Evaluation of Georgia State University's CREATE Teacher Residency Program – Cohorts 3 through 5

A FINAL REPORT OF A QUASI-EXPERIMENT IN GEORGIA

March 31, 2023

Andrew P. Jaciw

Sze-Shun Lau

Jenna Zacamy

Li Lin

Empirical Education Inc.

ACKNOWLEDGEMENTS

We are grateful to the research study participants, CREATE program staff, staff at Georgia State University's College of Education and Human Development (GSU CEHD), and participating schools and districts for their assistance and cooperation in conducting this research. We also appreciate Audra Wingard and Thanh Nguyen, former Senior Research Managers, and Zahava Heydel, former Senior Research Assistant at Empirical Education Inc., for their contribution to this project and report. This work has been supported by the U.S. Department of Education's Supporting Effective Educator Development (SEED) program, through Award Number U423A170027. The opinions expressed are those of the authors and do not represent views of the U.S. Department of Education. As the independent evaluator, Empirical Education Inc. was provided with independence in reporting the results.

ABOUT EMPIRICAL EDUCATION INC.

Empirical Education Inc. is a Silicon Valley-based research company that provides tools and services to help K–12 school systems make evidence-based decisions about the effectiveness of their programs, policies, and personnel. The company brings its expertise in research, data analysis, engineering, and project management to customers that include the U.S. Department of Education, educational publishers, foundations, leading research organizations, and state and local education agencies.

Reference this report: Jaciw, A. P., Lau, S., Zacamy, J., & Lin, L. (2023). Evaluation of Georgia State University's CREATE Teacher Residency Program – Cohorts 3 through 5: A Final Report of a Quasi-Experiment in Georgia. (Empirical Education Rep. No. Empirical_GSU-7035-FR-2023-O.2). Empirical Education Inc. https://www.empiricaleducation.com/create/

Table of Contents

Chapter 1. Introduction

The Collaboration and Reflection to Enhance Atlanta Teacher Effectiveness (CREATE) teacher preparation program seeks to raise student achievement in high-needs schools in the Atlanta metropolitan area by increasing teacher effectiveness and retention of both new and veteran racially diverse educators. CREATE aims to achieve this by developing compassionate, skilled, and antiracist educators who are committed to teaching practices that prioritize racial justice and interrupt inequities.

Empirical Education has partnered with CREATE to serve as an external evaluator since 2015. Our evaluative work of CREATE's teacher residency program for early career teachers began under an Investing in Innovation (i3) grant in 2015 (with the first two cohorts of participants) and continued with funding through the Supporting Effective Educator Development (SEED) grant issued in 2017. The work covered by this SEED grant allowed for the continuation of additional cohorts of CREATE residents and early career teachers (covering Cohorts 3 through 5, from 2017 through 2022), as well as additional components, including the evaluation of the effectiveness of the Professional Learning (PL) opportunities CREATE offers to veteran educators.[1]

This evaluation aims to determine whether there is a positive impact of the CREATE Teacher Residency program on teacher and student outcomes by comparing the experiences of study participants who are in the CREATE residency program to a comparison group of study participants also working towards their teaching credential through Georgia State University College of Education and Human Development's (GSU CEHD) traditional credentialing program, but are not participants of the CREATE program. Results of this study will inform teacher preparation, effectiveness, and retention policies and practices across the state of Georgia. It will also contribute to the limited but growing body of research on the effectiveness of teacher preparation programs for preservice teachers (Meadows & Theodore, 2012), and particularly of residency programs that support a racially diverse educator workforce.

This chapter of the report continues to summarize the CREATE study under the SEED grant, including a description of the CREATE and comparison group programs, a synopsis of findings generated through the analysis under the i3 grant, and contextual information for the research. Chapter 2 presents an overview of the study methodology, participant recruitment processes, and data collection sources. Chapter 3 includes the analysis of the key components of fidelity of implementation (FOI), as well as descriptive findings from the experiences of both CREATE and comparison group participants. Chapter 4 provides results on the impact of CREATE on teacher self-compassion, levels of burnout, and stress management and empathy related to teaching, as measured through surveys. Chapter 5 presents the impact of CREATE on student achievement in, as measured by the Georgia Milestones Assessment System (Georgia Milestones). Chapter 6 provides findings

[1] Since receiving this original SEED grant in 2017, CREATE has been awarded two additional SEED grants: one in 2020, and another in 2022. These grants allow for a longitudinal evaluation of ten total cohorts of participants.

of the impact of CREATE on teachers' early career trajectories and on teacher retention. We discuss the significance and implication of the findings, and offer conclusions, in Chapter 7.

OVERVIEW OF THE CREATE RESIDENCY PROGRAM

At the core, CREATE is a three-year teacher residency program that offers support for those working towards their teaching credential at GSU CEHD and into their first two years serving as an educator. Participation as a CREATE resident begins during a teacher-in-training's preservice teaching year, while residents are completing their credential at GSU CEHD. It continues for three years through their second year of teaching as a credentialed teacher-of-record. During the three-year residency, CREATE residents receive intensive support through their entry into their teaching career in the form of frequent meetings with their resident cohort (which stays consistent throughout the residency), opportunity to co-teach during the first year as a classroom teacher, receiving support from CREATE-trained mentors, and participating in summer and school-year professional learning. CREATE moves towards its vision of developing "compassionate, skilled, antiracist, educators for every student" (personal communication, 2022) by 1) fostering collaboration around justice-centered practices, 2) challenging oppressive schooling structures, and 3) elevating liberatory practices in each of its PL opportunities. Appendix A includes further details about the programming CREATE offers for novice and experienced educators.

The programming intends to 1) increase teacher collaboration through mentoring and involvement in collaborative learning communities, 2) reduce the stress that often accompanies the early years of teaching, 3) increase collegiality and teacher support, and 4) increase opportunities and abilities to apply antiracist and critical conscious lenses to the work of teachers. CREATE hypothesizes that these short-term outcomes will mediate teachers' use of research-based pedagogical strategies that impact students' acquisition of key knowledge and skills and the development of flexibly-managed, safe, and culturally responsive environment conducive to learning. In turn, CREATE conjectures that these teacher and classroom outcomes will lead to positive effects on student achievement and retention of teachers, particularly teachers of color (Figure 1).[2]

In this report, "Year 1" refers to study participants' first year in the study, which is their preservice teaching year, as well as the CREATE residents' first year in the CREATE program. "Year 2" refers to study participants' second year in the study and first year as teachers, as well as CREATE residents' second year in the CREATE program. "Year 3" refers to study participants' third year in the study and second year as teachers, as well as CREATE residents' third and final year in the CREATE program. Table 1 displays this visually.

[2] The CREATE program has evolved over the years. The logic model in Figure 1 reflects the CREATE program at the time of the study.

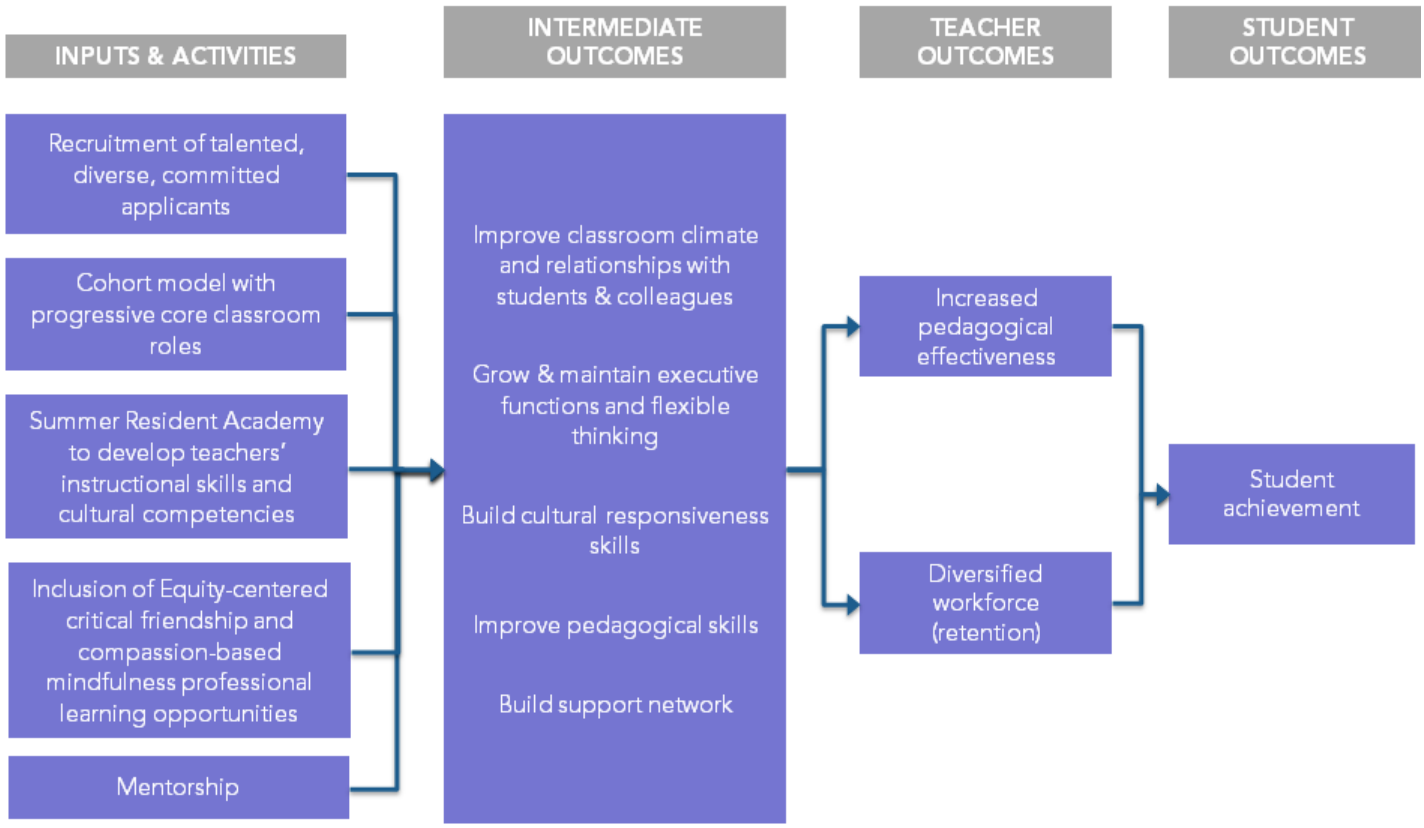

FIGURE 1. CREATE TEACHER RESIDENCY PROGRAM LOGIC MODEL

OVERVIEW OF THE COMPARISON PROGRAM

Study participants in the comparison (non-CREATE) group complete the traditional credentialing program at GSU CEHD in the Early Childhood and Elementary (ECEE) and Middle and Secondary Education (MSE) departments, with the goal of teaching in kindergarten through eighth grade classrooms in core subject areas. Students in GSU's traditional credentialing program complete a variety of requirements, including GSU CEHD coursework and an in-school practicum (which CREATE residents also complete). Students in this program complete their student teaching during their final year of enrollment at GSU CEHD, and then, after graduating with their teaching credential, typically move on to serve as full-time teachers-of-record without participating in further programming by GSU CEHD. Study participants in the comparison group may complete their practicum either in Atlanta Public School (APS) or in another nearby district. While some comparison participants completed their student teaching in APS schools, the majority completed their practicum across nine neighboring districts in the Atlanta metropolitan area.

SUMMARY OF I3 FINDINGS OF INITIAL CREATE COHORTS

The analysis conducted under the i3 grant followed the first two cohorts of CREATE residents and matched comparison groups. Cohort 1 began their residency with CREATE during their student teaching year in the

2015–16 school year, and Cohort 2 began their CREATE residency during their student teaching year in the 2016–17 school year. As with Cohorts 3 through 5, the study followed participants for three years, through their second year serving as a teacher-of-record.

For teachers, we examined exploratory intermediate outcomes including measures of executive functioning, self-efficacy, commitment to teaching, and retention in teaching, as well as confirmatory outcomes including ratings of instructional strategies and positive learning environment. For students, we examined confirmatory outcomes including ELA achievement, math achievement, and general (ELA and math) achievement.

The study found that for early career teachers**,** there was not a statistically significant impact of the CREATE program across the full sample on any of the measures of executive functioning, self-efficacy, or commitment to teaching. However, the impacts of CREATE on Black educators, were all in a positive direction: 0.388 (p = .175) for resilience, 0.146 (p = .460) for mindfulness, 0.370 (p = .191) for self-efficacy in teaching, 0.026 (p = .895) for commitment to teaching, and 0.700 (p = .042) for stress management. The differential impacts on Black educators are also in a positive direction for all five outcomes, with two being statistically significant (resilience and self-efficacy) and one being marginally significant (stress management related to teaching).

Teacher effectiveness was measured using teachers' scores on Teacher Assessment on Performance Standards, or TAPS (the observation component of the teacher evaluation system in Georgia). The study found no statistically significant effect of CREATE on the TAPS instructional strategies professional standard (ES = -0.339, p = .221) or on the positive learning environment performance standard (ES = -0.557, p = .192). We speculated that the low level of variability observed in TAPS scores points to a broader difficulty in evaluating early-career teachers, and the tendency of existing rating systems—both in human factors pertaining to judgments of behaviors of early career teachers, and in the technical limitations of the instrument themselves—to be unable to effectively distinguish between low- and high-performing teachers.

There were no statistically significant effects of CREATE on ELA achievement (ES = -0.122, p = .454), math achievement (ES = -0.175, p = .569), or general (ELA and math) achievement (ES = -.139, p = .234), as measured by the Georgia Milestones Assessment System (Georgia Milestones). However, given limitation in the availability of data in specific grades (4–8) and content areas, the sample size for these analyses was very small, with the potential for random sampling error dominating the results.

Analysis of retention of participants in the teaching field revealed promising trends. Exploratory analyses showed a positive and statistically significant impact on undisrupted retention over a three-year time period (spanning graduation from GSU CEHD, entering teaching, and retention into the second year of teaching) for the CREATE group, compared to the comparison group (p = .038). We also observed that higher continuous retention among Black educators in CREATE, relative to those in the comparison group (p = .021), is a large driver of the favorable impact.

In our continuation of the evaluation of the CREATE program through additional funding from SEED, we expected the program's focus on fostering racial equity work and collaborative spaces to strengthen CREATE's

impact on multiple measures of effectiveness for new educators. For more information, please refer to the full report of results found under the i3 grant, which can be accessed here, and the associated appendices, located here.

Context for the Research Conducted Under the 2017 SEED Grant

It is imperative to acknowledge the sociopolitical context within which this research took place, particularly in the years during which many of the outcomes were assessed. The experiences of teachers and in school communities were impacted tremendously by shifts in structures of everyday life and political upheaval.

The COVID-19 pandemic that struck in spring 2020 set into motion changes to the classroom that required teachers, students, and other school staff to adjust quickly, including school closures, a shift to remote learning, and added responsibilities of considering public health policies in their work. During the time these and other pandemic-related adjustments went into effect, the participants in this study were in either their student teaching year, or their first or second year as full-time-teachers. These participants, who were all in the early stages of their careers as teachers, suddenly were witness to, and responsible for, managing technology needs, supporting their students while they learned from home, adjusting to a completely new modalities of teaching, and monitoring everchanging instructions from their district and local leaders. Participants who were students at GSU simultaneously had to adjust to shifts in modalities that their own schooling underwent, while continuing to balance their personal lives. CREATE supports, including Together Time meetings and professional learning opportunities for veteran educators, transitioned to a virtual format. Through the following school year (2020–21), schools in Georgia employed remote learning models, and then began to integrate hybrid models that required teachers to teach both virtually and in-person to serve the health and safety needs and preferences of their students and their families. Teachers continue to adapt their instructional styles in response to the ongoing pandemic.

Across the country, these changes in the way schools are run prompted contentious debates about when and how school re-openings should take place, accompanied by arguments about how the jobs of teachers impact the ability for the economy to function properly. In addition to an increased national discourse that pinpointed in-person schooling as a central tenant in allowing parents and caregivers of school-aged children to resume their own work effectively, sociopolitical conversation and action around racial justice and police brutality were ignited in the summer of 2020 following the murders of George Floyd and Breonna Taylor by law enforcement officers. Protests and other forms of activism gained fuel, calling for the addressing of anti-Black racism and the existing system of policing that are harmful to communities, particularly communities of color. These calls for racial equity and investing in community occurred across the United States and were broadcast widely, bringing issues of systemic and community injustice to the forefront of the minds of many. Furthermore, in Georgia in early 2021, both senate seats were won by Democrats for the first time since 2005, bringing national attention to Georgia's 2020 primary and 2021 runoff elections, as they ultimately tipped the U.S. senate to be majority Democrat.

This research was also conducted in the context of continuous debate over what teachers should be responsible for or barred from teaching, in terms of history and oppression, longstanding fear and lack of policy change around gun ownership and gun violence in schools, and the exacerbation of a shortage of teachers and support staff. In Georgia itself, House Bill 1084 was passed in 2022, limiting and skewing the discussions of race in schools. Because effective teaching requires considering the impact of social forces on students' lives, teachers at this time often held an additional role in supporting their students through processing the social action they were hearing about and experiencing in their own communities. For teachers, this responsibility was an added layer on top of processing and adjusting their own personal values, actions, and community membership. Many of the challenges also exacerbated existing inequities in educational access and opportunities.

When we asked in one of our study surveys during the 2020–21 school year about the biggest challenges participants were facing in their roles as educators, a second-year teacher responded: "Student interaction has decreased dramatically. I host Office hours every day, but it has been rare for students to join, and even rarer that we discuss history." A first-year teacher shared: "It has been hard to enter school as a first-year teacher where everything is so different for how a normal school year would typically look or operate. It has been a very big change as a teacher and a big transition period to teaching. It has also been hard not being able to meet my students and build those in person bonds." Many responses mentioned disengagement from students or families, distractions at home as teaching occurred virtually, and the lack of ability to develop personal connections with students. We also saw drastic changes in study participation between the 2019–20 and 2020–21 school year, when multiple participants left the study after being deemed ineligible because they were unable, or chose not, to secure a teaching position.

A study published in November 2022 found that teachers in the United States were 40 percent more likely to report experiencing anxiety symptoms than healthcare workers between September 2020 and March 2021 (Kush et al., 2022). Additionally, it has been reported that between February 2020 and May 2022, approx. 300,000 public school teachers and staff left the field. This further reduced the workforce by 3% (Dill, 2022). Recognizing the shifts and events that were taking place on a broader level inform our understanding of influential factors that may have had an impact on the wellbeing of teachers, both those participating in this study and beyond. While it is important to recognize this context when considering any research around education conducted during this time, this context is especially crucial when the research focuses on programs that aim to both strengthen supports offered to early-career teachers and increase the retention of equity-driven effective teachers.

In addition to the turbulence and sudden changes that occurred in school operations affecting the wellbeing of the education, careers, and personal lives of the teachers participating in this research study, changes in the availability and accessibility of student and teacher data also impacted the study. Teachers' TAPS scores, used as measures of teacher effectiveness based on performance on defined standards, were unavailable to us through the Georgia Department of Education for spring 2020 and 2021. Additionally, Georgia Milestones scores, which assess student achievement in ELA and math, were only available for the 2020–21 school year at the time of analysis. Due to COVID, schools did not administer standardized tests in spring 2020.

KEY RESEARCH QUESTIONS

In alignment with standards for the SEED grant, this study addresses distinct confirmatory and exploratory research questions.

The implementation evaluation investigates the following questions.

1. Were the key components of the CREATE teacher residency logic model implemented with fidelity?
2. What is the experience of study participants in the CREATE teacher residency program and in the comparison group, specifically with regard to level of support and mentorship?

The impact evaluation of the CREATE teacher residency program addresses the following confirmatory research questions.[3]

3. Is there a positive impact of CREATE on student achievement in elementary and middle grades on average three years after the start of residency?
4. Is there a positive impact of CREATE, compared to the GSU BAU program, on teacher retention three years after start of residency?

We also measure exploratory impacts on intermediate outcomes and differential impacts on teacher retention with a focus on the following research question.

5. Is there a positive impact of CREATE on self-reported levels of self-compassion, burnout, and stress management and empathy related to teaching, three years after entry into the residency program?
6. Do the impacts on the outcomes in question 5 increase with consecutive cohorts?
7. Is there a positive impact of CREATE, compared to the GSU BAU program, on teacher retention three years after start of residency among Black educators?
8. Is there a positive impact of CREATE, compared to the GSU BAU program, on teacher retention through the third year after start of residency, among teachers who completed teaching in the second year of residency?
9. What is the impact of CREATE on completion of the teacher preparation program at GSU CEHD and teacher retention into the first year of teaching for the overall sample?
10. Are impacts on teacher retention after 1, 2, and 3 years increasing by cohort for the overall sample and among Black educators?

[3] We planned to address a third confirmatory research question: "Is there a positive impact of CREATE on teacher instructional outcomes, as assessed through the Teacher Assessment of Performance Standards (TAPS), on average three years after start of residency?" by assessing the performance standards of quality of instructional strategies and quality of learning environment; however, due to the unavailability of TAPS scores, we were unable to analyze this question.

Chapter 2. Study Methods

CHAPTER OVERVIEW

We conducted a quasi-experimental study to evaluate the impact of CREATE on 1) teachers' measures of self-compassion, burnout and stress management, and empathy; 2) student achievement; and 3) teacher retention. The design compared outcomes for CREATE participants with those of similar GSU CEHD participants who did not enroll in CREATE.

PARTICIPATION TIMELINE

The quasi-experimental evaluation follows three staggered cohorts of study participants for three years each. The three years of participation in the study comprise study participants' preservice teaching year, followed by their first two years as full-time classroom teachers. CREATE expects their residents to spend the three study years working in APS, first as student teachers, and then as teachers-of-record. Study participants in the comparison group serve as student teachers and then teachers-of-records in APS, as well as neighboring districts and other districts in the state of Georgia. This SEED grant funded the study participation of the third (referred to as "Cohort 3"), fourth ("Cohort 4"), and fifth cohorts ("Cohort 5") of the study. Cohort 3's participation in the research began in the 2017–18 school year, funded by an Investing in Innovation (i3) grant, and continued to be funded via this SEED grant beginning in the 2018–19 school year. Cohort 4's participation in the study began in the 2018–19 school year, and Cohort 5's participation began in the 2019–20 school year (Table 1). Appendix B provides information about the process of recruiting CREATE and comparison participants for the research study.

TABLE 1. CREATE RESEARCH STUDY TIMELINE FOR COHORTS 3, 4, AND 5

	2017–18	2018–19	2019–20	2020–21	2021–22
Cohort 3	Year 1	Year 2	Year 3		
Cohort 4		Year 1	Year 2	Year 3	
Cohort 5			Year 1	Year 2	Year 3

Note. Light blue cells indicate groups that were funded under the i3 grant, rather than the SEED grant.

SAMPLE

At the time of recruitment, all study participants 1) were enrolled in GSU CEHD in either the ECEE or MSE departments, 2) were completing their student teaching requirement during their first year of participation in the study, 3) planned to graduate from GSU CEHD at the end of their first year of the study, and 4) planned to teach in an elementary or middle school in Georgia public schools following graduation. Recruitment efforts resulted in 14 CREATE residents and 48 comparison study participants who agreed to participate in the research in Cohort 3, 16 CREATE residents and 51 comparison participants who agreed to participate in Cohort 4, and 17 CREATE residents and 25 comparison participants who agreed to participate in Cohort 5. The

analytic samples differ across outcomes, depending on survey response rates and study attrition. Appendix C details participant participation/attrition from the study across the three years for each cohort.[4]

DATA SOURCES AND COLLECTION

This report uses multiple sources of data that include student achievement, teacher certification, personnel records and performance, and participant surveys (Figure 2). We also collected participant background information through the consent process and program data from the CREATE program team in the form of rosters, attendance logs, and mentor observation logs.

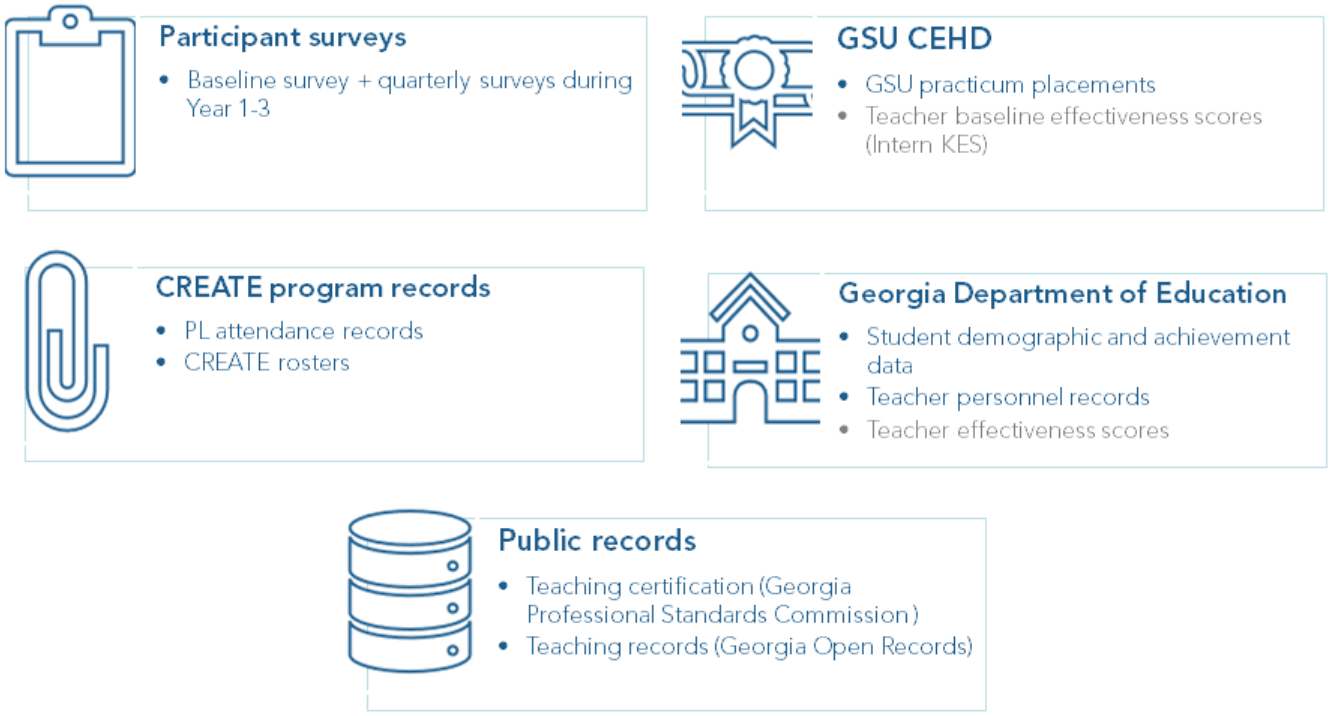

FIGURE 2. DATA SOURCES FOR CREATE RESEARCH STUDY

Note. Gray text indicates data that were not available for analysis for this study.

From participants who completed consent forms agreeing to take part in the study, researchers collected a one-time baseline survey, followed by quarterly online surveys during each of the three school years of study participation. The consent form and baseline survey provided background information about participants, as well as information about their interests in the context of teaching. The quarterly surveys included questions

[4] These values will vary to some extent from the numbers used to track attrition in the chapter on retention outcomes. For retention outcomes, we triangulated teaching status using multiple sources, including public records.

about the support participants were receiving at the student teaching sites and schools where they were employed, classroom experiences, mindfulness, resilience, perceptions of social capital, and plans for continued teaching at the time of completing the survey.

From GaDOE, we collected TAPS scores to assess teacher effectiveness, and Georgia Milestones scores to assess student achievement. Availability of data from GaDOE was limited due to interruptions caused by the COVID-19 pandemic. GaDOE did not calculate teacher TAPS scores during the 2019–2020 or 2020–21 school years. Additionally, this analysis uses only Georgia Milestones scores from the 2020–21 school year.[5]

From CREATE, we collected program data that includes PL attendance lists, resident rosters with information about classroom placement and assigned mentors, and mentor logs.

Finally, researchers also collected publicly available data about participants regarding their certification status and teaching status.

Appendix D includes detailed information about various forms of data collection and specific scales used in the study. Appendix E includes a schedule of major data collection milestones. Appendix F lists survey response rates for each survey completed by Cohort 3 through 5 participants.

GENERAL DESIGN

To address questions about the effects of CREATE on the main outcomes, we used a comparison group design to obtain estimates of interest.[6] That is, we compared outcomes for the CREATE group with those of a matched sample of similar comparison cases. We used three design and analysis strategies to establish equivalence of CREATE and comparison groups and to reduce potential for selection bias.

The first strategy was to select a comparison group that was similar to the CREATE group. Study participants in both the CREATE and comparison groups were from a pool of students enrolled in GSU CEHD. This ensured that the comparison group participants were similar to the CREATE residents, in terms of important characteristics, including motivation to enter the teaching profession in this region of Georgia and qualifications for entering the preservice teaching program at GSU. The comparison group participants chose not to join CREATE for a variety of reasons. For example, a comparison group participant may have been interested in joining CREATE but may have not wanted to teach in APS due to the distance from their home. Likewise, while they might have been committed to teaching long-term, they may not have wanted to make a three-year commitment to a specific program. Having a comparison group that was similar to the CREATE

[5] Our original analysis plan was to evaluate impacts on student achievement in a teacher's third year of teaching for all three cohorts. However, due to the suspension of Georgia Milestones testing in spring 2020 due to COVID-19, student achievement was measured for teachers three years after they started the program *on average*. Specifically, student achievement is measured in Year 4 for Cohort 3, Year 3 for Cohort 4, and Year 2 for Cohort 5.

[6] Note that the descriptive analysis of survey questions related to implementation (Chapter 3) is based on responses collected from *all* CREATE and *all* comparison teachers who completed the surveys. The sample was not limited to matched cases as described in the General Design section.

group on these factors was much more preferable than if we had selected a comparison group of study participants in colleges of education in other institutions or states.

The second strategy for ensuring a comparison group design with less potential for bias was to conduct additional matching within each cohort. This involved limiting the pool of study cases to achieve greater similarity between CREATE and comparison cases on baseline characteristics. The goal was to achieve a difference no larger than 0.25 standard deviations on any of the baseline covariates used to evaluate equivalence. The What Works Clearinghouse uses this criterion for assuming a tolerable level of bias (provided the analysis also adjusts for the covariate if baseline equivalence is greater than 0.05 standard deviations). In assessing impacts on teacher surveys and retention outcomes, we assessed baseline equivalence on measures of confidence in general teaching skills, motivation to enter teaching, and self-reported levels of math anxiety at the time participants completed the baseline survey. (For a select analysis of teacher retention, we also focused on baseline equivalence in student achievement and minority status in the classes of teachers.) For assessing impact on student math and ELA achievement, we used student pretest scores in the corresponding subject(s) from the year before they entered classes of study teachers.

Specific analyses used certain adaptations of the matching methods. For example, where samples were small, we used a very basic method of trimming the sample to allow overlap between CREATE and comparison cases on the baseline measure. For analyzing confirmatory impacts on student achievement, we matched students in terms of propensity scores (i.e., estimated probabilities of being in the CREATE group) informed by both student and teacher covariates.[7]

A third strategy for achieving greater accuracy in impact estimates was to adjust the results through analysis. Once we matched cases, we used fairly straight-forward, regression-based adjustment methods. That is, our estimates of differences in outcomes between the CREATE and comparison groups adjusted for any remaining differences between the groups on baseline characteristics that can affect outcomes. The success of the methods depends more on the quality and completeness of the covariates used to make the adjustment than the sophistication of the methods (Bloom et al., 2005). We report the regression models in appendices for corresponding chapters.

GENERAL APPROACH TO ANALYSIS

After matching cases separately by cohort for each main analysis, we analyzed impacts using a series of regression-based methods. Most often, we used standard one-level linear regressions. The two exceptions were 1) the analysis of impact on student achievement, for which we used a two-level regression model (students nested in teachers) and 2) the analysis of impact on teacher retention, for which, in addition to linear probability models, we used logistic regression to model the log odds of being retained. In each regression

[7] We will note the adaptations of the methods, as necessary, and report the results of the baseline equivalence tests alongside the impact results in the respective chapters.

model, we included a variable indicating membership in the CREATE group or the comparison group; a variable indicating membership in Cohort 4 or 5; a series of covariates; and terms for random effects.

For all confirmatory and certain exploratory results, we tested for baseline equivalence between the CREATE and comparison groups on specific covariates that, in theory, might be related to the outcome variable. This involved regressing each baseline covariate against a variable indicating membership in the CREATE or the comparison group, and variables indicating membership in Cohort 4 or Cohort 5. To determine the degree of equivalence, we examined the estimate of the regression-adjusted difference in the baseline covariate, reported in units of the pooled standard deviation of that covariate. We assessed baseline equivalence for a covariate in terms of difference between CREATE and comparison in their average score on that covariate expressed in units of the pooled standard deviation.

As part of the exploratory analyses, to test whether specific variables moderated impacts, we used the standard regression models and included a term for the interaction between the variable indicating membership in CREATE or comparison and the moderating characteristic of interest. For example, to examine whether impacts on retention increased across years for residents and teachers at the same point in their career trajectory, we used terms for interactions between membership in CREATE and cohort membership in a model with retention as the outcome.

Chapter 3. Implementation Results

In this chapter, we present findings related to fidelity of implementation across the three cohorts of study participants and provide descriptive findings from survey data about levels of support for teaching, reported success as teachers, mentorship, and participation in Together Time meetings.

✨ Research Questions 1 and 2 ✨

Were the key components of the CREATE teacher residency logic model implemented with fidelity?

What is the experience of study participants in the CREATE teacher residency program and in the comparison group, specifically with regard to level of support and mentorship?

FIDELITY OF IMPLEMENTATION RESULTS

We worked with the CREATE program team to establish key components for FOI based on CREATE's logic model. We collected data on these components and calculated whether fidelity was met for each of the key components. We have assessed implementation fidelity for the following key components: (1) progressive core classroom roles, (2) equity-centered professional learning, (3) compassion-based PL, (4) mentoring, and (5) Summer Resident Academy (SRA). Figure 1 shows the CREATE logic model. Thresholds used for FOI are in the FOI matrix in Appendix G.

We assessed FOI using CREATE program rosters and resident responses on surveys. CREATE rosters were the primary data source, and we used resident self-reported attendance to fill in any cases in which there were missing data in the CREATE rosters. The FOI assessment included active CREATE residents in a given year. See Appendix C for more information about reasons residents left the CREATE program.

We present results that show which indicators within the key program components were implemented with fidelity during the three years of CREATE programming for Cohort 3, Cohort 4, and Cohort 5 (combined) in Table 2. These results are detailed in Appendix H.

The CREATE team implemented three of the five key components of the CREATE residency program—progressive core classroom roles (Component 1); equity-centered PL (Component 2); and SRA (Component 5)—with fidelity for the years in which they were measured.

The CREATE team implemented compassion-based PL component (Component 3) with fidelity in Year 3 of programming for Cohorts 3 through 5, but not in Year 1 or Year 2. Specifically, the second indicator of this component, which measures residents' participation in compassion-based PL during Together Time meetings, did not meet fidelity during the 2019–2020 school year for Cohort 4 (then in Year 2) and Cohort 5 (then in Year

1). Through communications with CREATE staff, we learned a few residents were dissatisfied with CREATE programming (though they did not elaborate) and, as a result, did not attend multiple Together Time meetings. Also of note is that schooling and CREATE programming during the 2019–2020 school year underwent many unexpected changes due to the COVID-19 pandemic, including moving to online format. This may have impacted participation in these meetings.

The CREATE team did not implement the component multiple forms of mentoring (Component 4) with fidelity in either of the two years (years 2 and 3) in which they were measured. In Year 2, of the School Based Mentors (SBM) assigned to CREATE residents, 20% attended mentor training in the summer prior to the mentoring year (Indicator 1), and 0% attended mentor training during the school year (Indicator 2). During Year 3, 80% of SBMs assigned to Cohort 3 residents attended mentor training during the school year. For Cohort 4, 73% of SBMs assigned to residents attended mentor training in the summer prior to the school year, and none of the SBMs attended training during the school year. For Cohort 5, 86% of the SBMs assigned to residents attended mentor training during the summer prior to the mentoring year, and 71% of SBMs attended mentor training during the school year. The target for both Indicator 1 and Indicator 2 is to have 90% of mentors of CREATE residents attend mentor trainings prior to and during the mentoring year. Because CREATE cultivates a community of educators who share a commitment to collaborating to foster CREATE's mission, it has become common for these educators to serve as mentors with CREATE for multiple years. Adjusted training requirements for experienced mentors (that is, not requiring training each year) may explain these implementation scores that fall below the target thresholds.

Cohorts 3 through 5 met fidelity for the other two indicators that make up this component for both Year 2 and Year 3: attending semi-monthly mentor meetings with their mentors, including both SBMs and Instructional Mentors (IM) (Indicator 3), and participating in mentor-resident observation cycles with their mentors (Indicator 4). Also of note is that the sample size for this indicator is smaller than the total number of residents because some residents opted not to have a SBM during a given year.

TABLE 2. FIDELITY OF IMPLEMENTATION RESULTS FOR COHORTS 3, 4, AND 5 COMBINED

Component	Program level threshold	Year 1	Year 2	Year 3
Component 1: Progressive Core Classroom Roles	Year 1: 90% or more of residents must meet fidelity on 2+ indicators Year 2: 80% or more of residents must meet fidelity on 2+ indicators Year 3: 85% or more of residents must meet fidelity on 2+ indicators	49/50 (98%) met fidelity on 2+ indicators **Overall: Fidelity MET**	32/35 (91%) met fidelity on 2+ indicators **Overall: Fidelity MET**	34/35 (97%) met fidelity on 2+ indicators **Overall: Fidelity MET**

TABLE 2. FIDELITY OF IMPLEMENTATION RESULTS FOR COHORTS 3, 4, AND 5 COMBINED

Component	Program level threshold	Year 1	Year 2	Year 3
Component 2: Equity-Centered Professional Learning	Year 1: Fidelity met for Indicator 1 and at least one other indicator Year 2: Fidelity met for Indicator 1 and at least one other indicator Year 3: Fidelity w met for Indicator 1, Indicator 3, and at least one other indicator	Indicator 1: CREATE administrators host 2 or more institutes Indicator 2: 223/240 (93%) Indicator 3: Not measured in Y1 Indicator 4: 42/50 (84%) **Overall: Fidelity MET**	Indicator 1: CREATE administrators host 2 or more institutes Indicator 2: 182/192 (95%) Indicator 3: Not measured in Y2 Indicator 4: 30/35 (86%) **Overall: Fidelity MET**	Indicator 1: CREATE administrators host 2 or more institutes Indicator 2: 180/183 (98%) Indicator 3: 16/35 (46%) Indicator 4: 35/35 (100%) **Overall: Fidelity MET**
Component 3: Compassion-Based Professional Learning	Year 1: Fidelity met for two indicators Year 2: Fidelity met for two indicators Year 3: Fidelity met for two indicators	Indicator 1: CREATE administrators host 1 or more institutes Indicator 2: 42/50 (84%) **Overall: Fidelity NOT MET**	Indicator 1: CREATE administrators host 1 or more institutes Indicator 2: 30/35 (86%) **Overall: Fidelity NOT MET**	Indicator 1: CREATE administrators host 1 or more institutes Indicator 2: 35/35 (100%) **Overall: Fidelity MET**
Component 4: Multiple forms of mentoring	Years 2 and 3: All indicators must meet fidelity	Not measured in Year 1	Indicator 1: 27/35 (77%) Indicator 2: 25/35 (71%) Indicator 3: 35/35 (100%) receive a score of 2 and none receive a score of zero Indicator 4: 35/35 (100%) **Overall: Fidelity WAS NOT MET**	Indicator 1: 19/23 (83%) Indicator 2: 9/23 (39%) Indicator 3: 35/35 (100%) receive a score of 2 and none receive a score of zero Indicator 4: 33/35 (94%) **Overall: Fidelity WAS NOT MET**

TABLE 2. FIDELITY OF IMPLEMENTATION RESULTS FOR COHORTS 3, 4, AND 5 COMBINED

Component	Program level threshold	Year 1	Year 2	Year 3
Component 5: Summer Resident Academy	Year 2: Indicator 1 must meet fidelity	Not measured in Year 1	Indicator 1: 35/35 (100%) receive a score of 2 **Overall: Fidelity MET**	Not measured in Year 3

Note. All indicators that did meet fidelity thresholds are in green. All indicators that did NOT meet fidelity thresholds are in red.

Experiences of CREATE Residents and Comparison Group Early Career Teachers

In this section, we present descriptive findings from survey data across the three years of the study where participants responded to questions about their experiences with various CREATE PLs, their access to mentorship, how supported they felt at their schools, how successful they felt in a variety of professional areas, and other forms of support they received.

As a developing program, CREATE continues to evolve and adjust programming each year. It is important to assess whether or not these programmatic changes are producing the desired results. While the summary of survey results below does not answer this question, it seeks to provide a description of how CREATE has evolved throughout the experiences of Cohorts 3, 4, and 5. Each year of the CREATE teacher residency is unique from the other years in terms of expectations and content, so we find it most useful to look across cohorts, but within each year of the residency, to see how both resident experiences may or may not have evolved.

Progressive Core Classroom Roles

In their first year of teaching (and their second year of the residency), CREATE residents teach in classrooms in pairs. Co-teaching provides a scaffolded experience in which residents have support from a peer as they navigate their first year of teaching. This co-teaching arrangement is intended to prepare residents to serve as solo instructors in Year 3, their second year of teaching, with full responsibility of their classes.

In Year 2, surveys asked all participants whether they were sharing their classroom with a co-teacher, serving as the sole teacher of record in their classroom, or teaching in another arrangement. Across Cohorts 3 through 5, 93% ($n = 28$) of CREATE participants reported that they were sharing their classroom with a co-teacher, while 27% ($n = 14$) of comparison participants reported the same. Three percent ($n = 1$) of CREATE participants and sixty-one percent ($n = 31$) of comparison participants were serving as the sole teacher of record in their classroom. CREATE considered residents' needs and preferences in classroom arrangements, and honored the choices of residents around whether they desired a co-teaching arrangement, or preferred to teach a class on

their own. Those who had other classroom arrangements worked in special education or dual language immersion classes, served as paraprofessionals, worked with instructional coaches in their classrooms, or co-taught some, but not all, of their classes.

We asked the same question in Year 3 and found that 41% (n = 12) of CREATE participants and 31% (n = 14) of comparison participants shared a classroom with another teacher in their second year of teaching, while 53% (n = 16) of CREATE participants and 69% (n = 31) of comparison participants were the sole teacher of record. The remaining teachers worked with paraprofessionals or associate teachers. When looking at results for all years, it is important to consider that one of CREATE's partnership schools uses a co-teaching model throughout the school, regardless of teachers' experience levels.

Experiences With CREATE Equity-Centered PLs

Equity Centered Critical Friendship (ECCF) is the equity programming CREATE offers for residents as part of their Together Time meetings, and for veteran educators at separate PL institutes. ECCF practices are intended to guide educators in engaging in identity work, and to equip educators with the ability to identify and address issues of privilege and oppression in their schools. Throughout the residency experiences of Cohorts 3, 4, and 5, CREATE consistently met its intended implementation threshold for the number of equity-centered PLs held each year. While CREATE residents were able to register for and attend the ECCF Institutes that CREATE held for veteran educators, they also consistently received training in ECCF practices during Together Time meetings. We asked CREATE study participants about the degree to which they agreed with the three following statements regarding the ECCF protocols, with response options as follows: *Strongly disagree, Disagree, Neutral, Agree, Strongly agree*.

1. My teaching/instructional skills have improved as a result of Equity-Centered Critical Friendship protocols.
2. Equity-Centered Critical Friendship protocols have contributed to a greater sense of social support and belonging for me.
3. Equity-Centered Critical Friendship protocols are contributing to a more reflective practice on my part.

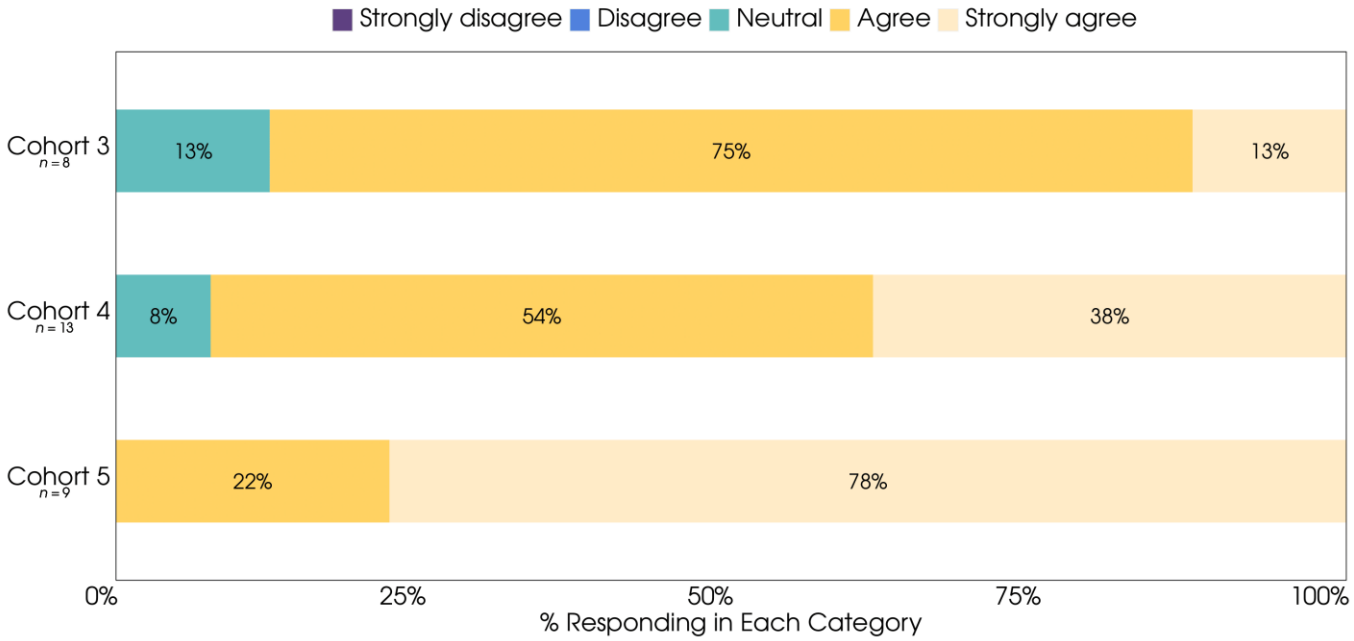

FIGURE 3. YEAR 3 CREATE RESIDENTS' RESPONSES TO "MY TEACHING/INSTRUCTIONAL SKILLS HAVE IMPROVED AS A RESULT OF EQUITY-CENTERED CRITICAL FRIENDSHIP PROTOCOLS"

Source: Quarterly surveys

Figure 3, Figure 4, and Figure 5 show the extent to which participants agreed with each of the above statements when they completed the first quarterly survey in Year 3, a point in the residency at which all residents would have received training around ECCF protocols through attendance at Together Time meetings, and potentially at ECCF Institutes. Of note is that none of the respondents *disagreed* (or *strongly disagreed*) with any of the three statements. These descriptive findings also reflect higher ratings given by Cohort 4 and Cohort 5 residents than by Cohort 3 residents.

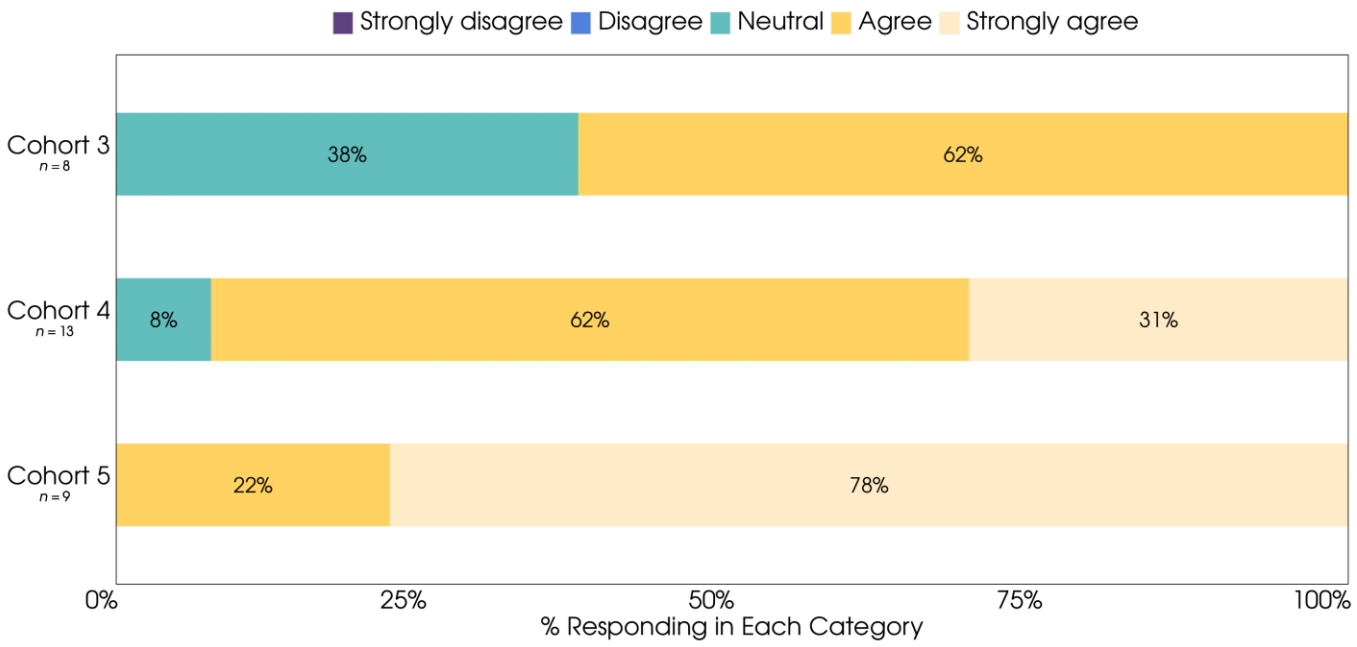

FIGURE 4. YEAR 3 CREATE RESIDENTS' RESPONSES TO "EQUITY-CENTERED CRITICAL FRIENDSHIP PROTOCOLS HAVE CONTRIBUTED TO A GREATER SENSE OF SOCIAL SUPPORT AND BELONGING FOR ME"

Source: Quarterly surveys

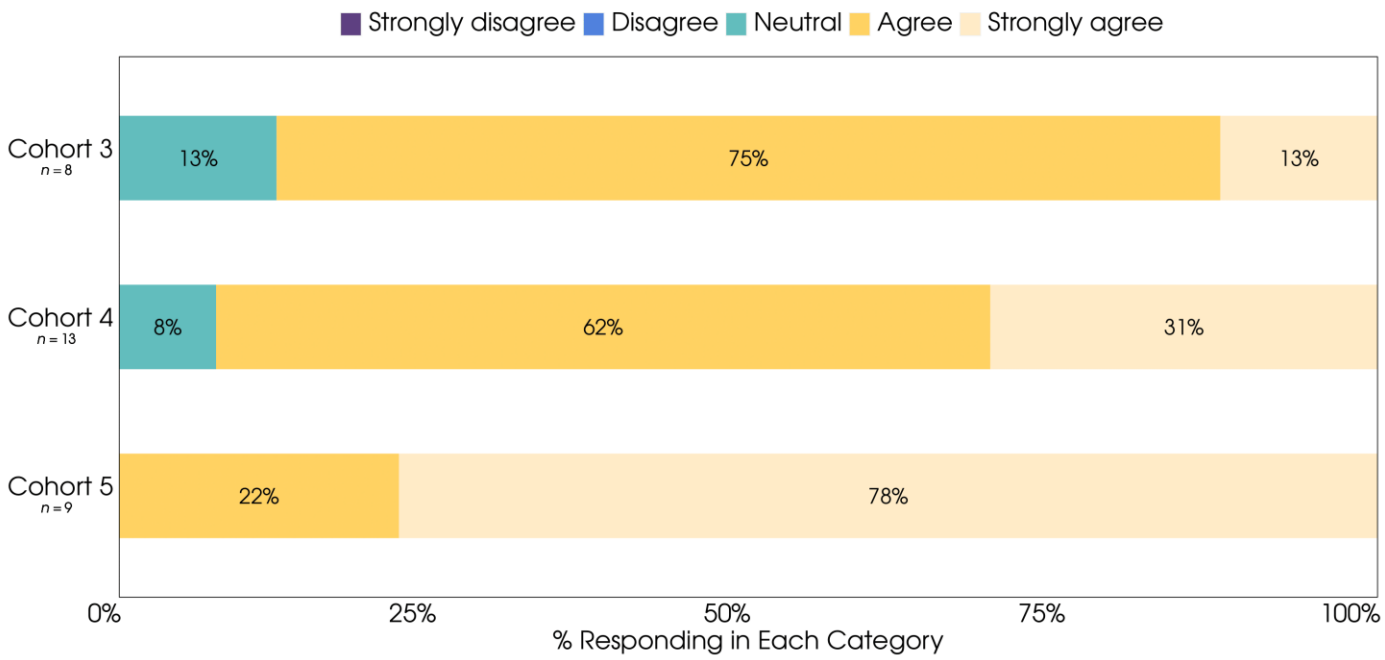

FIGURE 5. YEAR 3 CREATE RESIDENTS' RESPONSES TO "EQUITY-CENTERED CRITICAL FRIENDSHIP PROTOCOLS ARE CONTRIBUTING TO A MORE REFLECTIVE PRACTICE ON MY PART"

Source: Quarterly surveys

Through ECCF, the centering of equity-based approaches and practices in all CREATE PLs, and social justice-steeped supports, CREATE actively invests in supporting educators' understanding and commitment to antiracist, equity-centered practices in schools. To assess the degree to which participants felt capable of addressing issues of race and equity in their schools and the opportunities available for them to do so, we added survey questions about this to surveys during the 2021–22 school year. Both CREATE and comparison participants in Cohort 5 answered questions about the extent to which they were integrating and discussing issues related to race and equity at their schools. We asked all participants the following question: How confident are you in your understanding of concepts related to race and equity in schools? All CREATE respondents reported being at least *moderately confident* in their understanding of race- and equity-related concepts in schools, with 89% ($n = 8$) of CREATE teachers feeling *very* or *extremely confident* in their understanding, while 67% ($n = 6$) of comparison teachers felt *very* or *extremely confident* (Figure 6).

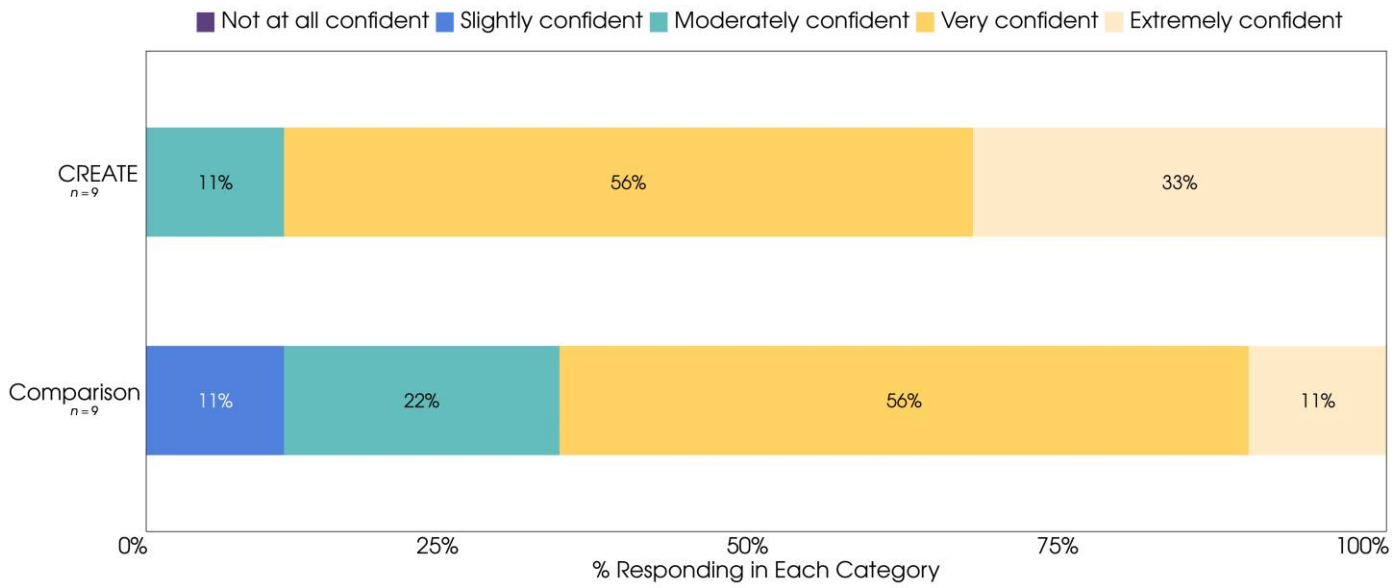

FIGURE 6. YEAR 3 CREATE RESIDENTS' RESPONSES TO "HOW CONFIDENT ARE YOU IN YOUR UNDERSTANDING OF CONCEPTS RELATED TO RACE AND EQUITY IN SCHOOLS?" (COHORT 5 ONLY)

Source: Quarterly surveys

Participants were also asked the following question: At your school, how often are you encouraged to think deeply about race-and equity-related issues? Of the Cohort 5 participants who responded, 78% ($n = 7$) of CREATE participants reported *frequently* or *almost always* being encouraged to think deeply about race- and equity-related issues, while 44% ($n = 4$) of comparison participants reported doing so *frequently* or *almost always* (Figure 7). These descriptive findings shed some light onto how equipped CREATE participants feel in understanding issues of race and equity in schools as compared to their comparison counterparts, as well as the degree to which CREATE and comparison schools integrate an understanding about race and equity in their work.

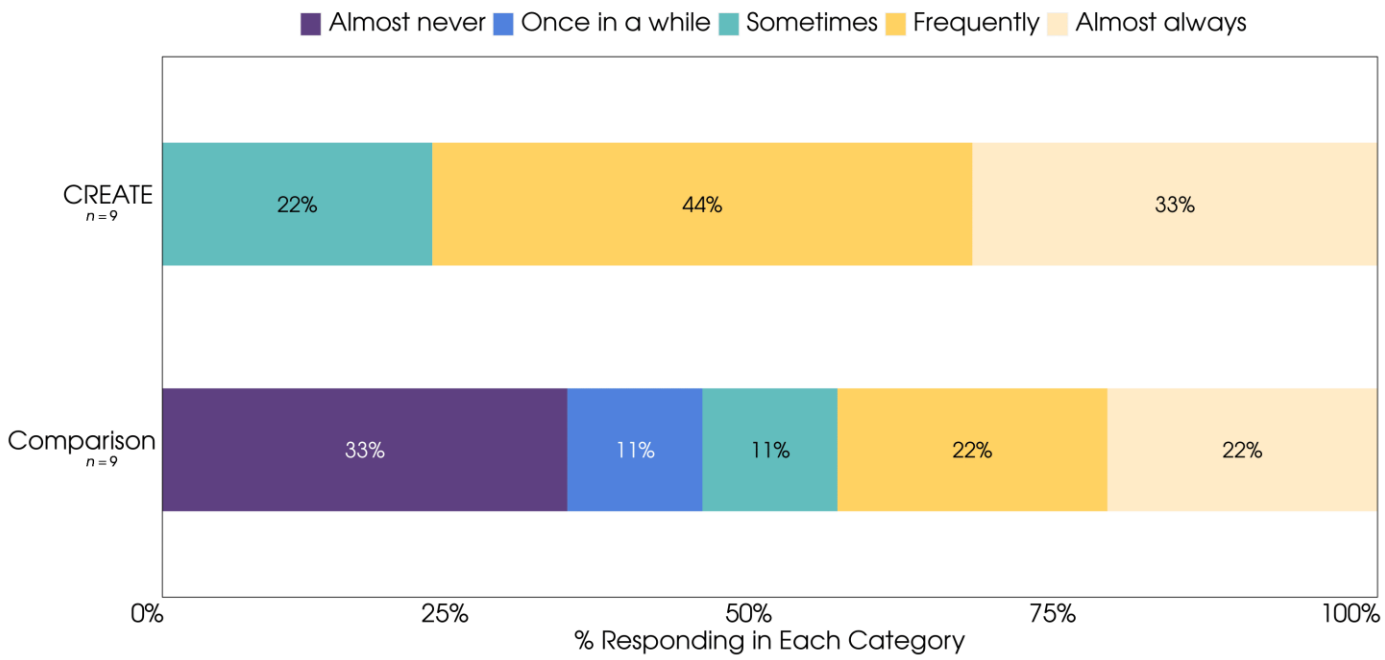

FIGURE 7. YEAR 3 CREATE RESIDENTS' RESPONSES TO "AT YOUR SCHOOL, HOW OFTEN ARE YOU ENCOURAGED TO THINK DEEPLY ABOUT RACE- AND EQUITY-RELATED ISSUES?" (COHORT 5 ONLY)

Source: Quarterly surveys

Experiences with CREATE Compassion-Based PLs

Compassion-based practices are integrated into the Together Time meetings that residents attend, and are central to the Power, Presence, Impact (PPI) institutes offered by CREATE to experienced educators, which are also open for residents to attend for additional PL. PPI protocols learned during Together Time meetings include meditative practices using the Cognitively-Based Compassion Training (CBCT©) framework and methods for reducing the stressors of teaching. Through CREATE's compassion-based PLs, participants use compassion-based mindfulness to cultivate non-judgmental awareness and acceptance of their thoughts and emotions. The goal of PPI is to support educators in using mindfulness and contemplative practices to increase emotional intelligence and self-compassion in the face of teaching-related stressors.

On the second quarterly survey, residents who reported having had received PPI training through attendance at Together Time meetings and/or PPI institutes rated the following statement: To what extent have Presence, Power, Impact practices improved your capacity to be an effective teacher? About half of the participants in Cohort 3 and Cohort 4 reported that PPI practices have *somewhat* improved their capacity to be effective teachers, while 75% of Cohort 5 participants reported that PPI practices have *very much* improved their capacity to be effective teachers (Figure 8).

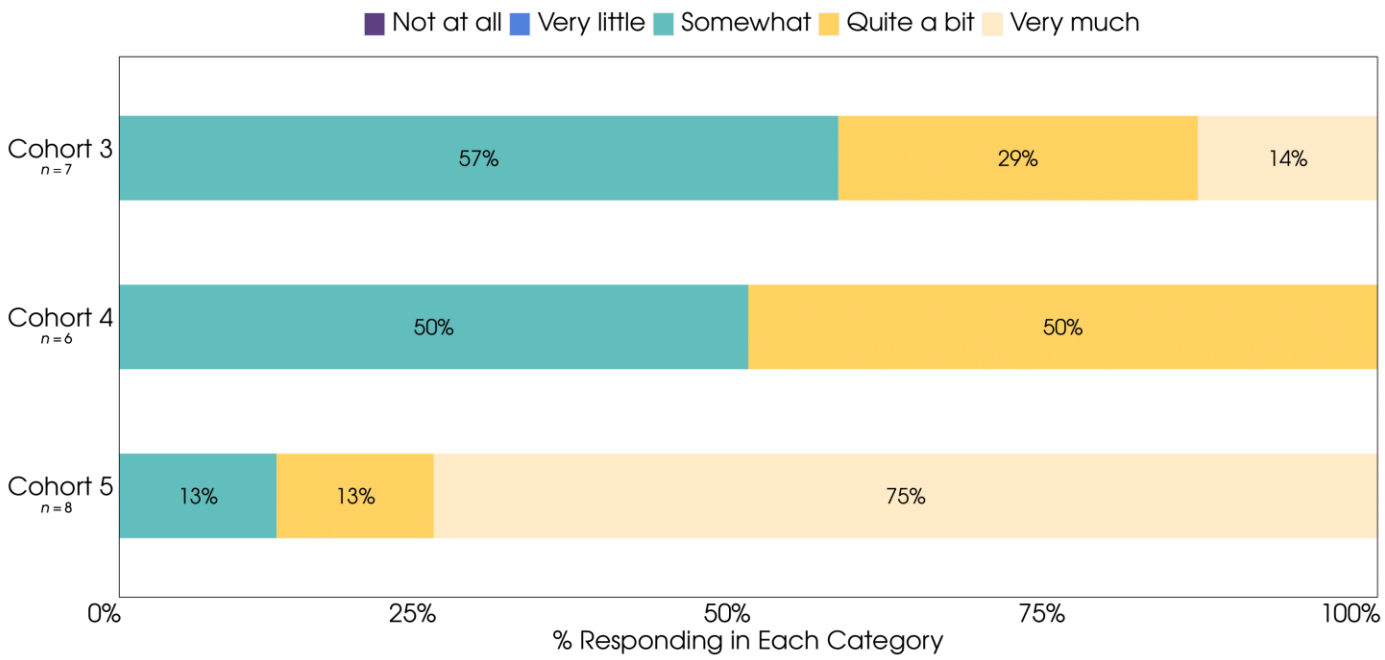

FIGURE 8. YEAR 3 CREATE RESIDENTS' RESPONSES TO "TO WHAT EXTENT HAVE PRESENCE, POWER, IMPACT PRACTICES IMPROVED YOUR CAPACITY TO BE AN EFFECTIVE TEACHER?"

Source: Quarterly surveys

We also asked residents who received any PPI training to respond to the following statement: To what extent have Presence, Power, Impact practices improved your ability to manage the day-to-day challenges of working in the classroom? Similar to the descriptive trends reported above, Cohort 5 participants assigned PPI higher ratings than Cohort 4 and Cohort 3 residents did. While 63% of Cohort 5 participants reported that PPI practices have *very much* improved their ability to manage the day-to-day challenges of working in the classroom, the majority of Cohort 3 and Cohort 4 participants reported that PPI practices *somewhat* improved their ability to manage these challenges (Figure 9).

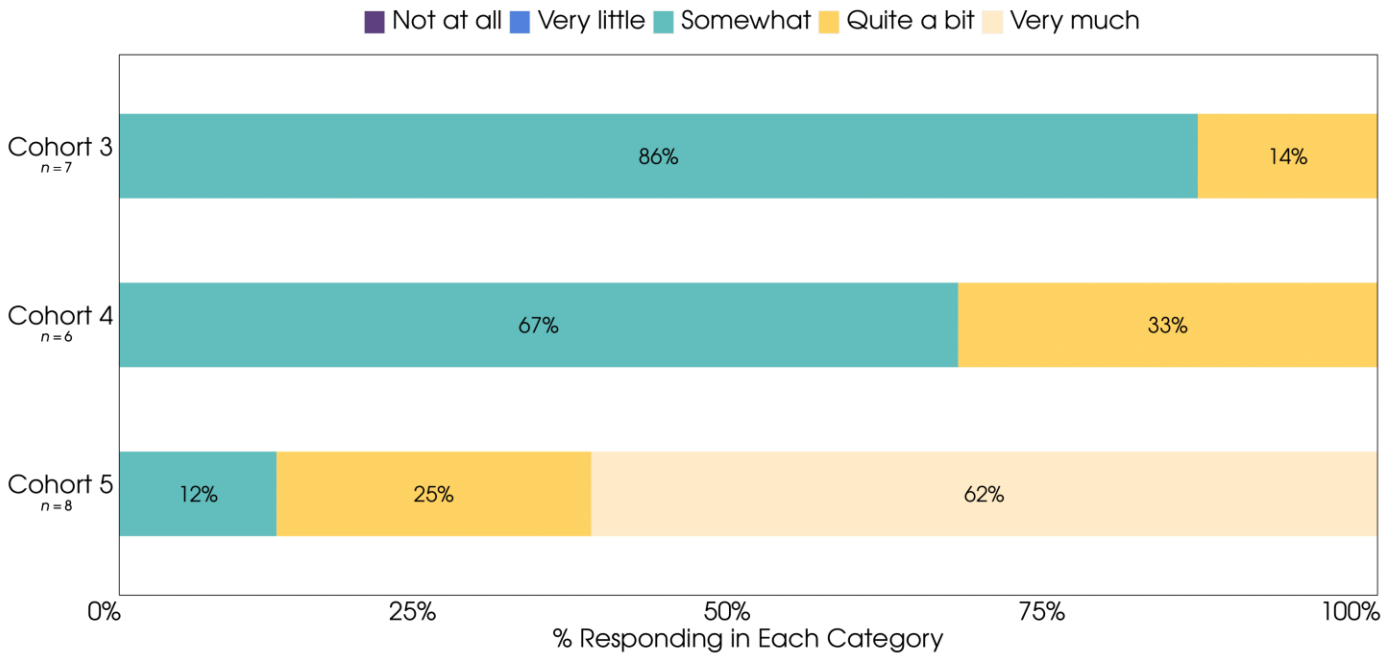

FIGURE 9. YEAR 3 CREATE RESIDENTS' RESPONSES TO "TO WHAT EXTENT HAVE PRESENCE, POWER, IMPACT PRACTICES IMPROVED YOUR ABILITY TO MANAGE THE DAY-TO-DAY CHALLENGES OF WORKING IN THE CLASSROOM?"

Source: Quarterly surveys

Summer Resident Academy (SRA)

SRA is a paid summer academy facilitated by university faculty, CREATE leaders, and district representatives that prepare residents for the upcoming school year. SRA 2 occurs in the summer after the first year of residency, prior to their first year of teaching. During the five-week SRA 2, residents gain experience teaching in summer schools while learning about equity, privilege, and justice in schools. This experience allows residents to practice applying their learnings to their experiences in the classroom, and to learn how to manage their emotions and stress around teaching in a supported environment. SRA 3, which was offered for the first time during the summer of 2020 for Cohort 4, occurs for two weeks in the summer after their first year of co-teaching, prior to their third and final year in the residency. SRA 3 focuses on reflection on the previous year of teaching and planning for the next year.

FOI calculations show that residents consistently meet the expected threshold for SRA 2 attendance by attending all or most days of the 5-week SRA. On the first survey during Year 2, we ask CREATE residents participating in the study to answer the following question: How valuable was the Summer Resident Academy to your practice as a classroom teacher?

The majority of CREATE residents participating in the study reported that the intensive summer training they participated in during the summer prior to their first year of teaching was *more than moderately valuable* or *very valuable* (Figure 10). We note that the proportion of participants who reported SRA as *very valuable* increases by cohort, with 70% ($n = 7$) of Cohort 5 participants reporting at this highest level.

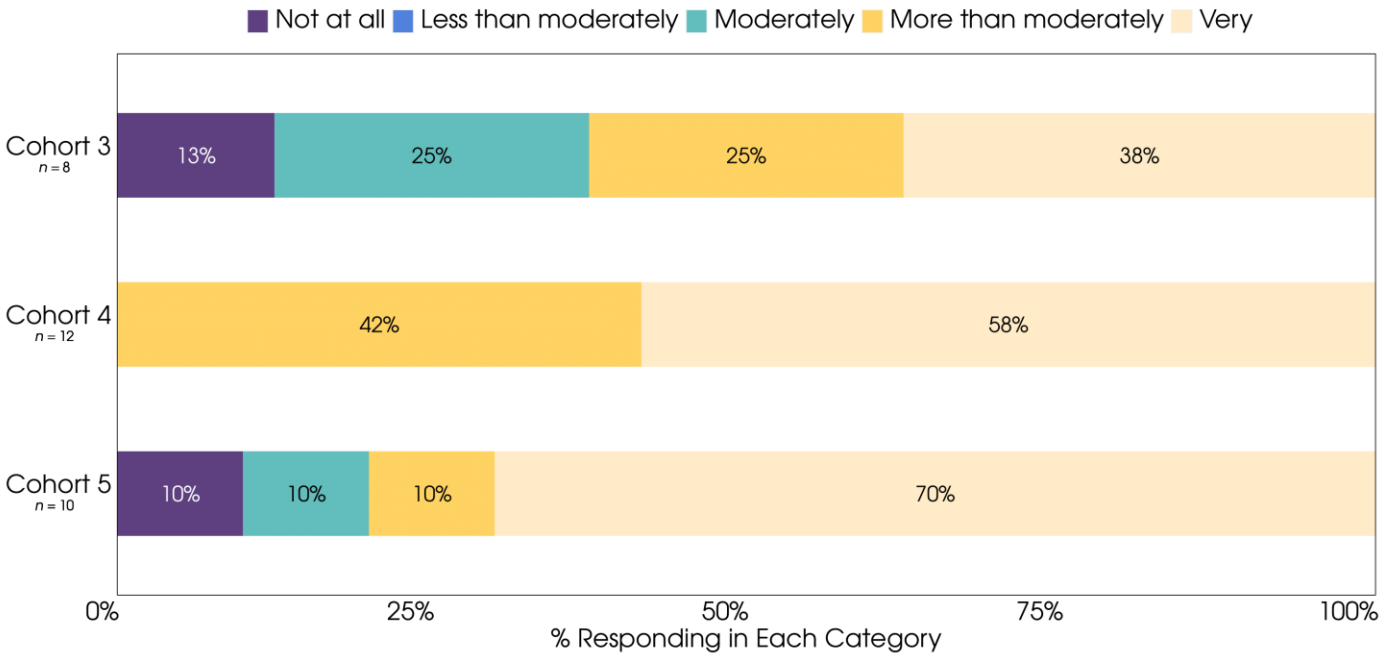

FIGURE 10. CREATE RESIDENTS' RESPONSE TO VALUE OF SUMMER RESIDENT ACADEMY 2 TO PRACTICE AS A CLASSROOM TEACHER

Source: Quarterly surveys

In a follow-up, open-ended, question, we asked participants to describe what they did and did not find valuable about SRA with respect to their practice as a classroom teacher. In response to this question, residents frequently mentioned that benefits of participating in SRA included 1) having the opportunity to plan and collaborate with a co-teacher and with fellow CREATE residents who were also new to teaching, 2) learning strategies for organizing their classroom community sharing resources and strategies, and 3) preparing for teaching virtually. Challenges surrounding the SRA mentioned by participants included facing unexpected classroom visits by school administrators, and having less control over the classroom than was desired due to having other teachers intervening during class time. Participants also noted that they would have liked to see more resident feedback incorporated into the SRA.

Though not included in the FOI matrix, quarterly surveys also asked about residents' experiences at SRA 3. Again, the majority of residents rated this SRA as *more than moderately valuable* or *very valuable* (Figure 11), with

100% (n = 11) of Cohort 5 reporting that SRA 3 was *very valuable*. Residents specified that the value of SRA 3 included the sharing of resources and strategies that helped prepare them to teach in virtual settings, opportunities to organize materials before the school year began, and learning to de-escalate behavioral issues in order to foster supportive classroom environments.

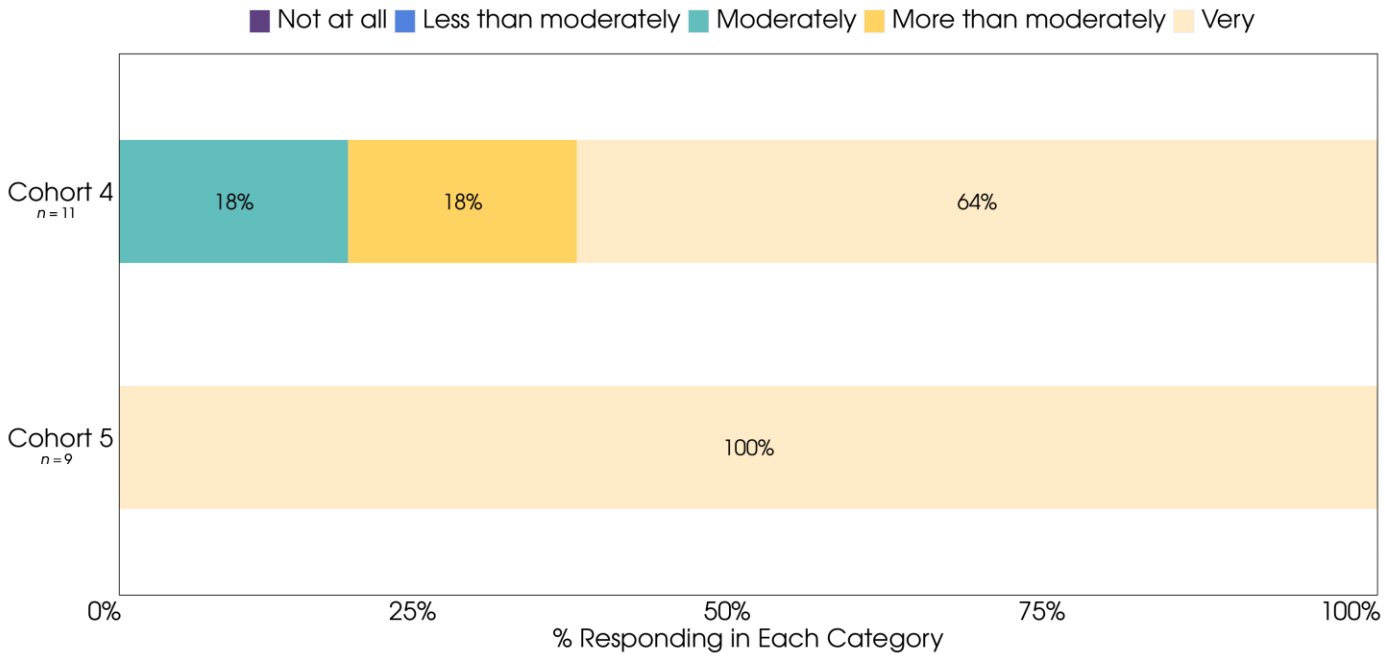

FIGURE 11. CREATE RESIDENTS' RESPONSE TO VALUE OF SUMMER RESIDENT ACADEMY 3 TO PRACTICE AS A CLASSROOM TEACHER

Source: Quarterly surveys

Mentorship

On each of the surveys participants take, we ask both CREATE and comparison participants how helpful various supports have been to their practice as a teacher. For CREATE participants, these supports include CREATE mentors, other CREATE residents, CREATE administrators, PPI practices, ECCF protocols, school administrators, and interactions with veteran educators. Overall, Cohort 4 and Cohort 5 residents reported higher ratings of helpfulness across various CREATE supports than Cohort 3 residents did. CREATE Instructional Mentors (IMs) were the highest rated support for CREATE residents in terms of helpfulness. IMs, with training from CREATE, meet with Year 2 and Year 3 residents weekly to support them in their command of content, and in the integration of the equity-centered instructional practices and compassion-based mindfulness skills they have learned. In addition to an IM, each Year 2 and Year 3 CREATE resident is assigned

a School-based Mentor (SBM) who provides mentorship at their school site and supports residents' development of a sense of belonging at their school.

To paint a picture of the experiences that comparison participants have with mentors, we asked about access to mentorship on surveys and found that 80% ($n = 41$) of comparison teachers across the three cohorts had mentors during their first year of teaching, while 39% ($n = 17$) of comparison participants had mentors during their second year of teaching. Out of mentor teachers, school administrators, and other teachers in their schools, comparison participants rated the other teachers in their schools as the most helpful form of support.

Other Forms of Support

Each survey asks how supported both CREATE and comparison participants feel at their student-teaching placement or school site. We found that a higher percentage of CREATE participants reported feeling *more than moderately* or *very* supported at their student-teaching placement and school site (Figure 12 through Figure 14 for Cohort 3, 4, and 5).

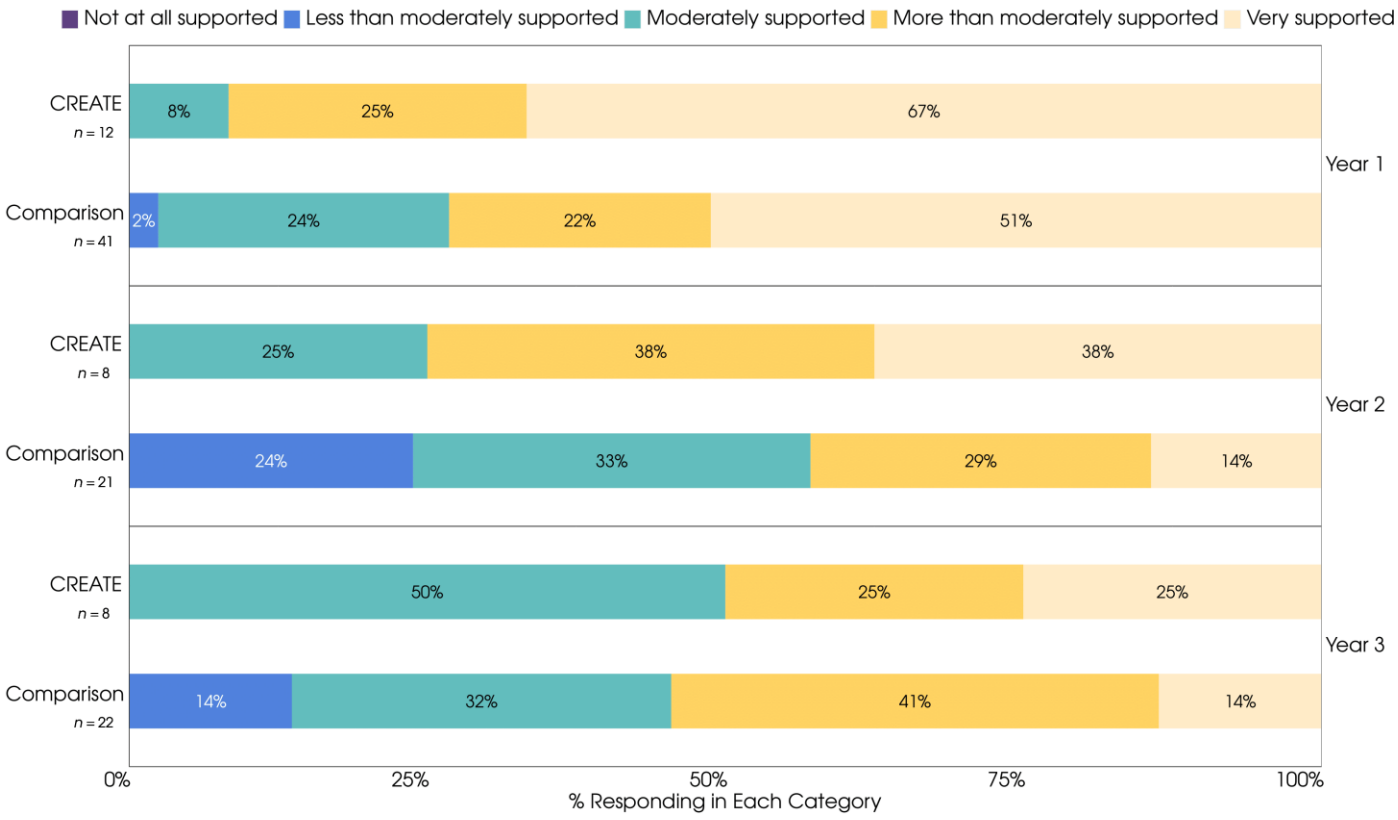

FIGURE 12. LEVEL OF SUPPORT AT PRACTICUM PLACEMENT OR SCHOOL SITE FOR COHORT 3 IN YEARS 1, 2, AND 3

Source: Quarterly surveys

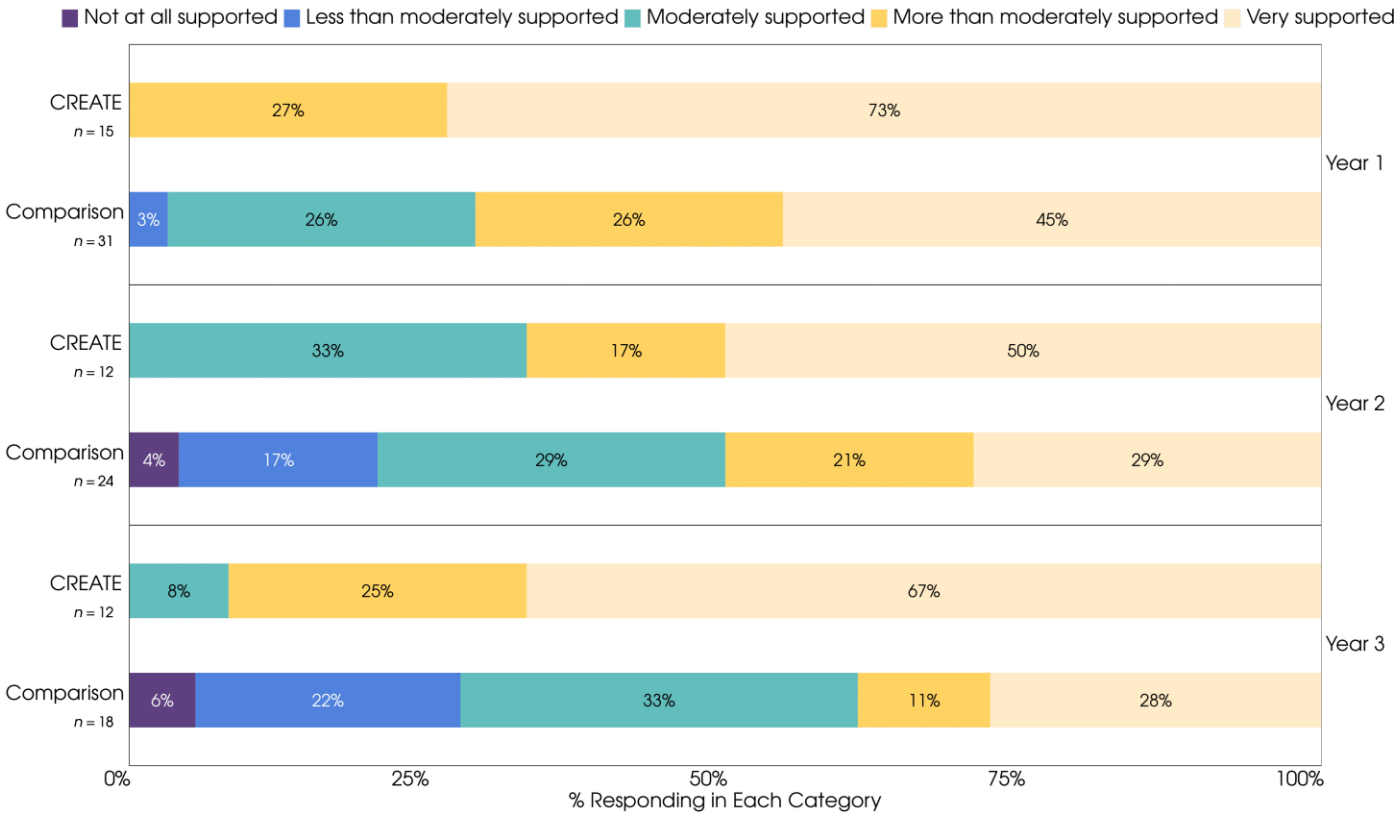

FIGURE 13. LEVEL OF SUPPORT AT PRACTICUM PLACEMENT OR SCHOOL SITE FOR COHORT 4 IN YEARS 1, 2, AND 3

Source: Quarterly surveys

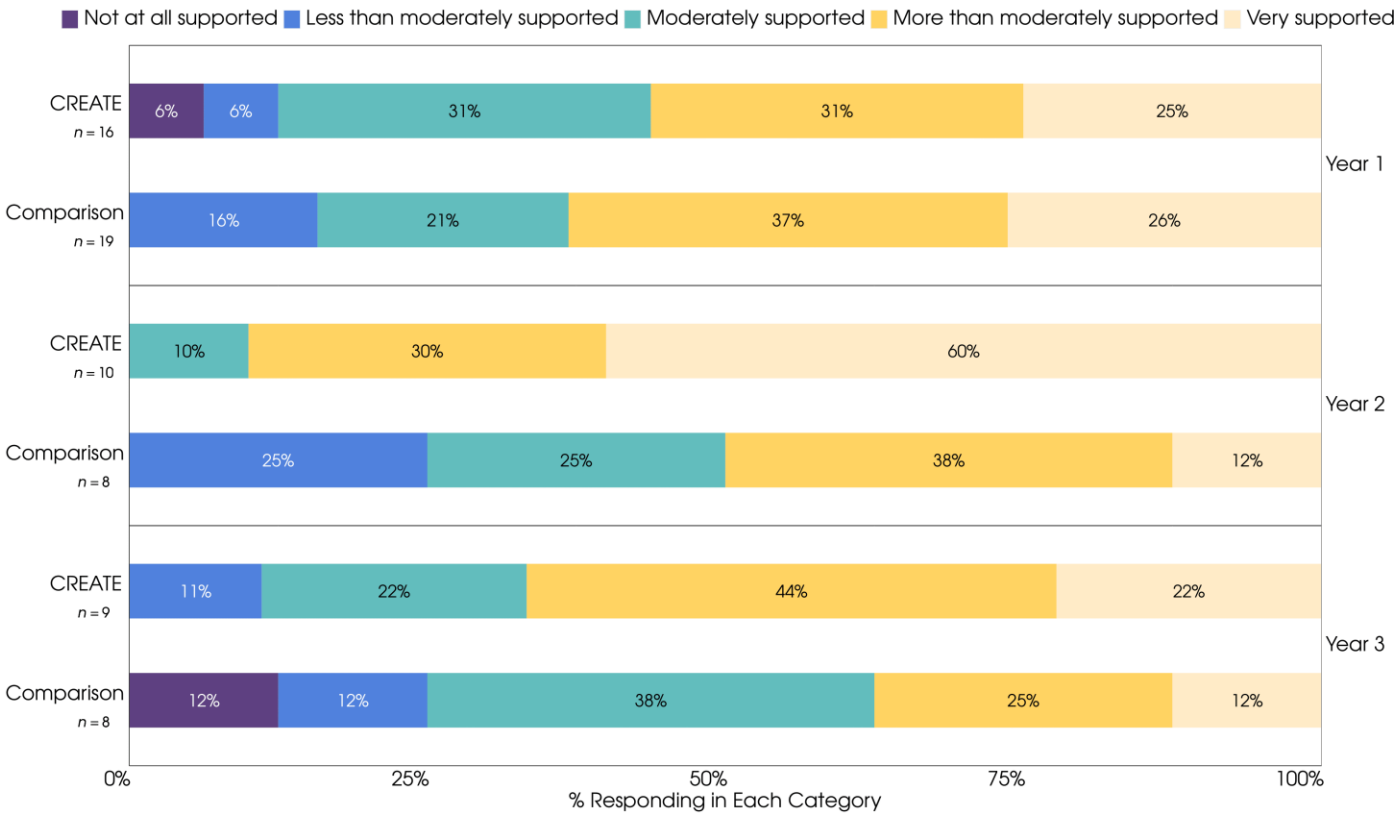

FIGURE 14. LEVEL OF SUPPORT AT PRACTICUM PLACEMENT OR SCHOOL SITE FOR COHORT 5 IN YEARS 1, 2, AND 3

Source: Quarterly surveys

Additionally, participants were asked about how successful they felt in the following areas: pedagogical knowledge, content knowledge, classroom management, lesson planning, and balancing their school responsibilities (in Year 1), general work life, and personal life (Figure 15 through Figure 17 for Cohort 3, 4, and 5). Overall, across the three cohorts in Year 3, CREATE participants felt more successful in classroom management and balancing the multiple responsibilities than comparison participants did. Comparison participants felt slightly more successful in their grasp of content knowledge than CREATE participants did. Of the three cohorts of CREATE residents, Cohort 4 generally felt the least successful in these areas than Cohort 3 and Cohort 5 residents did.

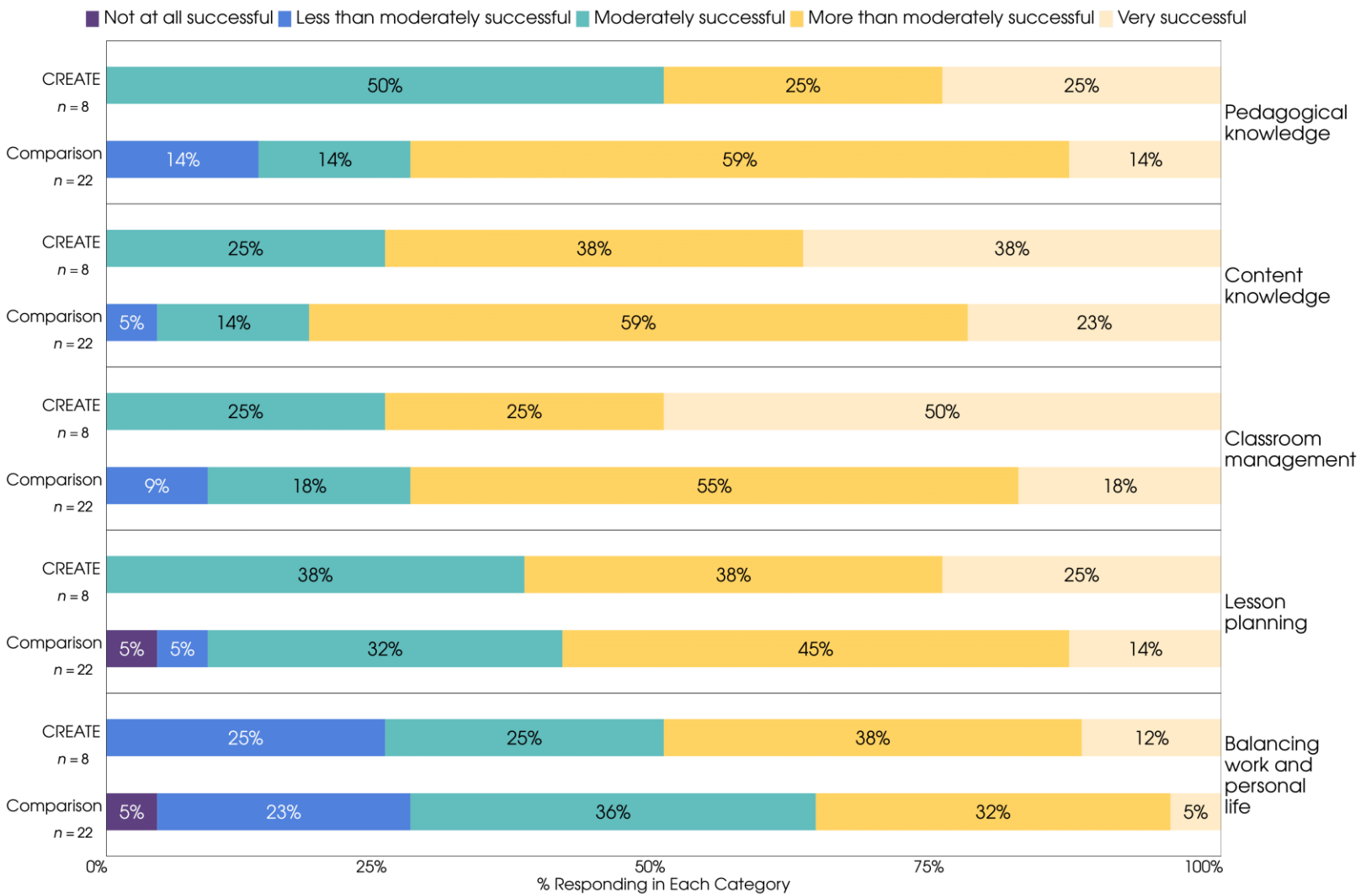

FIGURE 15. LEVEL OF SUCCESS IN TEACHING-RELATED AREAS FOR COHORT 3 IN YEARS 1, 2, AND 3

Source: Quarterly surveys

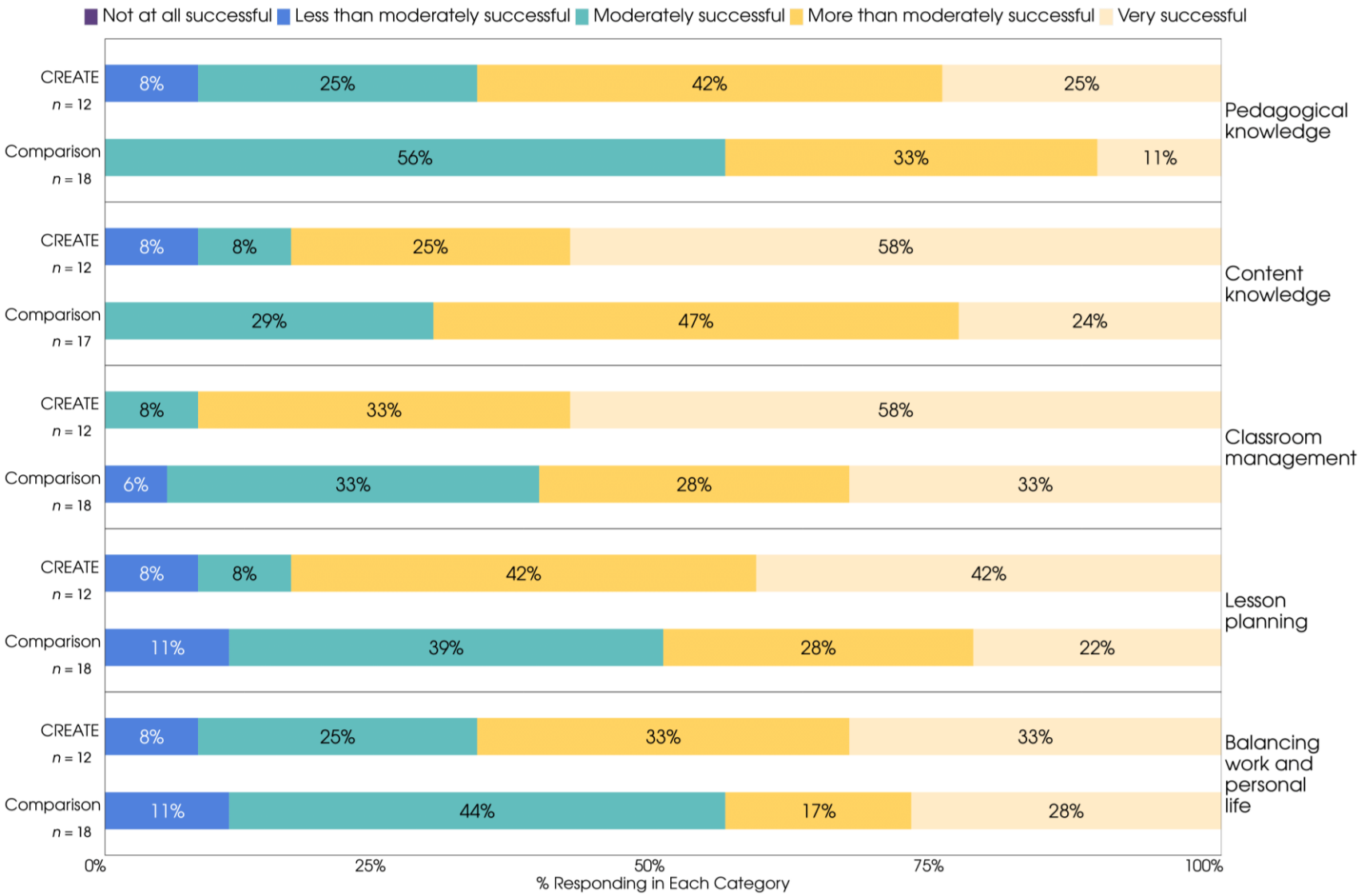

FIGURE 16. LEVEL OF SUCCESS IN TEACHING-RELATED AREAS FOR COHORT 4 IN YEARS 1, 2, AND 3

Source: Quarterly surveys

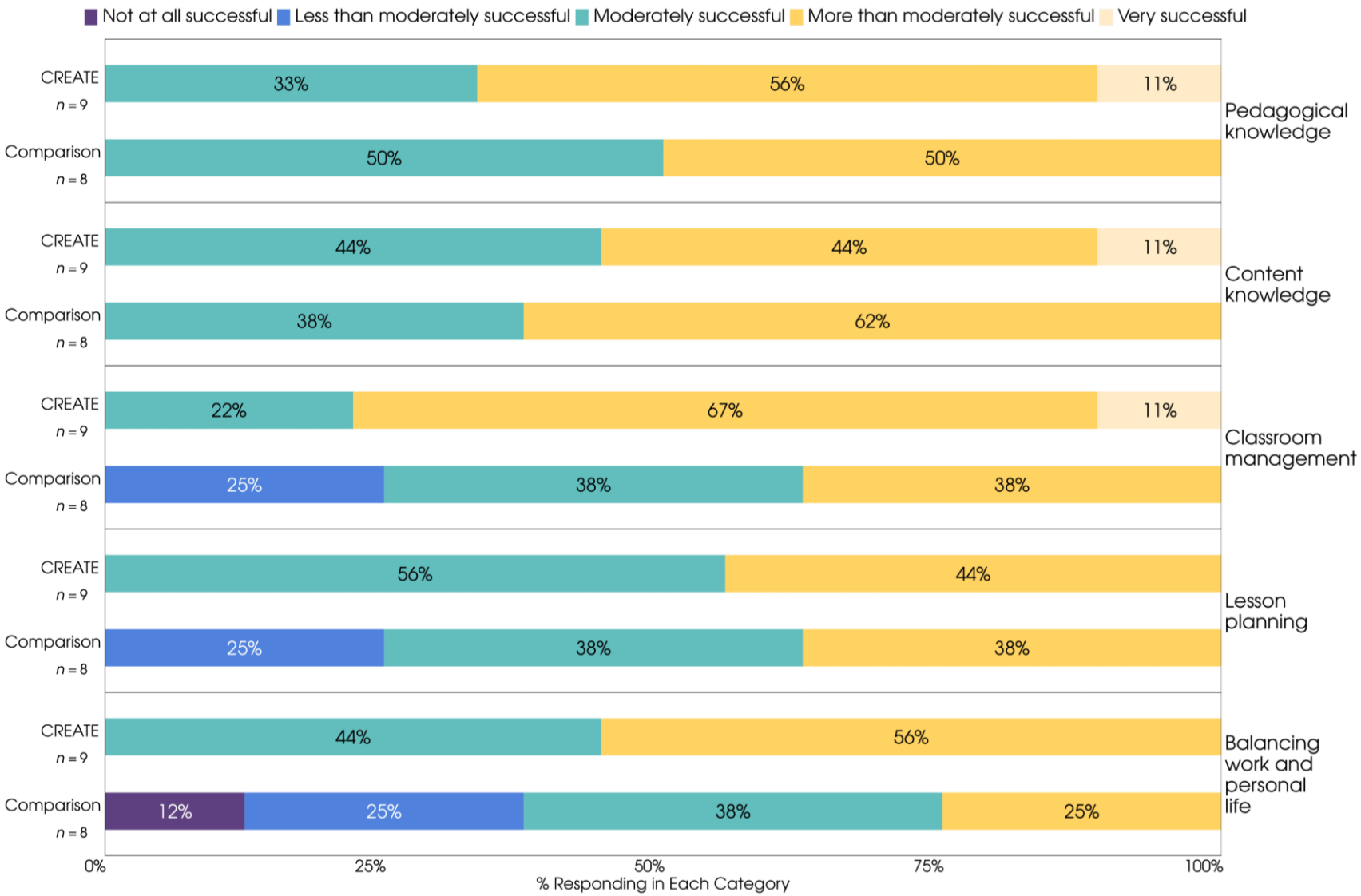

FIGURE 17. LEVEL OF SUCCESS IN TEACHING-RELATED AREAS FOR COHORT 5 IN YEARS 1, 2, AND 3

Source: Quarterly surveys

We also asked participants about the value of receiving financial support from CREATE. Of the Cohort 3, 4, and 5 residents who responded to this item in Year 1, during which they receive a $5,000 stipend while student teaching, 70% (n = 31) reported that they would have experienced significant financial hardship if they had not received financial assistance from CREATE. This finding reflects research that reports that the cost of becoming a teacher is the number one barrier to entering the teaching profession, particularly for Black teacher candidates. Teacher residency programs like CREATE restructure the financial responsibility related to the cost of becoming an educator, relieving teacher candidates of some of the financial burden and fostering financial sustainability in the teaching profession. As a result, a more diverse teacher workforce is generated (NCTR, 2023).

Chapter 4. Exploratory Impacts on Teachers' Self-Compassion, Levels of Burnout and Stress Management and Empathy as Related to Teaching

✨ Research Questions 5 and 6 ✨

Is there a positive impact of CREATE on self-reported levels of self-compassion, burnout, and stress management and empathy related to teaching, three years after entry into the residency program?

Do these impacts increase with each consecutive cohort?

IMPACTS AFTER TWO YEARS AS TEACHERS OF RECORD

By the end of their second year of teaching, CREATE teachers reported more favorable levels of self-compassion, teacher burnout, and stress management and empathy related to teaching, than teachers in the comparison group. These outcomes are theoretically posited to mediate longer-term outcomes of teacher effectiveness, retention, and student achievement.

Specifically, among the 27 CREATE and 32 comparison matched cases,[8] CREATE teachers reported higher levels of self-compassion (Standardized Effect Size [ES] = .51, p = .06), and higher levels of stress management and empathy related to teaching (ES = .34, p = .25). On dimensions of teacher burnout, CREATE teachers reported lower emotional exhaustion (ES = -.18, p = .52), lower levels of depersonalization (ES = -.27, p = .34), and higher levels of personal accomplishment (ES = .29, p = .31)[9].

The large effect sizes across outcomes that are all in the hypothesized direction of positive impacts are noteworthy; however, none of the results reached statistical significance (at significance level α = .05) (Kraft, 2020).

EXPLORATION OF LONGITUDINAL TRENDS

The impacts described above are assessed towards the end of implementation to reflect maximum exposure. However, they do not reveal the variability in responses across cohorts and over time, which provides additional context for results, especially for the period of program adaptations necessitated by COVID-19. Year-end trends in mean scores by cohort for each of the five scales are displayed in Figure 18 through

[8] Cases were limited to samples of teachers who were matched within cohort in terms of baseline covariates, and with non-missing covariate and outcomes data

[9] Baseline equivalence between the CREATE and comparison groups was achieved on the three covariates used in analysis, with standardized mean differences of .017 SD on a measure of confidence in general teaching skills, -.176 SD on a scale assessing motivation for entering the teaching profession, and .008 SD on a measure of math anxiety. See Appendix I for a description of the measures used.

Figure 22. Descriptively, we observe variation in responses across cohorts and over time, with the largest differences between CREATE and comparison observed for Cohort 5. The results should be interpreted keeping in mind the disruptions of COVID-19 and its impacts on teacher well-being. To test the hypothesis that impacts increase from one cohort to the next, potentially reflecting program development and intensification, we assessed whether 3-year impacts (that is, after the second year of teaching) varied across cohorts. We did not observe a difference in impact that reached conventional levels of statistical significance (α = .05) across cohorts after three years.

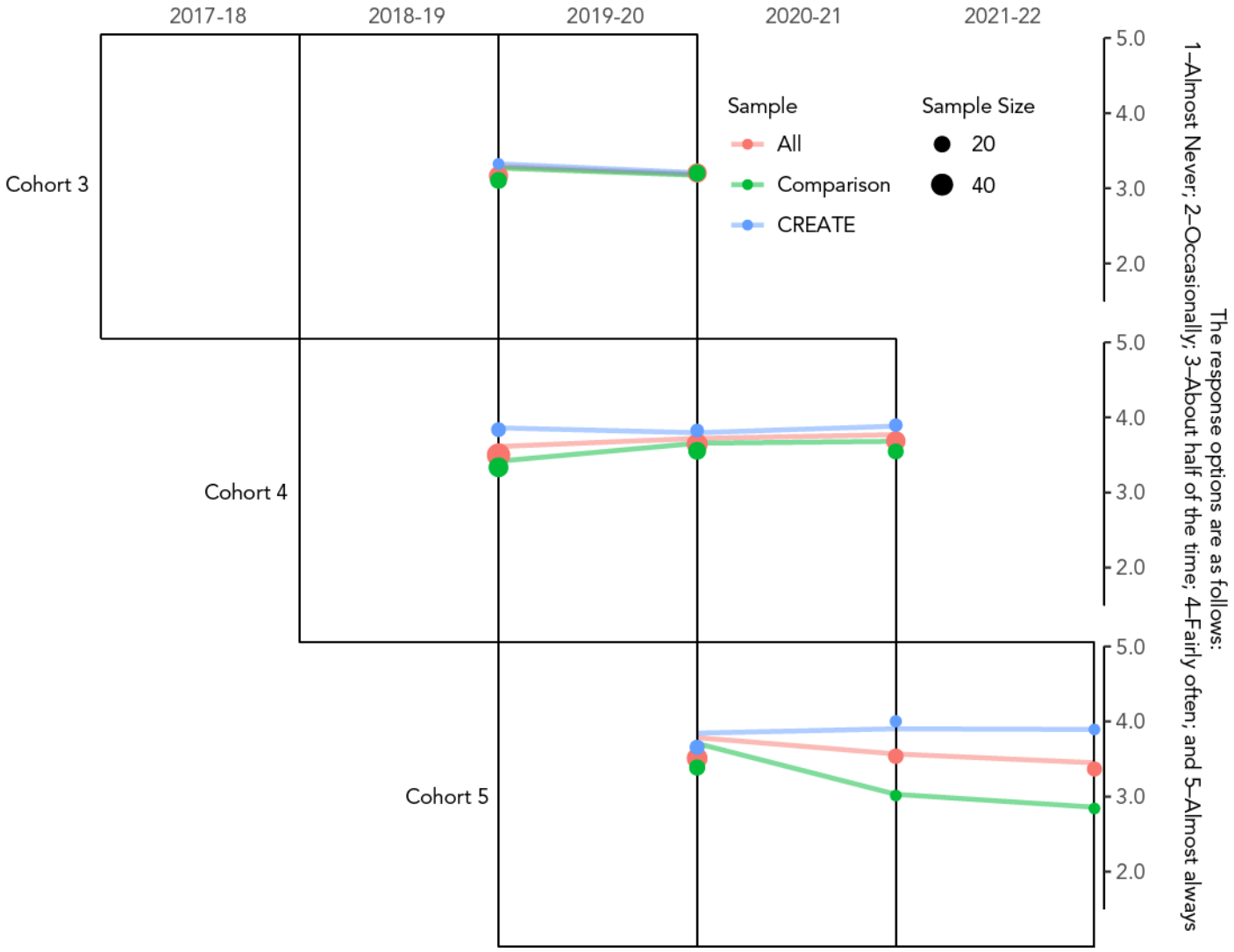

FIGURE 18. CHANGES IN COHORT 3–5 CREATE AND COMPARISON PARTICIPANTS RESPONSES OF OVER TIME ON THE SELF-COMPASSION SCALE

Note. The 26-item self-compassion scale reflects frequency of feelings of self-compassion. Higher scores correspond to greater self-compassion. The lines connect score averages for the consistent sample of 59 matched cases (27 in CREATE and 32 in comparison); the dots represent outcomes for all available teachers responding in a given cohort and year. This scale was added during the 2018–19 school year, so we do not have data for Cohort 3 during their student teaching year.

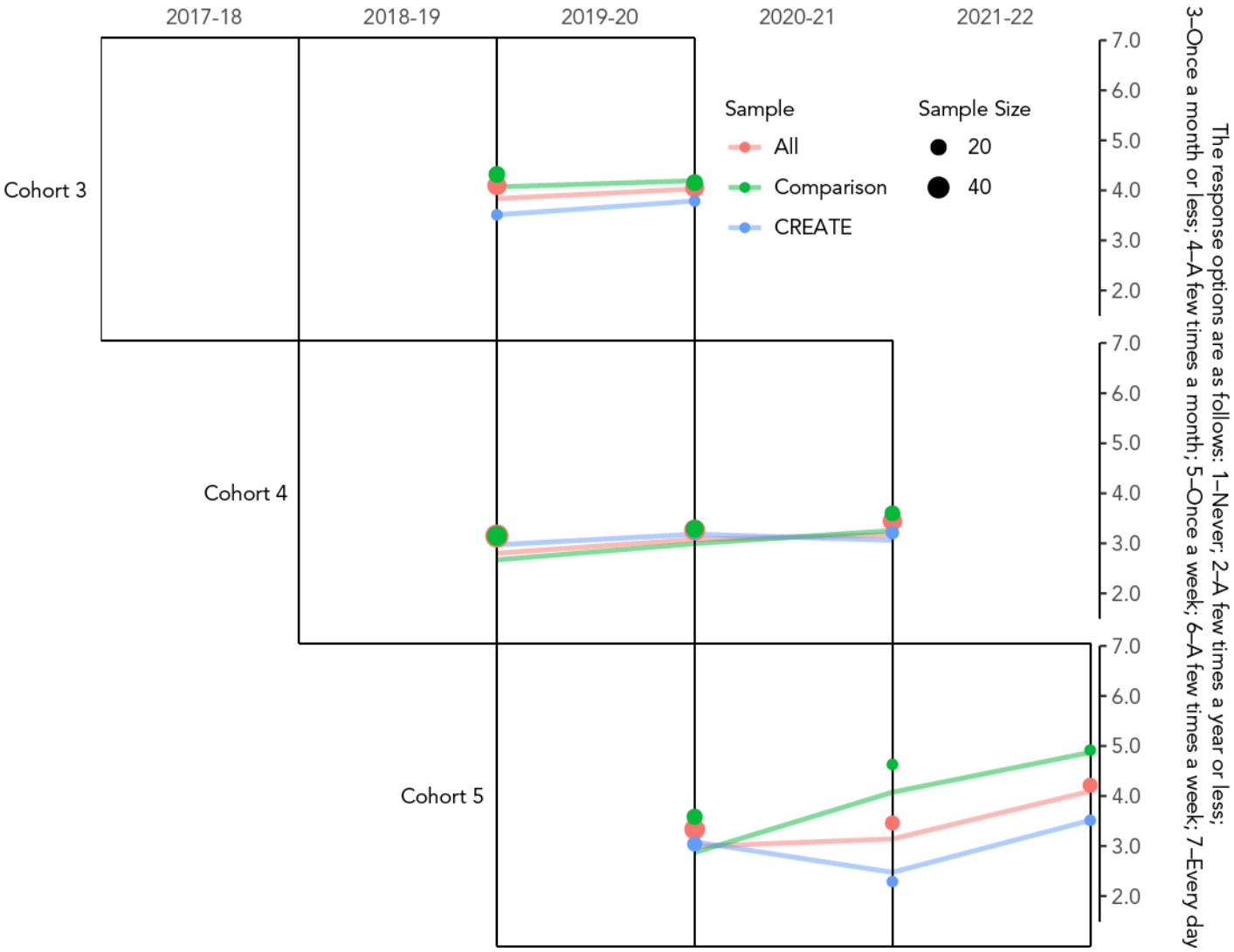

FIGURE 19. CHANGES IN COHORT 3–5 CREATE AND COMPARISON PARTICIPANTS RESPONSES OVER TIME ON THE MASLACH BURNOUT: EMOTIONAL EXHAUSTION SUBSCALE

Note. The 9-item Emotional Exhaustion (EE) subscale of the Maslach inventory assesses feelings of being emotionally overextended and exhausted by work. Higher scores correspond to greater experienced burnout. The lines connect score averages for the consistent sample of 59 matched cases (27 in CREATE and 32 in comparison); the dots represent outcomes for all available teachers responding in a given cohort and year. This scale was added during the 2018–19 school year, so we do not have data for Cohort 3 during their student teaching year.

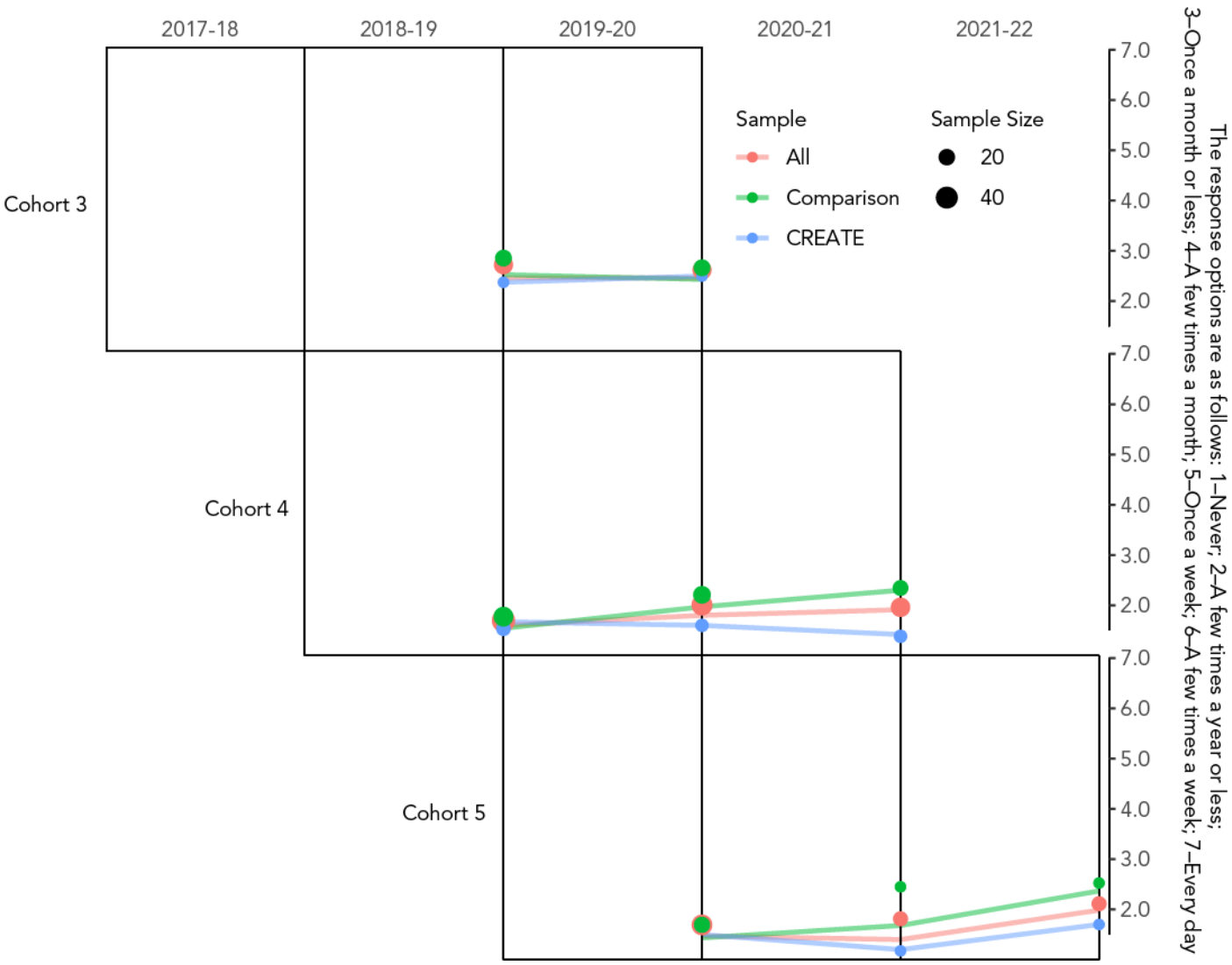

FIGURE 20. CHANGES IN COHORT 3–5 CREATE AND COMPARISON PARTICIPANTS RESPONSES OVER TIME ON THE MASLACH BURNOUT: DEPERSONALIZATION SUBSCALE

Note. The 5-item Depersonalization (DP) subscale of the Maslach inventory measures an unfeeling and impersonal response toward recipients of the service, care, treatment, or instruction. Higher scores correspond to greater degrees of experienced burnout. The lines connect score averages for the consistent sample of 59 matched cases (27 in CREATE and 32 in comparison); the dots represent outcomes for all available teachers responding in a given cohort and year. This scale was added during the 2018–19 school year, so we do not have data for Cohort 3 during their student teaching year.

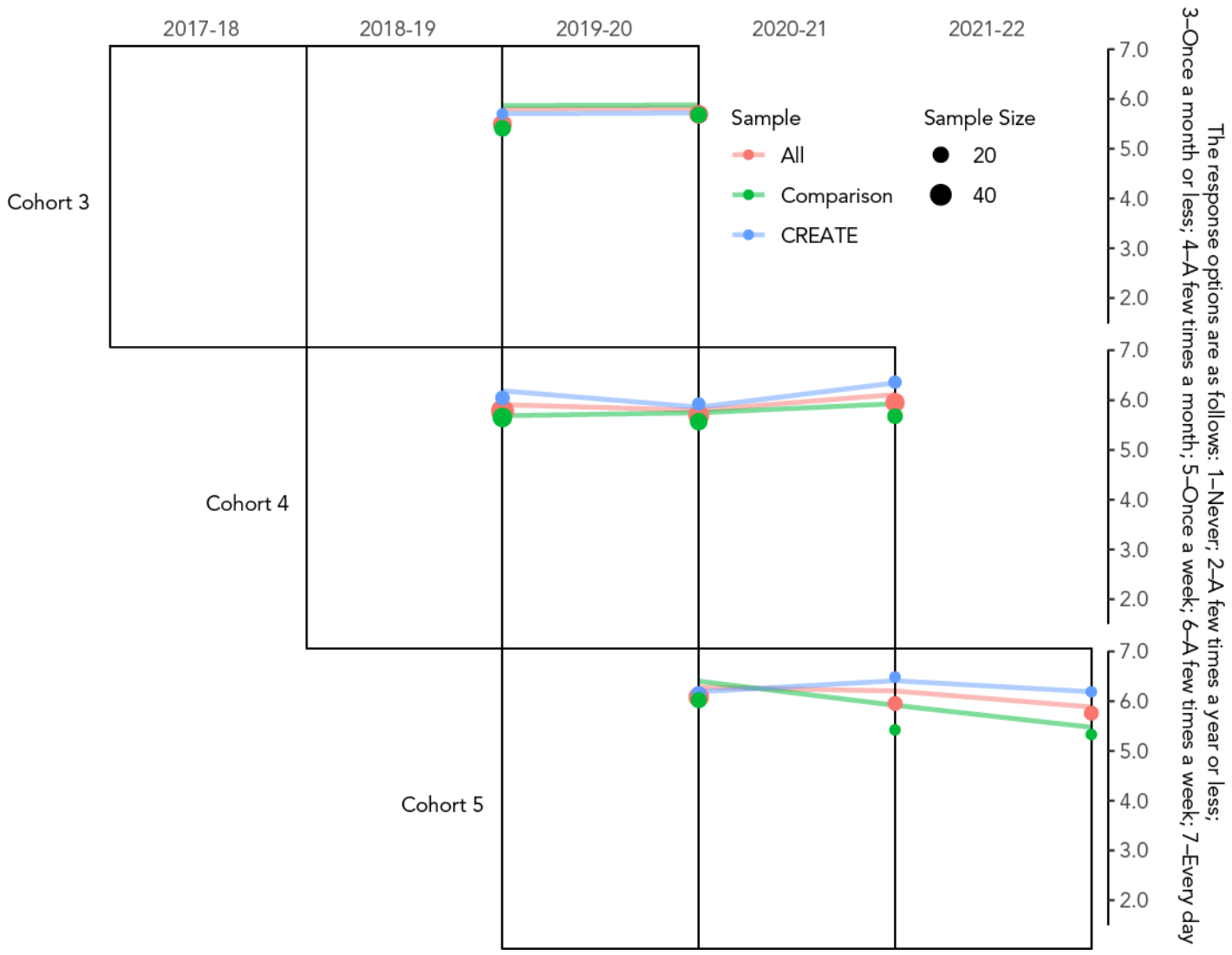

FIGURE 21. CHANGES IN COHORT 3–5 CREATE AND COMPARISON PARTICIPANTS RESPONSES OVER TIME ON THE MASLACH BURNOUT: PERSONAL ACCOMPLISHMENT SUBSCALE

Note. The 8-item Personal Accomplishment (PA) subscale of the Maslach inventory assesses feelings of competence and successful achievement in the work with people. Lower scores correspond to greater experienced burnout. The lines connect score averages for the consistent sample of 59 matched cases (27 in CREATE and 32 in comparison); the dots represent outcomes for all available teachers responding in a given cohort and year. This scale was added during the 2018–19 school year, so we do not have data for Cohort 3 during their student teaching year.

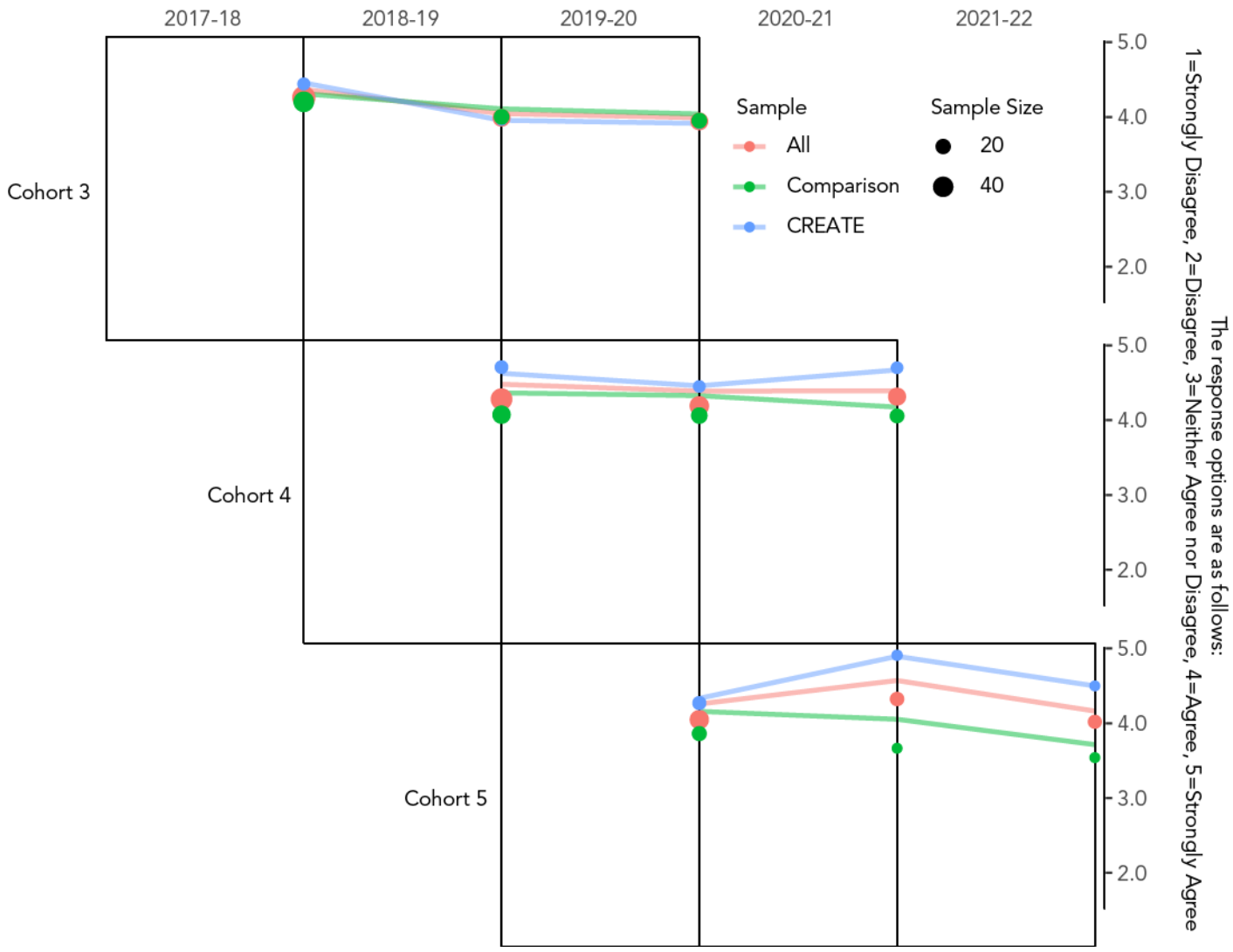

FIGURE 22. CHANGES IN COHORT 3–5 CREATE AND COMPARISON PARTICIPANTS RESPONSES OVER TIME ON THE STRESS MANAGEMENT AND EMPATHY RELATED TO TEACHING SCALE

Note. The 6-item Stress Management and Empathy Related to Teaching scale is a researcher-developed six-item scale. Higher scores correspond to greater degrees of stress management and empathy related to teaching. Note. The lines connect score averages for the consistent sample of 59 matched cases (27 in CREATE and 32 in comparison); the dots represent outcomes for all available teachers responding in a given cohort and year. This scale was added during the 2018–19 school year, so we do not have data for Cohort 3 during their student teaching year.

Chapter 5. Confirmatory Impacts on Student Achievement

✨ Research Question 3 ✨

Is there a positive impact of CREATE on student achievement in elementary and middle grades on average three years after the start of residency?

By the end of their second year of teaching, students of CREATE teachers did not outperform matched samples of students in classes of comparison teachers on ELA achievement, as measured by the Georgia Milestones Assessment System.

Specifically, among the 29 CREATE and 52 matched comparison students, the estimated impact on ELA outcomes based on the benchmark model, with adjustment for pretest and other covariates, was -10.78 scale score units (standardized effects size of -.176), (p = .537). We show this covariate-adjusted result in Table 3, (as the "adjusted effects size"). We also show the impact without covariate adjustments (as the "unadjusted effect size"). Baseline equivalence on the pretest was achieved for the confirmatory analysis of impact on ELA.

TABLE 3. IMPACT OF CREATE ON STUDENT ELA ACHIEVEMENT (CONFIRMATORY ANALYSIS) FOR COHORT 3–5

	Condition	Means	Standard deviations	No. of students	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	476.86	57.06	52	3	-0.613	.063
	CREATE	439.35	67.90	29	2		
Adjusted effect size[b]	Comparison	476.86				-0.176	.532
	CREATE	466.08					

Note. The *p* values are for the corresponding impact estimates in the regression model.

[a] The unadjusted effect size is the impact estimate from a model with cohort indicators and teacher and student random effects, divided by the pooled standard deviation of the outcome variable.

[b] The adjusted effect size estimate is the point estimate for impact from the benchmark model, which includes all the components of the unadjusted model, with the addition of student covariates divided by the pooled standard deviation of the outcome variable.

Impacts on student achievement in mathematics could not be estimated because of a lack of outcomes data, attributable in part to interruption in data collection during COVID-19. The situation is explained further in Appendix J, which describes the sampling and matching processes.

Chapter 6. Impacts on Early Career Teacher Retention

✨ Research Questions 4 and 7–10 ✨

Is there a positive impact of CREATE, compared to the GSU BAU program, on teacher retention three years after start of residency (i.e., through the second year of teaching)?

Is there a positive impact of CREATE, compared to the GSU BAU program, on teacher retention three years after start of residency among Black educators?

What is the impact of CREATE on completion of the teacher preparation program at GSU CEHD and teacher retention into the first year of teaching for the overall sample?

Is there a positive impact of CREATE, compared to the GSU BAU program, on teacher retention through the third year after start of residency, among teachers who completed teaching in the second year of residency, for the full sample?

Are impacts on teacher retention after 1, 2, and 3 years increasing by cohort for the overall sample and among Black educators?

We evaluated the impact of CREATE on teacher residents' graduation from GSU CEHD and their uninterrupted retention in teaching in Georgia public schools for the subsequent two years. Outcomes for CREATE residents were compared to those in the comparison group. We examined impact one, two, and three years after joining the study for Cohort 3, 4, and 5. We examined impacts for the full sample and for the sample of Black educators.

We did not observe an impact of CREATE on retention through the second year of teaching, either for the full sample or the sample of Black educators. Similarly, we did not observe impact on rates of graduation or persistence in teaching through the first year, for the full sample or the sample of Black educators. There was no impact on rates of retention through the second year of teaching among matched samples of teachers who were retained through the first year of teaching, and impacts did not vary after 1, 2, and 3 years of teaching across the three teacher cohorts.

DETERMINING TEACHER STATUS

We rely on a variety of sources to determine the status for each study participant at the three time points. For graduation from GSU CEHD, we rely on data from our participant tracker, participant surveys, data provided to the research team by GSU or the CREATE program team, and teacher certification records from the Georgia

Professional Standards Commission (Georgia Professional Standards Commission, 2022). For teaching in year 1 and teaching in year 2, we triangulate data received on teacher surveys, database received from the CREATE program team, data from GADOE, and teaching records from Open Georgia: Transparency in Government travel and salary database (Open Georgia, 2008). We describe the process we used to determine graduation and teaching status in Georgia Public schools in Appendix K.

For this analysis, we base retention status on on-time teaching. On-time graduating and transitioning to teaching means that a teacher graduates from GSU CEHD during the expected year and becomes a teacher of record in the two consecutive years after graduating from GSU CEHD. That is, we are not analyzing whether a teacher—at some point within the three-year trajectory—graduates and enters teaching. We are analyzing whether a teacher transitions into teaching on time, indicating a faster and higher return on investment in a teaching career.

RETENTION RATES FOR THE FULL SAMPLE OF STUDY PARTICIPANTS

Table 4 and Figure 23 display levels of graduation and retention for the matched samples of participants for both the CREATE and comparison group, as well as percentages of individuals whose status could not be confirmed (e.g., these teachers may have moved to non-public schools within Georgia, moved to schools outside of Georgia, or may not be teaching at all).[10] Each bar shows the constant total of teachers retained, not retained, and for whom outcomes are missing, through graduation, teaching after one year, and teaching after two years.

[10] Individuals were matched separately within each cohort. The description of the matching procedure is in Appendix K. The matched samples were similar in terms of baseline reported levels of math anxiety, confidence in general teaching skills, and motivation for entering teaching (there was less than .25 standard deviations difference between conditions on baseline characteristics). The results of the baseline equivalence tests are reported in Appendix K.

TABLE 4. NUMBERS AND PERCENTAGES OF COHORT 3–5 CREATE AND COMPARISON TEACHERS RETAINED, NOT RETAINED OR WITH RETENTION STATUS MISSING AMONG MATCHED SAMPLES

	CREATE			Comparison		
	Retained	Not retained	Status unknown	Retained	Not retained	Status unknown
Graduated	44 (95.65%)	2 (4.35%)	0 (0.00%)	91 (86.67%)	6 (5.71%)	8 (7.62%)
Teaching in year 1	37 (80.43%)	6 (13.04%)	3 (6.52%)	76 (72.38%)	17 (16.19%)	12 (11.43%)
Teaching in year 2	37 (80.43%)	6 (13.04%)	3 (6.52%)	68 (64.76%)	13 (12.38%)	24 (22.86%)

Note. Some percentages may not add to 100% due to rounding.

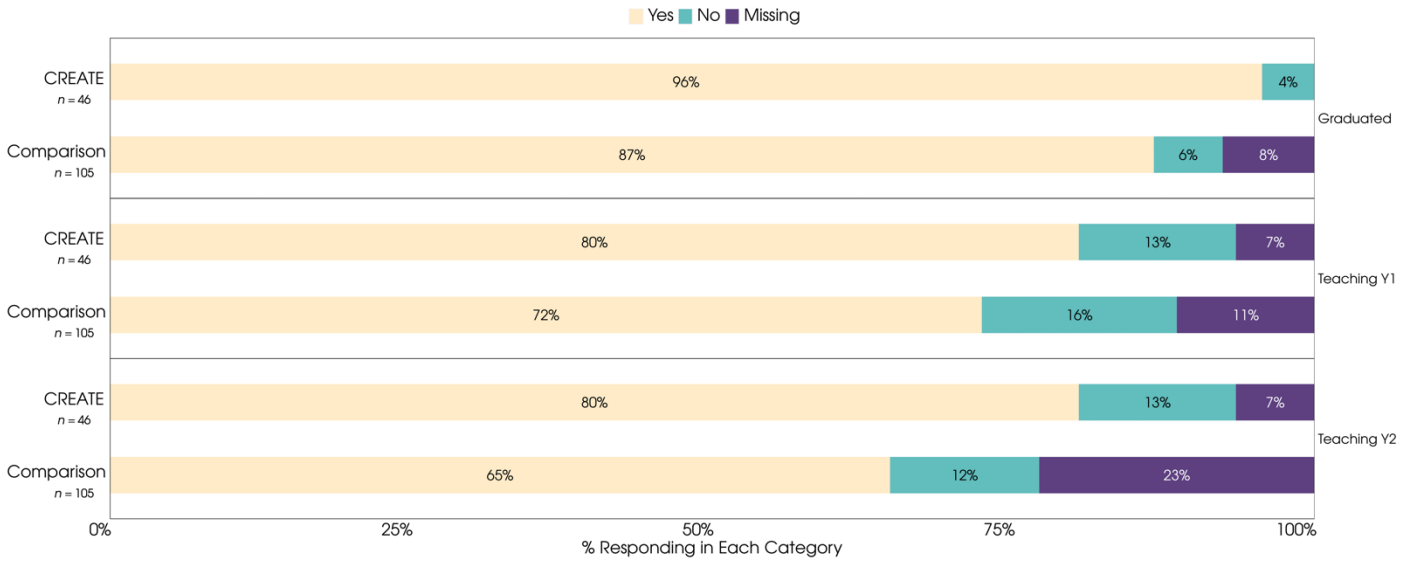

FIGURE 23. PERCENTAGES OF COHORT 3–5 CREATE AND COMPARISON TEACHERS RETAINED AMONG MATCHED SAMPLES

Note. The full sample sizes for generating the bars in this graph (inclusive of teachers in Cohort 3–5) are 46 teachers in CREATE and 105 in comparison. These samples include individuals with non-missing values for covariates and after matching CREATE and comparison cases within cohort. The full sample across cohorts, and before matching, consisted of 47 teachers in CREATE and 124 in comparison. Participant counts at each stage of limiting the data are displayed in Appendix K in Tables K3 through K14.

The proportion retained for the available samples is the ratio of the percentage retained (percentage in peach in Figure 23) to the percentage of retained plus not retained (percentage in peach plus the percentage in teal in Figure 23). These proportions are reported as percentages in Table 5 below. The proportions retained through the second year of teaching among non-missing cases is 86% (37 of 43) for CREATE and 84% (68 of 81) for comparison.

TABLE 5. PERCENT OF TEACHERS GRADUATED FROM GSU CEHD AND RETAINED IN CREATE AND COMPARISON GROUPS AMONG MATCHED INDIVIDUALS WITH NON-MISSING COVARIATES AND OUTCOMES FOR RETENTIONS STATUS, COHORTS 3–5

	CREATE	Comparison	Difference
Graduated	95.65%	93.81%	1.84%
Teaching in year 1	86.05%	81.72%	4.33%
Teaching in year 2	86.05%	83.95%	2.10%

INFERENTIAL RESULTS

Main Impact Finding

We evaluated the difference in the probability of retention three-years after joining the study (through the second year of teaching). We used a linear probability model (Appendix K), adjusting for effects of cohort and possible differences between conditions in the distributions of covariates used to evaluate baseline equivalence, and in the proportions of Black educators in the two conditions. The regression-adjusted difference between conditions in the probability of being retained was .015 (SE = .072) and is not statistically significant (p = .83). We expect a difference of this magnitude or larger to be observed 83% of the time due to sampling variation, when the difference is actually zero.

We substantiated this result by also conducting a logistic regression adjusting for the same covariates, which yielded a difference between CREATE and comparison in the log odds of retention of .186 (SE = .56), p = .74 (with corresponding odds ratio of 1.20, and Cox ratio of .186/1.65 = .113.

Additional Analyses

Impacts on Retention

We additionally explored the difference in the probability of retention both one and two years after joining the study (i.e., through graduation and the first year of teaching). Applying the linear probability model, conditional on cohort, and three baseline covariates (scale scores for levels of math anxiety, confidence in teaching, and motivation to enter teaching), the regression-adjusted difference between conditions **in the probability of graduating** is .023 (SE = .041), and is not statistically significant (p = .587). Alternatively, using logistic regression and adjusting for the same covariates, yielded a difference between CREATE and

comparison **in the log odds of being retained through a first year of teaching** of .550 (SE = .918), $p = .549$ (with corresponding odds ratio of 1.73, and corresponding Cox ratio of .550/1.65 = .333).

Applying the linear probability model, the regression-adjusted difference between conditions **in the probability of being retained through the first year of teaching** is .052 (SE = .070) and is not statistically significant ($p = .458$). Alternatively, using logistic regression and adjusting for the same covariates, yielded a difference between CREATE and comparison **in the log odds of being retained through a first year of teaching** of .302 (SE = .545), $p = .580$ (with corresponding odds ratio of 1.352, and corresponding Cox ratio of .302/1.65 = .183).

Examining Potential Confounding Effects of Enrollment in the MAT Program

A concern raised by the CREATE team was that a much larger proportion of teachers were pursuing a Master's of Arts in Teaching (MAT) degree at GSU in the comparison group than in the CREATE group, and that this could have had a bearing on the retention outcome.[11] For example, MAT candidates are more likely to be career-switchers and, therefore, may be older. Additionally, while the practicum and student teaching experience is the same for both groups, students pursuing their B.S.Ed have a whole year (their junior year) where they are taking education-specific courses about pedagogy, equity/social justice, and philosophy of education, so they likely have more exposure to these topics than MAT students.

Table 6 shows the distribution for the matched sample and breaks down the numbers in terms of retention status after two years of teaching, by condition and by degree enrollment.

It is clear that the comparison students are more likely to be enrolled in the MAT program (35 of 105) compared to their CREATE counterparts (3 of 46). To evaluate the relationship between retention and MAT program enrollment, we examined the cross-tabulation of retention status and MAT program enrollment for the comparison group (the comparison group had larger cell numbers to supports a statistical test). Applying Fisher's Exact Test, we observed no correlation between retention status and MAT enrollment ($p = .782$) indicating that enrollment in the MAT program did not confer an advantage to being retained. We further examined the impact of CREATE on retention, conditional on MAT program enrollment status. The regression-adjusted difference between conditions in the probability of being retained was .018 (SE = .076) $p = .811$. We note a limitation of this result is that the parametric test depends on the very small count of individuals in CREATE who enrolled in the MAT program.

[11] Eligible participants were either perusing the following degrees at GSU: Elementary Education, B.S.Ed., Middle Level Education, B.S.Ed., or a Master of Arts in Teaching.

TABLE 6. NUMBERS OF TEACHERS IN MAT PROGRAM VERSUS THE BACHELOR OF SCIENCE IN EDUCATION PROGRAM DURING THEIR ENROLLMENT AT GSU FOR THE MATCHED SAMPLE, COHORTS 3–5

	Bachelor of Science in education			Master of Arts in teaching		
	Retained	Not retained	Status unknown	Retained	Not retained	Status unknown
CREATE	N = 35	N = 5 Sum of N = 43	N = 3	N = 2	N = 1 Sum of N = 3	N = 0
Comparison	N = 44	N = 10 Sum of N = 70	N = 16	N = 24	N = 3 Sum of N = 35	N = 8

Examining Whether Impacts on Retention Increased Across Years for Teachers at the Same Point in Their Career Trajectory

We also explored whether impacts on retention after 1, 2, and 3 years in the program were changing with each consecutive cohort. For participants with non-missing retention outcomes, there were no differences across the cohorts in impacts on rates of graduation (p = .259), completion of the first year of teaching (p = .637), and completion of the second year of teaching (p = .836).

Impacts on Retention Through the Second Year of Teaching Among Teachers Who Were Retained Through a First Year of Teaching

Finally, we report impacts of CREATE on retention through the third year of the program (i.e., the second year of teaching) among teachers who were retained through the second year of the program (i.e., through the first year of teaching). Although teachers in both conditions are teachers-of-record in their first year of teaching, in CREATE, teachers are offered the opportunity to co-teach with another resident. This makes the transition from their first year of teaching to their second year of teaching a critical juncture, because that is when teachers in both conditions are fully transitioned to standard teaching loads (i.e., no longer co-teaching). Most teachers in the comparison group transition to a standard teaching load immediately after graduating from GSU; whereas CREATE supports a slower progression into the role of sole teacher-of-record through co-teaching during their first year. A contrast of these two models allows a test of the question of whether slower progression supports longer retention, at least in the short horizon through their second year of teaching. The answer to this question may have value for the field generally (and it can support evaluation of whether longer retention offsets the added cost of co-teaching).

In the context of this study, the question and corresponding analysis also allows the test of a contrast, potentially reviewable using WWC evidence standards, given the availability of pretest and demographic data for students in classes of teachers in their first year of teaching. This allows an evaluation of baseline

equivalence of classes for teachers in both conditions, in terms of both the class composition and teacher baseline attributes.

We found that all 16 CREATE teachers who were teachers of record in their first year continued to teach in their second year. In the comparison group, 30 of 32 (94%) teachers of record in their first year remained in teaching the second year. (The number of teachers available for this analysis is more-limited because not all teachers had data available regarding the composition of the students in their classes in their first year of teaching.)

A Fisher's Exact Test showed no difference between conditions in the probability of retention ($p \sim 1.00$).[12]

RETENTION RATES FOR THE SAMPLE OF BLACK EDUCATORS

Table 7 and Figure 24 display levels of graduation and retention for the matched samples of participants for both the CREATE and the comparison group, as well as percentages of individuals whose status could not be confirmed among Black educators.[13]

TABLE 7. NUMBERS AND PERCENTAGES OF COHORTS 3–5 CREATE AND COMPARISON TEACHERS RETAINED, NOT RETAINED OR WITH RETENTION STATUS MISSING AMONG MATCHED SAMPLES OF BLACK EDUCATORS

	CREATE			Comparison		
	Retained	Not retained	Status unknown	Retained	Not retained	Status unknown
Graduated	27 (96.43%)	1 (3.57%)	0 (0.00%)	25 (83.33%)	1 (3.33%)	4 (13.33%)
Teaching in year 1	22 (78.57%)	4 (14.29%)	2 (7.14%)	23 (76.67%)	5 (16.67%)	2 (6.67%)
Teaching in year 2	22 (78.57%)	4 (14.29%)	2 (7.14%)	20 (66.67%)	5 (16.67%)	5 (16.67%)

The proportions retained for the available samples is the ratio of the percentage retained (percentage in peach in Figure 24) to the percentage of retained plus not retained (percentage in peach plus the percentage in teal in Figure 24). We report these proportions as percentages in Table 8 below. The proportions retained through the

[12] For this analysis, the sample was narrowed to include only teachers for whom we were able to obtain student achievement scores and demographic data in the year before they joined the study classes of the teachers in their first year as teachers of record. This allowed us to assess baseline equivalence between teachers in the two conditions, in terms of the composition of their students. For the sample of 48 teachers considered here, we observe baseline equivalence both on student-level pretest and on the proportion of students who are Black or Hispanic. Full technical details for WWC review are in Appendix K.

[13] The approach to within-cohort matching is the same as was described earlier for the full sample. The results of the baseline equivalence tests are reported in Appendix K.

second year of teaching among non-missing cases is 85% (22 of 26) for CREATE and 80% (20 of 25) for comparison.

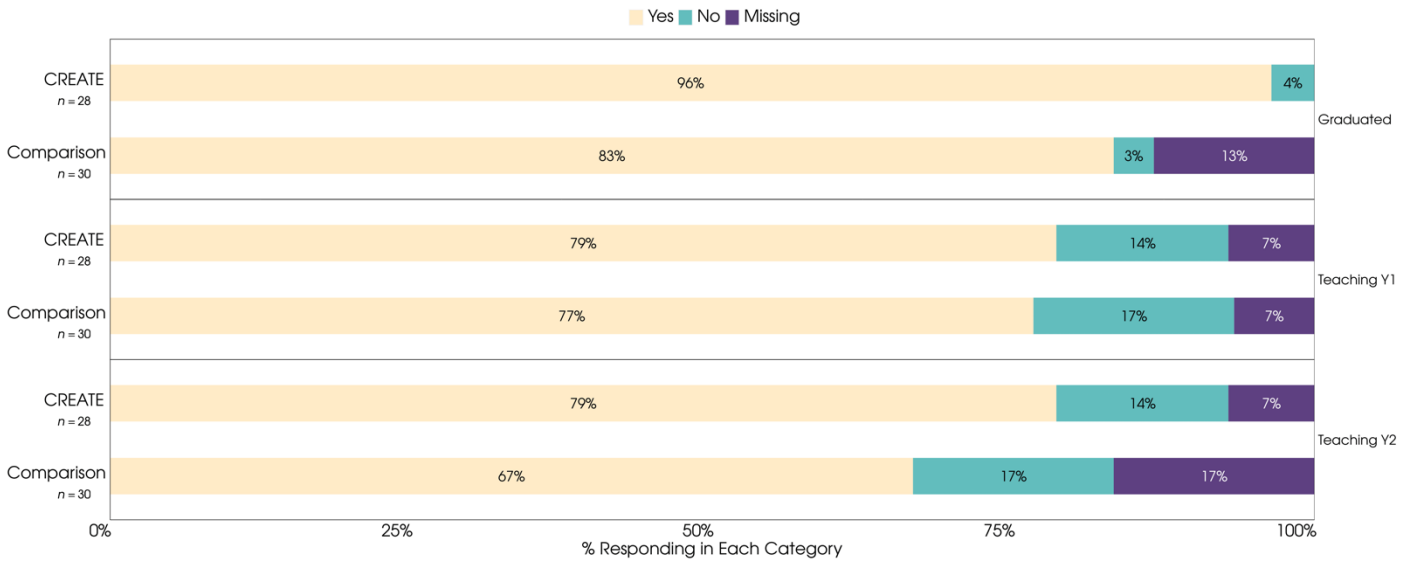

FIGURE 24. PERCENTAGES OF COHORT 3–5 CREATE AND COMPARISON TEACHERS RETAINED AMONG MATCHED SAMPLES FOR BLACK EDUCATORS

Note. The full sample sizes for generating the bars in this graph (inclusive of teachers in Cohort 3–5) are 28 teachers in CREATE and 30 in comparison. (These samples include individuals with non-missing values for covariates and after CREATE and comparison cases were matched within cohort.) The full sample across cohorts, before matching, consisted of 28 teachers in CREATE and 33 in comparison. Participant counts at each stage of limiting the data are displayed in Appendix K in Tables K15 through K26.

TABLE 8. PERCENT OF TEACHERS GRADUATED FROM GSU CEHD AND RETAINED IN CREATE AND COMPARISON GROUPS AMONG MATCHED INDIVIDUALS WITH NON-MISSING COVARIATES AND OUTCOMES FOR RETENTIONS STATUS, BLACK EDUCATORS, COHORTS 3–5

	CREATE	Comparison	Difference
Graduated	96.43%	96.15%	0.27%
Teaching in year 1	84.62%	82.14%	2.47%
Teaching in year 2	84.62%	80.00%	4.62%

INFERENTIAL RESULTS

Main Impact Finding

We evaluated the difference in the probability of retention three years after joining the study (through the second year of teaching). We used a linear probability model (Appendix K) adjusting for effects of cohort, and differences between conditions in the three covariates used to evaluate baseline equivalence. The regression-adjusted difference between conditions in the probability of being retained was .027 (SE = .100), and is not statistically significant (p = .79). We expect to observe a difference of this magnitude or larger 79% of the time due to sampling variation, when the difference is actually zero.

We substantiated this result by also conducting a logistic regression adjusting for the same covariates, which yielded a difference between CREATE and comparison in the log odds of retention of .256 (SE = .811), p = .752 (with corresponding odds ratio of 1.29, and Cox ratio of .256/1.65 = .155).

Additional Analyses

Impacts on Retention

We additionally explored the difference in the probability of retention one and two years after joining the study (i.e., through graduation and first year of teaching). Applying the linear probability model, conditional on cohort, and three baseline covariates (scale scores for levels of math anxiety, confidence in teaching, and motivation to enter teaching), the regression-adjusted difference between conditions **in the probability of graduating** is -.007 (SE = .047), and is not statistically significant (p = .882).[14]

Applying the linear probability model, the regression-adjusted difference between conditions **in the probability of being retained through the first year in teaching** is .020 (SE = .099), and is not statistically significant (p = .840). Alternatively, using logistic regression and adjusting for the same covariates, yielded a difference between CREATE and comparison **in the log odds of being retained through a first year of teaching** of .190 (SE = .771), p = .805 (with corresponding odds ratio of 1.21, and corresponding Cox ratio of .109/1.65 = .115).

Examining Potential Confounding Effects of Enrollment in the MAT Program

Table 9 shows the distribution for the matched sample and breaks the numbers down in terms of by retention status after two years of teaching, by condition, and by degree enrollment.

It is clear that the comparison students are more likely to enroll in a MAT program (11 of 30) compared to their CREATE counterparts (3 of 28). Applying Fisher's Exact Test, we observed no correlation between retention status and MAT program enrollment (p = .551). (As with the analysis for the full sample, we evaluated this

[14] The logistic regression could not produce a maximum likelihood estimate, likely reflecting the very high levels of retention on both groups (27/28 in CREATE and 25/26 in comparison) We do not report the difference in log odds of retention between conditions.

relationship for the comparison group only, given the marginally larger sample). We further examined the impact of CREATE on retention, conditional on MAT program enrollment status. The regression-adjusted difference between conditions in the probability of being retained was -.001 (SE = .103) p = .994. We note a limitation of this result is that the parametric test depends on the very small count of individuals in CREATE who enrolled in the MAT program.

TABLE 9. NUMBERS OF TEACHERS IN MAT PROGRAM VERSUS THE BACHELOR OF SCIENCE IN EDUCATION PROGRAM DURING THEIR ENROLLMENT IN GSU BAU FOR THE MATCHED SAMPLE AMONG BLACK EDUCATORS, COHORTS 3–5

	Bachelor of Science in education			Master of Arts in teaching		
	Retained	Not retained	Status unknown	Retained	Not retained	Status unknown
CREATE	N = 20	N = 3	N = 2	N = 2	N = 1	N = 0
		Sum of N = 25			Sum of N = 3	
Comparison	N = 14	N = 3	N = 2	N = 6	N = 2	N = 3
		Sum of N = 19			Sum of N = 11	

Examining Whether Impacts on Retention Increased Across Years for Resident and Teachers at the Same Point in Their Career Trajectory

We also explored whether impacts on retention after 1, 2, and 3 years were changing with each consecutive cohort. For participants with non-missing retention outcomes, there were no differences across the cohorts in impacts on rates of graduation (p = .260), completion of the first year of teaching (p = .811), and completion of the second year of teaching (p = .766).

PROJECTED RETENTION INTO THE SEVENTH YEAR OF TEACHING

In addition to evaluating impacts on achieved teacher retention through the second year of teaching, we asked teachers in both conditions to project what job they intended to engage in five years after participating in the final year of the study (that is, into their seventh year as an educator). Specifically, on the final survey of their third year in the study, we asked what job they expect to be doing in five years. Answer options included the following.

- Regular classroom teacher
- A different professional role in education (e.g., counselor, school administrator)
- A different non-professional role (e.g., school aide, office administrator) in education
- A different professional role (e.g., curriculum developer) in a commercial/business educational setting
- A different job not related to education

- Not working any job by choice
- Not working any job out of necessity
- Other, please explain: _____

We examine the contrast between conditions descriptively (without conducting a statistical test). The results are displayed in Figure 25. Two points are noteworthy. First, there is not a sharp difference between CREATE and comparison teachers in their expectation to remain classroom teachers five years out. Second, only approximately half the teachers in both conditions expect to continue as a classroom educator.

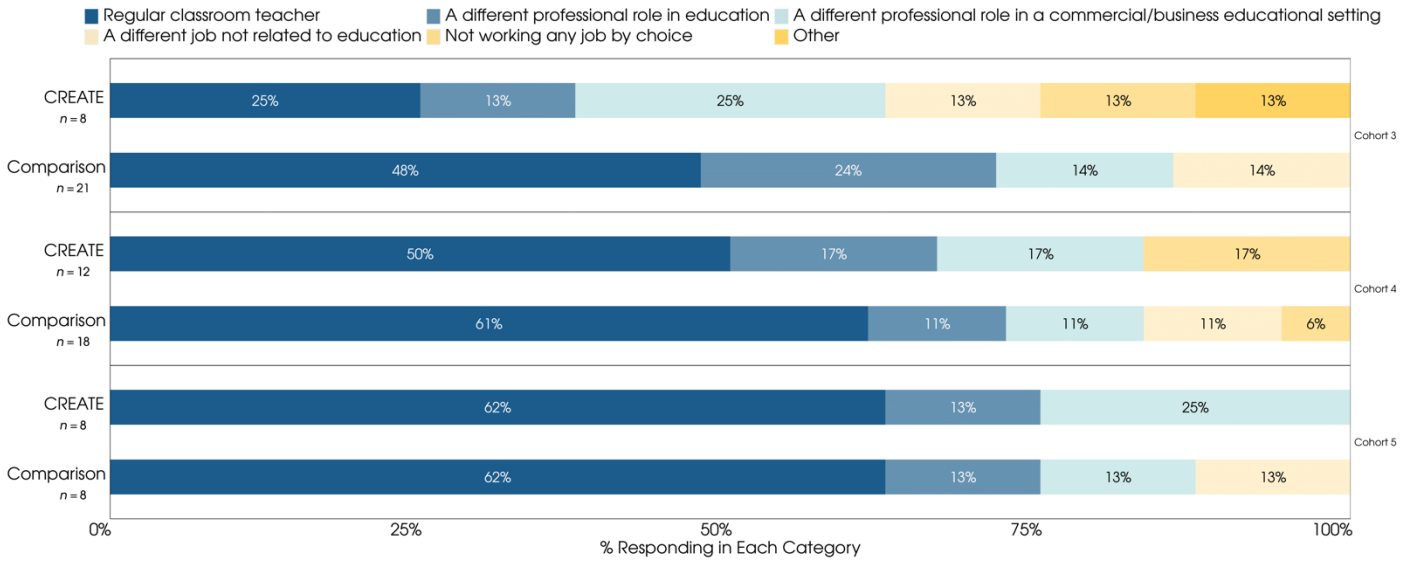

FIGURE 25. COHORT 3–5 CREATE AND COMPARISON PARTICIPANTS PROJECTED JOB EXPECTATION FIVE YEARS AFTER PARTICIPATING IN THE CREATE STUDY

Source: Quarterly surveys

Chapter 7. Discussion

Since its inception in 2015, the CREATE Teacher Residency program has received multiple grants at the state and federal levels and from private foundations—to provide support to new and veteran educators. CREATE recruits a diverse pool of committed educators to begin their residency during their final year of working towards their teaching credential at GSU CEHD. CREATE then provides three years of ongoing and consistent support to new teachers through their second year of teaching. Programming that includes PL opportunities, mentorship, and peer support through a cohort model, is designed to equip teachers to incorporate equitable and culturally relevant practices in their classrooms. At its core, CREATE's work is "in service of reimagining schooling and classroom spaces and to move toward the idea of education for liberation." This core idea drives the research questions and methods.

This independent evaluation of the implementation and impact of CREATE adds to the evidence base for the program from a prior development-level Investing in Innovation (i3) grant, which focused on the first two cohorts of early career teachers. It also bridges the follow-on Supporting Effective Educator Development (SEED) grants (awarded in 2020 and 2022) that hold promise for scaling the program while focusing on its core mission. This report discusses the evaluation of the impact of the CREATE program on student and teacher outcomes for Cohorts 3, 4, and 5, who began their residencies in the 2017–18, 2018–19, and 2019–20 school years, respectively. (The i3 grant supported residents in Cohort 1 and 2).

CREATE IMPLEMENTATION

This evaluation of Cohorts 3–5 demonstrates a predominantly successful implementation of CREATE, treatment-control contrasts on several program components, and trends towards favorable impacts on key potential mediators of impact. CREATE consistently met implementation for placing residents in progressive classroom roles through their three years in the residency, equity-centered PL, and attendance of residents at SRA before their first year of teaching. The Compassion-Based Professional Learning component did not meet fidelity for Year 1 and Year 2. This was due to fewer than expected Cohort 4 and Cohort 5 residents attending Together Time meetings where the PL took place during the 2019–20 school year (Year 2 of the residency for Cohort 4, and Year 1 for Cohort 5). All three cohorts met fidelity in Year 3 for the Compassion-Based Professional Learning component. The mentoring component did not meet fidelity in the years measured due to mentors not having attended the number of expected mentor trainings. This may point to the ongoing growth of a sustained community of veteran educators working with CREATE, who continue to take on mentorship roles year after year after having received their initial training in CREATE values and practices. Considering this, threshold measures for this implementation component should be revisited to incorporate any updated expectations for new and returning mentors.

Based on survey responses from participants regarding their schooling and career experiences during the beginning of the COVID-19 pandemic, as well as communications with CREATE, we know that pandemic-related teaching and employment challenges played a large part in the teaching experiences of early career

teachers during the years we assessed fidelity of implementation for these components. We believe the impact of some of these challenges are reflected in these implementation results.

IMPACTS ON TEACHERS' SELF-COMPASSION, LEVELS OF BURNOUT AND STRESS MANAGEMENT, AND EMPATHY AS RELATED TO TEACHING

We observed trends in impacts of CREATE in the direction of increased self-compassion, reduced burnout, and greater stress management and empathy related to teaching. We evaluated impacts across the three cohorts of teachers after their second year as teachers of record. The largest differences between CREATE and comparison were observed for Cohort 5. The observed variation in responses across cohorts and over time is not surprising given that each cohort experienced stresses and interruptions from COVID-19 at different times in their three-year participation in the program. Because we evaluated the main impacts per scale across cohorts at the end of the three-year program implementation, an additional test was conducted to assess whether 3-year impacts (that is, after the second year of teaching) varied across cohorts. One could expect, for example, that as a program matures, impacts may increase from one cohort to the next. A test of a difference across cohorts in impact after three years did not reach conventional levels of statistical significance (alpha = .05); that is, differences in impact across the three cohorts after three years of program participation were not so large that they would be hard to account for by chance variation.

The results should be interpreted in light of the challenging context of the COVID-19 pandemic during which the program implementation was conducted, which was an external shock that may have adversely affected participants' feelings regardless of which condition they were in. Another point to consider is that for some of the scales, CREATE participants were close to the ceiling of possible response options. For example, for the Personal Accomplishment subscale of the Maslach Teacher Burnout Inventory, which measures feelings of competence and successful achievement in work with other people, responses from CREATE teachers for Cohort 4 and 5 in their second year as teachers-of-record were between *a few times a week* and *every day* (i.e., the two most positive responses available). It may be unrealistic to expect almost all individuals to uniformly respond *every day*. Comparison group responses on average fell between *once a week* and *a few times a week*. Additionally, the large effect sizes, adequately large samples, but lack of statistical significance, imply substantial variability in responses. Exploring why individuals in either condition responded low may provide insight into areas for program improvement that in the future will produce more-uniformly positive responses among CREATE participants.

In general, the direction of the effects, taken together, indicates promise of the protective effects of CREATE on teachers' capacities to cope effectively with challenges of the profession, including potential stressors.

IMPACTS ON TEACHER INSTRUCTION (QUALITY OF INSTRUCTION AND TEACHING ENVIRONMENT)

COVID-19 curtailed the opportunity to carry out the intended evaluation of impacts of CREATE on important distal outcomes. Notably, in response to COVID-19, the state suspended collection of teacher performance data (including on project-critical dimensions of quality of instruction and educational environment). Several

districts also reasonably prohibited the alternative of using in-person classroom observations. Further, suspension of state testing in 2020 limited the sample of students for evaluating impacts on achievement.

IMPACTS ON STUDENT ACHIEVEMENT

We set out to evaluate impacts on student achievement in mathematics and ELA among three cohorts of teachers after their second year as teachers of record. Impacts on ELA, based on outcomes for 81 students, did not reach statistical significance. Impacts on mathematics could not be evaluated due to lack of availability of sufficient outcomes data. We did achieve baseline equivalence between the CREATE and comparison groups for the analytic sample for impact on ELA, but the small number of cases greatly reduces the scope and external validity of the conclusions. While we can feel confident that we are comparing outcomes for similar cases across conditions, we do so for only a small subset of the intended sample.

Several unexpected limitations arose during the study that ultimately resulted in small analytic samples. Georgia administers the Milestones state assessment in grades 4–8, and many participating teachers in both conditions taught in lower elementary school grades. For the analysis phase, many factors resulted in small student samples: reduced teacher samples, the technical requirement of matching of students across conditions within each cohort in order to meet WWC evidence standards, and the need to match students within grades, given the lack of vertically scaled scores.

Additionally, COVID-19 had several unexpected effects on the impact study with implications for the achieved sample sizes and the evaluation of impacts. Most notably, the lack of student achievement data from Spring 2020 resulted in a two-year lag between when baseline achievement measures were collected (Spring 2019) and achievement outcomes were assessed (Spring 2021). This reduced our sample size (we could not include Grade 4 outcomes because we lacked Milestones pretests from the prior year in third grade because of COVID, and Milestones is not administered in second grade). In turn, the smaller sample reduced statistical power to detect impact. Statistical power was likely further reduced from the weaker predictiveness of a lagged pretest on posttest performance. Additionally, because no student outcomes data were obtained for the school year 2019–20, it prevented us from analyzing impacts on Milestones outcomes for Cohort 3 in Spring 2020 as was planned in the original study proposal (developed in 2017). Instead, we evaluated student outcomes in Spring 2021 for Cohort 3, 4, and 5 (on average in the second year as teacher of record).

Some of the more persistent challenges of obtaining adequate student samples for analysis of impacts on achievement have led us to explore the use of other assessments for subsequent future cohorts of CREATE. This includes the possibility of using established district assessments that are acceptable to WWC, that evaluate a broader grade range, and that may have fewer limitations, given the design requirements of the quasi-experiment such as the need to match students across conditions within cohort.

IMPACTS ON TEACHER RETENTION

We observe weak trends of CREATE participants surpassing those in the comparison group in retention, among matched and non-missing cases, both for the full sample, and the sample of Black educators. However,

the difference in "on-time" teaching retention after three years (i.e., following two years of being teachers of record) are not statistically significant. This result is not consistent with findings from the i3 study (Jaciw et al., 2021), in which we found a statistically significant and positive impact of CREATE on uninterrupted retention of teachers through their second year as teachers of record among Black educators.

The following are some possible explanations for the lack of observed impact.

1. Missingness may be correlated with non-retention. If this is the case, then the greater proportion of missing cases among the comparison group in the third year may lead to overestimation of retention rates among comparison cases. Using multiple sources, we undertook a thorough effort to trace and contact all participants (Appendix K). The CREATE team also maintained a detailed database of residents and their teaching status. In August 2022, CREATE made a concerted effort to contact all residents from prior cohorts to determine their teaching status. They shared this information with us to triangulate our records.
2. GSU's traditional teacher preparation program may be becoming a more-formidable counterfactual. Experiences for both CREATE residents and comparison group participants are changing, with an evolving social landscape contributing to this change. Changes in the GSU program may be contributing to graduation and longer-term retention. Additionally, it may be useful to examine the kinds of employment supports and continuity into teaching that districts offer first year teachers in both conditions from GSU. The difference between this study and the i3 study in uninterrupted retention through the third year after enrolling in GSU CEHD among Black educators is almost entirely due to the difference between studies in graduation rates from GSU. In both studies, teachers stayed retained, provided they graduated. This adds credibility to the idea that the GSU traditional teacher preparation program has evolved to increase graduation rates among Black educators – comparable to rates for residents who go through CREATE.

Program implementation for all three cohorts overlapped with COVID-19. We evaluated retention rates in the periods shown in Table 1. The most-disruptive year (2020–21) affected Cohort 4 (in its second year of teaching) and Cohort 5 (in its first year of teaching). Effects on program continuity and retention in teaching were likely felt by participants in both conditions.

IMPLICATIONS

We look forward to continuing the evaluation of the CREATE program and its mechanisms and impacts, through the recently awarded SEED grant. It is well recognized that programs in the social sciences and their contexts change. The context for CREATE is no exception. Continuing evaluation efforts will clarify the conditions for impact and expand understanding of the supporting mechanism, especially as the program itself evolves in response to societal needs.

An important point of focus going forward, is to monitor the changing nature of the counterfactual program to which CREATE is compared. The flashpoints of the pandemic and social unrest have mobilized responses by

various stakeholders: government at all levels, program developers, schools of education, and schools across the nation. A possible result is a general strengthening of the teacher induction process. Specifically, this may include changes to components of standard one-year teacher preparation programs, and the mechanisms for supporting the transition from schools of education into teaching. Such changes may be reflected in programming in CEHD at GSU to encourage successful graduation, and in the support offered to secure employment upon graduation. For example, in the i3 CREATE study, the overall graduation rate from the non-CREATE GSU program was 85% and from the CREATE program, it was 98%. The results were across two cohorts enrolled in the teacher residency in their first year in 2015–16 and 2016–17. This contrasts with the graduation rates 94% from the non-CREATE GSU program and 96% for the CREATE program, for Cohort 3–5 (with enrollment in GSU CEHD in the years 2017–18, 2018–19 and 2019–20, respectively). The near-ceiling effect in graduation rates observed in the current study demonstrates the strong efficacy of both programs with a diminished contrast between the programs in their effects on graduation. Future work may be an opportunity for conceptual replication, with the range of counterfactual teacher preparation programs being expanded, and with CREATE continuing on its current track in development, to achieve a stronger contrast and establish the added value of CREATE program components.

Finally, an important question with tremendous social value is about CREATE's potential to mitigate the pandemic's damage to teaching and learning, and the well-being of teachers and students. CREATE's commitment to equity-centered induction is especially vital in recovery, given the disproportionate negative effects of the pandemic on people of color and on schools that serve predominantly students of color. CREATE's model, and its commitment to produce and use research in critical and impactful ways, will contribute to the national conversation about how to effectively address challenges that schools and districts are facing in attracting and retaining qualified teachers. We are hopeful that the ongoing evaluation of CREATE continues to reveal supports and values that are effective in bringing about sustainable structures for teacher induction programs, in Atlanta and beyond.

References

Bloom, H. S., Michalopoulos, C., & Hill, C. J. (2005). Using Experiments to Assess Nonexperimental Comparison-Group Methods for Measuring Program Effects. In H. S. Bloom (Ed.), *Learning more from social experiments*, 173-235. Sage Foundation.

Dill, K. (2022). School's Out for Summer and Many Teachers Are Calling It Quits. *The Wall Street Journal*. https://www.wsj.com/articles/schools-out-for-summer-and-many-teachers-are-calling-it-quits-11655732689

Georgia Professional Standards Commission (2022). *Check Certification Status*. https://www.gapsc.com/Certification/Lookup.aspx

Jaciw, A. P., Wingard, A., Zacamy, J., Lin, L., & Lau, S. (2021). *Final Report of the i3 Evaluation of the Collaboration and Reflection to Enhance Atlanta Teacher Effectiveness (CREATE) Teacher Residency Program: A Quasi-Experiment in Georgia.* (Empirical Education Rep. No. Empirical_GSU-7031-FR1-2021-O.1). Empirical Education Inc. https://www.empiricaleducation.com/create/

Kraft, M, A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher, 49*(4), 241-253.

Kush, J. M., Badillo Goicoechea, E., Musci, R. J., & Stuart, E. A. (2022). Teacher mental health during the COVID-19 pandemic. *Educational Researcher*. Prepublished November 15, 2022. https://www.doi.org/10.3102/0013189X221134281.

Meadows, L., & Theodore, K. (2012). *Teacher Preparation Programs: Research and Promising Practices.* Briefing Paper. Texas Comprehensive Center.

National Center for Teacher Residencies (NCTR). (2023, February 22). *Doing Better for Black Educators Webinar Recording*. [Webinar]. Black Educators Initiative. https://www.youtube.com/watch?v=L6ly_h2q-Mg

Open Georgia. (2008). State of Georgia Financial Reports. http://open.georgia.gov/

This report was prepared for the Office of Elementary and Secondary Education (OESE) under Contract U423A170027 funded by the Supporting Effective Educator Development (SEED) grant program. The content of the publication does not necessarily reflect the views or policies of OESE or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

The full report, appendices, and research summary are at https://www.empiricaleducation.com/create/.

Reference this report: Jaciw, A. P., Lau, S., Zacamy, J., & Lin, L. (2023). Evaluation of Georgia State University's CREATE Teacher Residency Program – Cohorts 3 through 5: A Final Report of a Quasi-Experiment in Georgia. (Empirical Education Rep. No. Empirical_GSU-7035-FR-2023-O.2). Empirical Education Inc. https://www.empiricaleducation.com/create/