regents Score	514	**65.3**	.74	1832	**62.2**	.39	p<.01

Spring 2018 NY Regents Examinations

Course Credit Earned by Students with Disabilities

When looking at schoolwide credit accumulation for students with disabilities across all 3 project years, we found a consistently favorable pattern of results. **All 4 subject-specific contrasts (ELA, math, science, social studies) indicated statistically significant higher credit earning rates at LP-UTR schools.**

TABLE 36. ANALYSIS OF COURSE CREDITS, SPECIAL EDUCATION STUDENTS YEARS 1–3 SCHOOLS

Outcome Area	LP-UTR Years 1 – 3 Schools			Non-LP-UTR Years 1 – 3 Comparison Schools			
	N Student	Adjusted Mean	SE	N Student	Adjusted Mean	SE	Sig
ALL SUBJECTS % Course Credits Earned	2539	**.75**	.01	8460	**.71**	.00	p<.01
ELA % Course Credits Earned	2520	**.77**	.01	8309	**.73**	.00	p<.01
MATH % Course Credits Earned	2462	**.72**	.01	7988	**.69**	.01	p<.01
SCIENCE % Course Credits Earned	2357	**.76**	.01	7739	**.70**	.01	p<.01
SOCIAL STUDIES % Course Credits Earned	2442	**.76**	.01	8176	**.73**	.00	p<.01

2017–18 School Year

Regents Examination Performance by English Language Learners

When looking at schoolwide Regents outcomes for ELs, 3 of the 9 contrasts showed LP-UTR schools averaged significantly higher scores, significantly lower on 2 contrasts, and the same on the remaining 4. The areas of advantage were in subjects most closely tied to language, specifically ELA and history as opposed to math and science. **While the examination results were mixed, the trend was that ELs performed at least similarly to their peers at matched schools.**

TABLE 37. ANALYSIS OF REGENTS SCORES, ENGLISH LANGUAGE LEARNERS YEARS 1-3 SCHOOLS

Outcome Area	LP-UTR Years 1 – 3 Schools			Non-LP-UTR Years 1 – 3 Comparison Schools			Sig
	N Student	Adjusted Mean	SE	N Student	Adjusted Mean	SE	
Regents ALL SUBJECTS # Passed at 55 or Above	1677	1.9	.04	4472	1.9	.02	p=.42
Regents ALL SUBJECTS Average Score	1677	**60.4**	.38	4472	**59.1**	.23	p<.01
Common Core ELA Regents Score	652	**50.5**	.97	2144	**48.0**	.54	p=.03
Common Core ALGEBRA Regents Score	915	61.1	.44	2600	61.6	.26	p=.38
Common Core GEOMETRY Regents Score	290	**57.8**	.97	727	**63.2**	.61	p<.01
ALGEBRA II / TRIGONOMETRY Regents Score	136	68.5	1.17	330	66.2	.74	p=.11
EARTH SCIENCE Regents Score	336	**55.5**	.95	726	**58.3**	.64	p=.02
LIVING ENVIRONMENT Regents Score	882	60.6	.49	2157	61.3	.31	p=.21
CHEMISTRY Regents Score	59	60.8	1.73	184	61.9	.98	p=.59
GLOBAL HISTORY Regents Score	815	**62.8**	.59	2139	**60.5**	.37	p<.01
U.S. HISTORY Regents Score	500	**67.8**	.72	1400	**64.0**	.43	p<.01

Spring 2018 NY Regents Examinations

Course Credits Earned by English Language Learners

When looking at schoolwide credit accumulation for ELs, the evidence of an LP-UTR effect compared to business-as-usual was more consistent than it was for exam performances. **All 4 of the contrasts indicated statistically significant advantages for ELs at LP-UTR schools** (see Table 38).

TABLE 38. ANALYSIS OF COURSE CREDITS, ENGLISH LANGUAGE LEARNERS YEARS 1– 3 SCHOOLS

Outcome Area	LP-UTR Years 1 – 3 Schools			Non-LP-UTR Years 1 – 3 Comparison Schools			Sig
	N Student	Adjusted Mean	SE	N Student	Adjusted Mean	SE	
ALL SUBJECTS % Course Credits Earned	2150	**.76**	.01	6121	**.71**	.00	p<.01
ELA % Course Credits Earned	2139	**.77**	.01	6059	**.72**	.01	p<.01
MATH % Course Credits Earned	2094	**.76**	.01	5833	**.70**	.01	p<.01
SCIENCE % Course Credits Earned	2039	**.73**	.01	5666	**.70**	.01	p<.01
SOCIAL STUDIES % Course Credits Earned	2109	**.77**	.01	5864	**.73**	.01	p<.01

2017–18 School Year

Thus, the overall finding is that while the examination results were mixed, ELs still tended to fare the same or better than peers at similar schools, and they also earned a higher percentage of course credits in all 4 domains examined (ELA, math, science, and social studies).

CONCLUSIONS

Over the five years of LP-UTR, project partners continued to successfully prepare effective teachers and teacher leaders in high-needs urban schools. A total of 108 new novice teachers entered the teaching field in New York City, close to 130 mentor teachers participated in cross-school learning, and nine school-based coaches worked to build teacher leadership capacity in their schools. Over the three cohorts, 83% of the residents completed the LP-UTR program. Of those who completed the residency (n=108), 77% went on to earn their Masters degrees and 69% obtained NYS Initial Certification. Combining the Learning Partners' framework and the UTR model for training teachers also extended learning beyond the classroom and school and into the larger teacher community.

This summative evaluation study sought to answer five key questions about LP-UTR—its evolution, its impact on novice and experienced teachers and school communities, and its impact on achievement.

Based on the above figures and findings shared throughout this report, the answer to the first question—*Did the LP-UTR partners meet their goal of recruiting teacher candidates who would succeed in the program and in the classroom?* — is a clear "yes."

The second question—*How did the LP-UTR program evolve over the three project years?* —is more nuanced because it is hard to document adjustments and incremental changes in a fluid project. What is clear is that changes were made in response to participants' needs. LFAs and problems of practice were revised to reflect triads' and schools' goals and give both more autonomy. Over the course of the program, the number of school-based coaches tripled to increase coaching capacity and allow more teacher leaders to emerge. To support the addition of TESOL candidates, Hunter increased offerings and supports each year, as needs arose. New Visions program staff continued to modify the suite of process monitoring assessments in an effort to better support residents' learning and mentors' and coaches' efforts to norm ratings. Data systems were improved to provide better data and easier access—and ensure that the work of the project was data-driven.

Answering the third question— *In what ways do novice and experienced teachers grow by participating in LP-UTR?* —is relatively easy for the LP-UTR residents, who were assessed on a suite of performance-based assessments to monitor and support learning. Residents were assessed by mentors and coaches at multiple points throughout the residency via the LP-UTR Danielson-based observation rubric. Over the year, average scores tended to show trajectories of growth in residents' ability to effectively plan instruction, manage their classroom, and use best practices. At the half-way point of the residency and again at the end of the year, residents participated in the capstone Defense of Learning assessment to demonstrate their impacts on student learning and use of assessment to inform their instruction. Each successive cohort achieved a greater percentage of its residents meeting or exceeding the established benchmarks on the Defense of Learning.

A core mission of the LP-UTR program is on the growth of novice *and* experienced teachers. While most teacher residency programs only focus on the growth of residents, LP-UTR focused on experienced teacher growth

through mentor and coach participation in cross-school collaborations. As part of teacher growth, residents and mentors received programmatic supports, such as, increased planning times and time for inquiry team meetings, in addition to supports from program staff and coaches. Residents and mentors also indicated they were more confident in a range of instructional skills and assessment and data practices through their participation in the LP-UTR program.

A core belief of the LP-UTR program that perhaps helps to distinguish it, as well as the UTR model, from similar teacher residencies is that—in addition to its main role of developing and retaining new cohorts of highly skilled and diverse teachers—the LP-UTR model is valued as a powerful means for supporting the professional growth of teacher leaders. The LP-UTR view is that the participating mentor teachers and coaches develop as leaders through the process of supporting their residents and in sharing with and learning from their mentoring and coaching peers. As it does for residents, the LP-UTR model provides ongoing training and other supports for its mentors and coaches. Moreover, LP-UTR residents, mentors, coaches, and even program leadership each engage in parallel learning processes of data-based decision making with a shared commitment to continually reflect and grow (Harford & Verdier, 2019). LP-UTR was therefore systematically directed towards capacity building at multiple levels within and across schools. As evidenced in other evaluations of initiatives focused on multi-level systems of learning (e.g., Moylan & Patel, 2018), such efforts can extend professional growth for educators across a spectrum of roles and experiences.

Perhaps the most difficult question to answer is *What effect did cross-school collaboration have on school support, teachers' growth, and collective efficacy?* Combining the Learning Partners framework and the Urban Teacher Residency program provided supports and training for mentor teachers and coaches to engage in a collaborative learning experience. Mentors from all three cohorts reported observing classrooms more frequently across schools as part of the triad work than within their own school. Residents reported having a less direct impact on combining the models, noting participation in visits to other schools increased in the second year of the program but involved less than half of residents. In general, residents did report being more engaged in school activities than in triad activities. About a quarter of mentors attributed growth in their ability to share skills and practices with teachers in other schools to LP-UTR cross-school sharing. Residents attributed the largest area of skill development to teaching and interacting with school colleagues, residents credited Hunter Coursework, LP-UTR Coaching, and New Visions PD and support with helping them to develop specific subsets of skills.

The question most often asked to gauge a project's value is does it work? The summative study findings answer *What impact does LP-UTR participation have on student achievement?* – in the affirmative.

Overall, students of LP-UTR teachers (both graduates and mentor teachers) outperformed the students of non-LP-UTR teachers at matched comparison schools, as measured by course examination scores and course credits earned. Additionally, across the core subjects, student achievement outcomes in LP-UTR graduate teacher classrooms tended to be comparable to or better than those observed in non-LP-UTR classrooms.

Students taught by the LP-UTR mentor teachers predominantly attained similar or greater levels of achievement than their peers taught by comparison mentor teachers. Additional comparative analyses of school-wide achievement data revealed that students with disabilities fared consistently better in core courses at LP-UTR schools than at non-LP-UTR schools. For English language learners, there was consistent evidence of a positive LP-UTR impact on course credit earning and evidence that average exam performances were at similar levels as found at comparison schools.

The introduction to this report alluded to views among educator and researchers that teacher residencies fill the space between traditional and alternative certification programs. In a focus group conducted at the end of LP-UTR's third year, a resident made the same observation, noting LP-UTR was the perfect middle ground between the traditional model of student teaching and the Teaching Fellows program. This and other feedback from participants indicate that the LP-UTR model not only occupies an important place in teacher preparation, but that it also helps resolve the perennial question about how to integrate theory and practice. Feedback and findings also suggest that LP-UTR should consider a reevaluation of the professional development that teachers typically receive in the post-training, post-credentialing part of their professional learning continuum. As LP-UTR has shown, teachers and teacher leaders, in collaboration with administrators, coaches, and even novice teachers, can serve as the providers of locally-defined PD that meets schools' and teachers' needs.

There are aspects of LP-UTR that this evaluation did not explore but that merit additional research, such as whether triad schools and teachers have continued to employ program structures, such as cycles of learning, or whether they continue to use strategies learned from cross-school colleagues. Tracking the continuing role and influence of school-based coaches and model teachers or teacher leaders could also gauge the durability of LP-UTR components and impact over time.

Our student achievement analyses also point to areas for further research that could shed light on the impact of residency-trained teachers on students' success and the impact of participation on mentors. Findings on credit accumulation showed that students with disabilities and English learners tended to fare better in LP-UTR schools. What the analyses didn't explore was why. Further data analyses, along with school visits, interviews, and tracking of these students, could provide more in-depth information about what supports success for these students. Findings also indicated that, in four out of five contrasts, students taught by LP-UTR mentors who participated multiple years out-performed their peers, at levels similar to those of students taught by LP-UTR graduates. Again, a more in-depth look, combined with larger samples and longitudinal tracking, could offer new insights.

The final item and long-term goal on LP-UTR's comprehensive logic model refers to sustainability. The durability of instructional strategies, coaching, and modeling is one measure of sustainability. Another valuable metric might be the degree to which LP-UTR participants—the returning mentors and schools, the model teachers, the LP-UTR graduates entering the workforce—represent a kind of critical mass that builds capacity beyond the LP-UTR project itself. Related research could also continue to track retention and mobility rates among LP-UTR graduates, past their four-year commitment, to understand the impact of the model.

References

American Association of Colleges for Teacher Education (2018). *A pivot toward clinical practice, its lexicon, and the renewal of educator preparation: A report of the AACTE Clinical Practice Commission*. Washington, DC: Author.

Bryk, A. S. (2009). Support a Science of Performance Improvement. *Phi Delta Kappan*, *90*(8), 597–600. doi: 10.1177/003172170909000815

Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting Ideas into Action: Building Networked Improvement Communities in Education. In M. T. Hallinan (Ed.), *Frontiers in Sociology of Education* (pp. 127–162). doi: 10.1007/978-94-007-1576-9_7

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to Improve: How America?s Schools Can Get Better at Getting Better*. Harvard Education Press.

Cochran-Smith, M., Villegas, A.M., Abrams, L., Chavez Moreno, L., Mills, T. & Stern, R. (2016). Research on teacher preparation: Charting the landscape of a sprawling field. In Gitomer, D. & Bell, C. (Eds.). *Handbook of Research on Teaching* (5th ed., pp.439–547). Washington, DC: American Educational Research Association.

Danielson, C. (2013). *Framework for Teaching Evaluation Instrument, 2013 Edition*. Chicago, IL: Danielson Group.

Darling-Hammond, L. (2004). Standards, accountability and school reform. *Teachers College Record, 106*(6), 1047-1085.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective Teacher Professional Development*. Palo Alto, CA: Learning Policy Institute.

Darling-Hammond L., Wei, R. C., Andree, A., Richardson N., & Orphanos S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad*. Palo Alto, CA: Stanford University, National Staff Development Council and The School Redesign Network.

Desimone, L. M., Garet, M. S. (2015). Best practices in teachers' professional development in the United States. Psychology, *Society and Education, 7,* 252–263.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*, 181–199. doi:10.3102/0013189X08331140

Egalite, A., Kisida, B., Winters, M.A.. "Representation in the classroom: The effect of own-race teachers on student achievement." Economics of Education Review, Volume 45, April 2015, 44-52.

Garet, M. S., Porte,r A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. American Educational Research Journal, 38, 915–945. doi:10.3102/00028312038004915

Gerdeman, D. R., Wan, Y., Molefe, A., Bos, J. M., Zhu, B., & Dhillon, S., (2017). Impact of TNTP's Teaching Fellows in urban school districts. Washington, D.C.: American Institute for Research.

Gershenson, Seth. "The alarming effect of racial mismatch on teacher expectations." Brookings, The Brown Center Chalkboard, Aug. 18, 2014. Available online at http://www.brookings.edu/blogs/brown-center-chalkboard/posts/2015/08/18-teacher-expectations-gershenson

Glazerman, S., Dolfin, S., Bleeker, M., Johnson, A., Isenberg, E., Lugo-Gil, J., Britton, E. (2008). *Impacts of comprehensive teacher induction: Results from the first year of a randomized controlled* study (NCEE 2009-4034). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://files.eric.ed.gov/fulltext/ED503061.pdf.

Glazerman, S., E., Isenberg, S., Dolfin, M., Bleeker, A., Johnson, Grider, M., & Jacobus, M. (2010). *Impacts of Comprehensive Teacher Induction: Final Results from a Randomized Controlled Study* (NCEE 2010-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Guha, R., Hyler, M.E., and Darling-Hammond, L. (2016). *The Teacher Residency: An Innovative Model for Preparing Teachers*. Palo Alto, CA: Learning Policy Institute.

Harford, M. & Verdier, R. (2019). Evaluating and Continuously Improving Program Quality in the Teacher Residency. In C.A. Torrez & M. Krebs (Eds.), *The Teacher Residency Model: Core Components for High Impact on Student Achievement* (pp. XX-XX). Chicago, IL: National Center for Teacher Residencies.

Hill, H. C. (2007). Learning in the teacher workforce. *Future of Children, 17*, 111–127. doi:10.1353/foc.2007.0004

Hill, H. C., Beisiegel, M., & Jacob, R. (2013). Professional development research: Consensus, crossroads, and challenges. *Educational Researcher, 42*(9), 476–487.

Ingersoll, R. & Strong, M. (2011). The Impact of Induction and Mentoring Programs for Beginning Teachers: A Critical Review of the Research. *Review of Educational Research, 81*(2), 201-233.

Joyce, B., & Showers, B. (1980). Improving inservice training: The messages of research. *Educational Leadership, 37*(5), 379-385.

Joyce, B., & Showers, B. (1982). The coaching of teaching. *Educational Leadership, 40*(1), 4–10.

Kennedy, M. M. (2016). How Does Professional Development Improve Teaching? *Review of Educational Research, 86*(4), 945–980. https://doi.org/10.3102/0034654315626800

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research, 88*(4), 547–588.

Moylan, A., Patel, R. (2018). *Evaluation of Leading with Learning i3 Development Initiative: Final Report*. San Francisco, CA: Rockman et al.

National Council for Accreditation of Teacher Education (2010). *Transforming teacher education through clinical practice: A national strategy to prepare effective teachers*. Report of the Blue Ribbon Panel on Clinical Preparation and Partnerships for Improved Student Learning. Washington, DC: Author.

New York City Department of Education (2019). Professional Learning. https://www.weteachnyc.org/approach/professional-learning/

Redding, C., Cannata, M., & Taylor Haynes, K. (2017). With Scale in Mind: A Continuous Improvement Model for Implementation. *Peabody Journal of Education 92*(5): 589-608. doi: 10.1080/0161956X.2017.1368635

Ronfelt, M., Farmer, S., McQueen, K., Grissom, J. (2015). *Teacher collaboration in instructional teams and student achievement.* American Educational Research Journal, 52(3), 475-514. Retrieved from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.915.1966&rep=rep1&type=pdf

Silva, T., McKie, A., Knechtel, V., Gleason, P., Makowsky, L. (2014). *Teaching Residency Programs: A Multisite Look at a New Model to Prepare Teachers for High-Need Schools* (NCEE 2015-4002). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Sloan, K., Allen, A., Blazevski, J., & Carson, F. (2017). *The MASTER Model: Preparation through Partnerships.* Bloomington, IN: Rockman et al.

Sloan, K., Allen, A., Blazevski, J., Carson, F., & Rockman, S. (2018). *A Different, More Durable Model: New Visions–Hunter College Urban Teacher Residency Project*. Bloomington, IN: Rockman et al.

Sloan, K., & Blazevski, J. (2015). *Measures of Success*. Bloomington, IN: Rockman et al.

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2015, July). *Teacher training, evaluation, and compensation intervention report: New Teacher Center Induction Model.* Retrieved from http://whatworks.ed.gov

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, September). *WWC review of the report: Impacts of comprehensive teacher induction: Final results from a randomized controlled study*. Retrieved from http://whatworks.ed.gov

Waterman, S., & He, Y. (2011). Effects of Mentoring Programs on New Teacher Retention: A Literature Review. *Mentoring & Tutoring: Partnership in Learning, 19*(2), 139–156. https://doi.org/10.1080/13611267.2011.564348

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, *37*(1), 3–14. https://doi.org/10.1016/j.stueduc.2011.03.001

Yoon K. S., Duncan T., Lee, S. W. Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007—No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest.

Zeichner, K. (2010). Rethinking the Connections between campus courses and field experiences in college- and university-based teacher education. *Journal of Teacher Education, 61*(1–2), 89–99. https://doi.org/10.1177/0022487109347671.

Appendix A: Process Monitoring

TABLE 39. RESIDENCY ASSESSMENT TIMELINE AND BENCHMARK TARGETS, COHORT 1

Assessment	Assessor	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
Danielson	Mentor		0		d1			d3		d4	
Danielson	Program Coach		0		d1			d3		d4	
Danielson	School Leader						d2				
Lesson Design	Mentor			48/68			54/68		54/68		
Lesson Design	Program Coach	0		48/68	48/68		54/68	54/68	54/68		
Unit Design	Mentor						50				
Unit Design	Program Coach						50		50		
Professionalism	Mentor	p1				p2				p3	
Defense of Learning	Program Coach					42/60	42				46/60

Key (cohort 1):

0 no minimum threshold / baseline
d1 Earning low 2 in at least 3 areas, including 2d
d2 Earning low 2 in at least 4 areas, including 2d
d3 Earning low 2 in all 6 areas, and a low 3 in 2d
d4 Earning 2 in all 6 areas, and a low 3 in 2d and 3d
p1 No 1s in any area except parent outreach, 35
p2 45, no 1s
p3 55, no 1s

TABLE 40. RESIDENCY ASSESSMENT TIMELINE AND BENCHMARK TARGETS, COHORT 2

Assessment	Assessor	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
Danielson	Mentor		0		d1			d3		d4	
Danielson	Program Coach		0		d1			d3		d4	
Danielson	School Leader						d2				
Lesson Design	Mentor	0		48			51/68		55/68		
Lesson Design	Program Coach	0		48			51/68		55/68		
Unit Design	Mentor						30/40				
Unit Design	Program Coach						30/40		30/40*		
Professionalism	Mentor	p1				p2				p3	
Defense of Learning	Program Coach (w/ Mentor input)					45/60					48/60

Key (cohort 2):

0 No minimum threshold / baseline
d1 Earning at least a 2 rating on 4 areas of rubric from both mentor and coach, including 2 in classroom management
d2 Earning at least a 2 rating on 4 areas of rubric from School Leader, including 2 in classroom management
d3 Earning high 2 rating on classroom management (2d); 2 ratings on all other areas of Danielson (no 1s or low 2s) from both mentor and coach
d4 Earning high 2 rating on classroom management (2d) and assessment (3d); 2 rating or above on all other areas of Danielson from mentor and coach
p1 51/72 (no 1s in any category except for parent outreach)
p2 54/72 (3s on punctuality, attendance, dress; 2s on other ok)

p3 58/72 (no 1s in any area, and no more than 2 categories at level 2)
* Mini unit design rubric for revisions only

TABLE 41. RESIDENCY ASSESSMENT TIMELINE AND BENCHMARK TARGETS, COHORT 3

Assessment	Assessor	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
Danielson	Mentor		0		d1			d3		d4	
Danielson	Program Coach		0		d1			d3		d4	
Danielson	School Leader						d2				
Lesson Design	Mentor	0		48/68			51/68		55/68		
Lesson Design	Program Coach	0		48/68			51/68		55/68		
Unit Design	Program Coach						18/24		21/28*		
Professionalism	Mentor	p1				p2				p3	
Defense of Learning	Program Coach (w/ Mentor input)					45/60					48/60

Key (cohort 2):
0 No minimum threshold / baseline
d1 Earning at least a 2 rating on 4 areas of rubric from both mentor and coach, including 2 in classroom management (unchanged from C2)
d2 Earning at least a 2 rating on 4 areas of rubric from School Leader, including 2 in classroom management (unchanged from C2)
d3 Earning high 2 rating on classroom management (2d); 2 ratings on all other areas of Danielson (no 1s or low 2s) from both mentor and coach (unchanged from C2)
d4 Earning high 2 rating on classroom management (2d) and assessment (3d); 2 rating or above on all other areas of Danielson from mentor and coach (unchanged from C2)
p1 no 1s in any category except for parent outreach
p2 2/36 (3s on punctuality, attendance, dress)
p3 no 1s in any area, and no more than 2 categories at level 2
* Mini unit design rubric for revisions only

TABLE 42. LP-UTR DANIELSON-BASED RUBRIC DATA, COHORT 2

Time	Residents with data available		Met benchmark*	
	Frequency	Percentage of All Enrolled Residents (n=48)	Frequency	Percentage of Residents with Data that Met Benchmark
T1 (Oct)	45	94%	N/A	N/A
T2 (Dec)	43	90%	43	100%
T3 (Feb)**	21	44%	21	100%
T4 (March)	40	83%	34	85%
T5 (May)	35	73%	29	83%

* All available scores had to meet or exceed benchmark for resident teacher to be identified as having met benchmark.
** By design, only school administrators administered the rubric during this third time period.
N/A = not applicable, no benchmark established for baseline

TABLE 43. LP-UTR DANIELSON-BASED RUBRIC SCORES BY TIMEPOINT, COHORT 2

	N	Minimum	Maximum	Mean	SD
T1 (Oct)	91	9.00	24.00	14.38	2.67
T2 (Dec)	81	10.00	24.00	15.79	2.84
T3 (Feb)	21	12.00	22.00	16.00	2.32
T4 (March)	73	10.00	23.00	16.41	2.53
T5 (May)	59	11.00	24.00	17.44	2.84

Sumscores could range from 6 to 24.

TABLE 44. LP-UTR DANIELSON-BASED RUBRIC BY ASSESSOR ROLE, COHORT 2

Time	Combined Mentor and Coach Scores			Assessor Role					
				Mentor			Coach		
	N	M	SD	N	M	SD	N	M	SD
T1	91	14.38	2.67	46	14.63	3.20	45	14.13	2.00
T2	80	15.81	2.85	37	16.51	3.21	43	15.21	2.37
T3	68	16.51	2.46	32	17.13	2.92	36	15.97	1.84
T4	57	17.53	2.84	26	18.35	3.32	31	16.84	2.19

For Cohort 3, there were four timepoints during the residency year when mentors and coaches used the LP-UTR Danielson-based Rubric and one timepoint (midyear) when school leaders used the rubric. In addition, some residents used the rubric for a self-assessment and submitted their scores.

TABLE 45. LP-UTR DANIELSON-BASED RUBRIC DATA, COHORT 3

Month	Residents with data available		Met benchmark*	
	Frequency	Percentage of All Enrolled Residents (n=48)	Frequency	Percentage of Residents with Data that Met Benchmark
T1 (Sep-Oct)	47	98%	N/A	N/A
T2 (Nov-Dec)	44	92%	39	89%
TA (Jan-Feb)**	17	35%	17	100%
T3 (Feb-Mar)	22	46%	19	86%
T4 (Apr-May)	40	83%	29	73%

* All available scores had to meet or exceed benchmark for resident teacher to be identified as having met benchmark.
Benchmarks for Cohort 3:
T1: Baseline. No benchmark established.
T2: Earning at least a 2 rating on 4 areas of rubric from both mentor and coach, including 2 in classroom management
TA: (School Administrators only) Earning at least a 2 rating on 4 areas of rubric from School Leader, including 2 in classroom management
T3: Earning high 2 rating on classroom management (2d); 2 ratings on all other areas of Danielson (no 1s or low 2s) from both mentor and coach
T4: Earning high 2 rating on classroom management (2d) and assessment (3d); 2 rating or above on all other areas of Danielson from mentor and coach
N/A = not applicable, no benchmark established for baseline
** Scores from school administrators only.

TABLE 46. LP-UTR DANIELSON-BASED RUBRIC SCORES BY TIMEPOINT, COHORT 3

	N	Mean	SD
T1 (Sep-Oct)	99	13.71	2.54
T2 (Nov-Dec)	91	15.81	3.17
TA (Jan-Feb)*	17	16.24	2.70
T3 (Feb-Mar)	32	16.13	3.30
T4 (Apr-May)	69	16.61	2.92

Sumscores could range from 6 to 24.
* Scores from school administrators only.

TABLE 47. LP-UTR DANIELSON-BASED RUBRIC BY ASSESSOR ROLE, COHORT 3

Time period	Combined Mentor and Coach Scores			Assessor Role					
				Mentor			Coach		
	N	M	SD	N	M	SD	N	M	SD
T1 (Sep-Oct)	91	13.52	2.47	51	13.63	2.81	40	13.38	2.00
T2 (Nov-Dec)	83	15.61	3.10	50	16.02	3.44	33	15.00	2.40
T3 (Feb-Mar)	28	15.57	3.06	14	14.86	3.66	14	16.29	2.23
T4 (Apr-May)	66	16.48	2.91	36	16.47	2.92	30	16.50	2.95

Appendix B: Data Needs & Logic Model

TABLE 48. ROCKMAN DATA NEEDS, YEARS 1-3

School-level Data Needed for School Selection (2014-15 Baseline Data, much of which is publicly available)	**2015-16 Teacher and Student Data Needed**
1. School Name 2. Participation in LP 3. LP Focus Area 4. Presence of UTR-trained Teachers 5. School DBN 6. Location (borough) 7. School Type (e.g., grade range) 8. Total enrollment 9. % Black 10. % Hispanic 11. % Female 12. % ELL 13. % Economically Disadvantaged 14. % with Disabilities 15. Average Grade 8 reading proficiency 16. Average Grade 8 math proficiency 17. Weighted Regents Pass Rate for a. English b. Math c. Science 18. Attendance Rate 19. 4-year graduation rate 20. AND for each of our LP-UTR schools, we need the list of "peer schools / comparison group" used in the most recent NYC DOE School Quality Report (see snapshot on p.3)	**Once the 35 schools have been selected, pull the following data:** **For <u>all teachers within the 35 schools</u> involved in the sample, the following data is needed:** 1. School Name & DBN 2. Teacher Name & ID* 3. Certification 4. Current Teaching Assignment 5. Years of Teaching Experience 6. LP-UTR "Resident" or "Mentor" designation **For <u>all students enrolled at one of the 35 schools</u> involved in the sample, the following data is needed:** 1. Student ID* 2. School Name & DBN 3. Grade level 4. Credits earned during SY 2015-16 in each of the following subject areas (reported separately): ELA, Math, Science, Social Studies <u>and</u> TOTAL credits earned during SY 2015-16 5. Regents Pass Rates for all Regents test taken during SY 2015-16 (passed at 55% & 65%) 6. (TESOL students) New York State English as a Second Language Achievement Test (NYSESLAT) or other appropriate measures 7. Ethnicity 8. ELL 9. Disability 10. Economically Disadvantaged 11. 8th Grade Reading Proficiency 12. 8th Grade Math Proficiency **Sub-Sample of Teachers where COURSE ROSTERS are NEEDED:** IF current teaching assignment is ELA, and years of teaching experience = "any", THEN pull the ELA course rosters for those teachers: **Course Rosters to Include:** 1. Teacher ID* 2. Course Name & Code 3. Semester 4. Student IDs* for all students enrolled in each teacher's ELA courses (roster)

Residency goals & outcomes Partnership goals & outcomes

LP-UTR Logic Model

REA evaluation LPP evaluation Shared evaluation

PROGRAM INPUTS & ACTIVITIES	SHORT-TERM OUTCOMES	INTERMEDIATE OUTCOMES	LONG-TERM OUTCOMES
1. Resident & Mentor Recruitment: New Visions–Hunter highly selective residency program, validated by research on impact on student achievement, teacher practice, & retention • Selection of residents through screening by NYCDOE, Hunter, New Visions; Involvement of school leaders, mentors • Selection of highly-qualified mentors in the same high-need content areas, for 1:1 mentoring for residents. 2. Resident Preparation & Induction • 18-month program: summer training, with CCSS-aligned Hunter coursework & support for exams • Supervised clinical experience in ELA, Math, Science, TESOL, Sped., co-teaching alongside skilled mentors. • Performance-based assessments • Placement & hiring support for graduates • Induction (Hunter practicum seminar & field supervision) & new teacher support in host schools. 3. Mentor Training & Development • 20-hour training course for mentors on best practices in • mentoring & inquiry process • On-going training through monthly coaching, quarterly PD • Culture of reflection, collaboration in host schools 4. Development of Learning Partner Host & Triad Sites • Partner support for host schools • Designated Site Director • NYCDOE Facilitator(s) • Principal Support, incentives • Model Teacher Designation 5. Inter- & Intra-School Collaboration • Intervisitations; online discussions • Resource sharing • Shared practices around learning focus areas (LFAs) • Action research & Implementation plans around LFAs 6. Expanded Partnerships • Cross-school learning, NYCDOE's Learning Partners Program • Shared expertise in particular Learning Focus Areas (STEM, ELLs, professional growth & leadership) • Increased collaboration: New Visions, Hunter, NYCDOE	1a. 168 highly-qualified residents recruited for the program, including approx. 25% STEM. 1b. 168 mentors recruited, trained. 2a-b. 85% of residents successfully complete 14-month residency (approx. 15% attrition), earn NYS teacher certification, & obtain NYC DOE jobs; earn Masters degree in 18 months. 2c. Residents gain skills, confidence in using data-driven inquiry; meet benchmarks on Danielson rubric & measures of student learning. 2d. Enhanced induction year support. 3a-d. Mentor competency rubric scores improve; mentors promote strong professional practice school-wide. 4a-b. Residents, mentors, site coordinators in triad partner schools incorporate shared practices. 4b. Increased support from NYCDOE & facilitators, around LFAs. 4c. Principals more actively support collaboration. 4d. Host LP-UTR model teachers share practices with triad partner schools. 5a-b. Teachers share practices & resources within, across LP-UTR sites. Additional support for residents, mentors, graduates via networking, online discussions. 5c. Schools share strategies, practices around LFAs (e.g., TESOL). 5d. Teachers share action research insights; consistency in implementation 6a. Increased stakeholder engagement, sharing. 6b. Development of support networks among LP-UTR schools. 6c. PD & shared practices in LFAs linked to LP-UTR goals.	1. 92% of residents successfully complete 1st year teaching in NYCDOE public schools. 2a–c. Higher %s of students in ELA, Math, Science, other classes taught by LP-UTR 1st-yr teachers accumulate required credits than peers taught by non-LPP-NTTP 1st-yr teachers, controlling for prior performance. 2d. Higher % of ELL & Sped learners taking standardized exam perform at levels closer to non-ELL, Sped peers than students taught by non-LPP-NTTP prepared teachers. 3a-d. Higher % of students in ELA, Math, Science, & other classes taught by all mentors (incl. LP-UTR) accumulate required credits than peers taught by non-LPPNTTP teachers, controlling for prior performance. 4a-c. Triad partner schools become LP-UTR host schools. 4d. Increased school responsibility & accountability for teacher development. 4e. Model teachers define & take on leadership roles. 5a-b. Higher performance on key measures among host schools. Improved performance among LP-UTR schools. 5c-d. Links between LPN-PPT implementation fidelity & performance, school quality. 6a. LP-UTR participants, partners promote stronger professional practice school-wide. 6b. Improvements in LFAs among LP-UTR schools. 6b. Changes in program, coursework reflect emphasis on shared practice in LFA areas. 6d. Increased focus on new-teacher-support & improved practice among partners.	1. 80% of UTR teachers successfully complete 3rd year of teaching in NYC public schools. Retention rates among UTR-trained teachers exceed city-wide rates by at least 5%. 2a–c. ELA, Math, & Science Regents &/or other standardized exam passing rates for LPP-NTTP taught students will be higher than for non LPP-NTTP-taught students, controlling for prior performance. 3a–b. ELA, Math, & Science Regents &/or other standardized exam passing rates for students taught by LP-UTR mentors will be higher than for students taught by matched comparison teachers, controlling for prior performance. 3- 4a. A higher % of students in ELA, Math, Science, & other classes in LP-UTR s will accumulate required credits in each subject than peers in matched comparison schools, at statistically significant levels, controlling for prior performance. 4a-e. ELA, Math, Science Regents &/or other standardized exam passing rates for students in LP-UTRs will be higher than for students in matched schools, controlling for prior performance. 5a-d. Through LPN-PPT collaboration, learning across schools will increase system-wide. 6a-c. Evidence of residency & LPP models' sustainability. Evidence of residency & LPP scalability through LPN-PPT collaboration.