Appendix

Evaluation of Georgia State University's CREATE Teacher Residency Program—Cohorts 6 through 8

A FINAL REPORT OF A QUASI-EXPERIMENT IN GEORGIA

September 2026

Andrew P. Jaciw

Sze-Shun Lau

Chelsey Nardi

Jenna Zacamy

Li Lin

Lindsay Maurer

Mayah Waltower

Empirical Education Inc.

Table of Contents

Appendix A. Description of CREATE Programming 1
Appendix B. Participant Recruitment 5
Appendix C. Details about Study Participation 6
Appendix D. Data Collection Sources 12
Appendix E. Data Collection Activity Response Rates 20
Appendix F. Fidelity of Implementation Matrix 25
Appendix G. Fidelity of Implementation Results 35
Appendix H. Technical Details of Teacher Well-Being Analysis 52
Appendix I. Detailed Analysis of Emotional Exhaustion Impacts (to Support WWC Review) 76
Appendix J. Retention Data Sources, Coding, and Decision Tree 94
Appendix K. Technical Details on Retention Analysis 98
Appendix L. Interview Protocol 105
Appendix M. Interview Coding Scheme 107
Appendix N. Economic Evaluation 128
References 140

Reference this report: Jaciw, A. P., Lau, S., Nardi, C., Zacamy, J., Lin, L, Maurer, L., & Waltower, M. (2026). Evaluation of Georgia State University's CREATE Teacher Residency Program—Cohorts 6 through 8: A Final Report of a Quasi-Experiment in Georgia. (Empirical Education Rep. No. Empirical_GSU-7036-FR-2026-O.3). Empirical Education Inc. https://www.empiricaleducation.com/create/

Appendix A. Description of CREATE Programming

General Residency Structure and Support

CREATE's three-year teacher residency program is designed to support and foster a community of critically-conscious, compassionate, and skilled educators who are committed to racial justice and interrupting inequities in education. During Year 1 of the residency, while in their final year of their teacher credentialing program at Georgia State University's College of Education and Human Development (GSU CEHD), residents complete their preservice teaching fieldwork requirement by serving as student teachers in K–8 classrooms within CREATE partner schools in Atlanta Public Schools (APS). Year 1 residents work closely with the teacher-of-record in the classroom, referred to as the Cooperating Teacher (CT), who serves in a mentor role to CREATE residents for the entire duration of the school year. CREATE CTs attend a Summer Mentor Institute each year and receive invites to workshops and coaching calls throughout the academic year. These professional learning (PL) opportunities introduce CTs to CREATE's mentorship model, *Coaching for Equity* by Elena Aguilar (2020), and provide a space for CTs to share their experience and build their skills alongside other mentors. Residents also work with a University Supervisor Coach (USC), who works to navigate and align the needs and understandings of both the resident and the CT. This support is designed to maximize CT-guided resident learning by prompting new understandings found at the intersection of university coursework and school-based contexts. The grouping of the USC, CT, and resident is termed "the triad."

As residents move through the three-year residency model, their role within the classroom changes. In Year 2 of the residency, once residents have completed their credentialing program at GSU CEHD, CREATE residents apply for and take positions to complete their first year as a teacher-of-record. During this year, residents have the option to be paired with another CREATE teacher, who serves as their co-teacher, in a single classroom. CREATE designed the co-teaching model to allow opportunities for collaboration and learning alongside a peer who is in the same place in their teaching journey, as well as to allow for the sharing of teaching responsibilities around planning, instruction, assessment, and data management. While working as a co-teacher in Year 2, each resident receives a full-time teaching salary, as well as support and mentorship from an experienced teacher working at the same school, referred to as their SBM. The SBM helps residents build social connections, well-being, and a sense of belonging among equity-minded, experienced educators. Residents also meet regularly with a CREATE Instructional Mentor (IM) who works with a small caseload of residents, provides targeted support in content mastery, equity-centered instructional practices, and compassion-based mindfulness skills. Residents benefit from personalized coaching and observation cycles with their IM. Additionally, IMs lead Together Time (TT), fostering professional growth through structured discussions and reflective practice (described below). In Year 3, residents become the sole teacher-of-record in their own classroom, and continue to receive support from an SBM and an IM.

In addition to the progressive classroom roles and multiple forms of mentorship that evolve with CREATE residents' growth as teachers throughout the three-year residency, residents receive intensive summer training and ongoing support from their resident cohort and the CREATE program team each year. Residents participate in a Summer Resident Academy (SRA) for a five-week intensive training during the summer after their student teaching year (that is, between Year 1 and Year 2 of the residency), and for a two-week intensive training during the summer after their first year as full-time teachers (that is, between Year 2 and Year 3 of the residency). Through SRA, CREATE guides residents in developing social-emotional competencies, pedagogical skills, content knowledge, understanding about the context of where they are teaching, skills to increase their well-being, and the confidence they will need for success in their teaching careers. The SRA is a structured learning experience designed to support CREATE residents in their journey as educators, progressing from foundational learning to hands-on teaching and reflective practice. In the summer after Year 1, SRA focuses on

classroom practices, as residents teach summer school under the guidance of university faculty, CREATE leaders, and district representatives. The curriculum emphasizes creating a warm and welcoming classroom culture while deepening residents' understanding of equity, justice, privilege, and oppression in schools. Residents connect these critical topics and their real-world classroom experiences while also learning strategies for managing their emotions and stress. In the summer after Year 2 of residency, SRA is a space for reflection and future planning as residents prepare for their second year as credentialed teachers in APS. Led by CREATE and GSU CEHD faculty, the first week is dedicated to reflecting on the previous year, identifying successes, and addressing ongoing challenges. In the second week of SRA, the focus shifts to lesson planning, with an emphasis on promoting students' academic growth, social emotional learning (SEL), and development as critically conscious, engaged citizens.

TT meetings, which occur approximately monthly throughout the school year, serve as opportunities for residents to gather, collaborate, and problem-solve with those in their resident cohort. Led by IMs, the goal of the meetings is to foster teacher well-being, community, professional growth, and equity-centered teaching practices throughout the residency. TT meetings include a compassion-based component that supports SEL, emotional regulation, and compassion towards self and others, as well as an equity-based component that provides opportunities to explore racial identity, privilege, and systemic oppression. The meetings provide a structured space for residents to increase well-being by (a) deepening their connections with peers, (b) engaging in mindfulness practices to increase emotional regulation and self-compassion, and (c) examining their role in the context of broader structures. Residents develop their instructional expertise by (a) critically examining education in the context of metro Atlanta, (b) learning culturally responsive lesson planning guided by Universal Design for Learning, (c) exploring equitable methods of formative and summative assessment, and (d) practicing building inclusive learning environments. They engage in professional problem solving and receive guidance from CREATE USCs. Additionally, TT meetings offer trauma-informed, well-being support to help residents heal from racialized trauma and build resilience. Through reflective discussions, coaching, and structured learning experiences, TT meetings equip residents with the knowledge, skills, and emotional support necessary to grow as effective, equity-driven educators.

Moreover, CREATE residents receive financial support in the form of yearly stipends throughout their participation in the residency, in order to work towards program accessibility.

Recruitment Into CREATE

CREATE recruits students who are entering their final year of college, pursuing their Bachelor of Science in Education (BSEd) in Birth Through Five (B-5), Elementary Education, and Middle Level Education, as well as those pursuing their Masters of Arts in Teaching (MAT) in Elementary Education, Middle Level Education, or K-12 English for Speakers of Other Languages (ESOL). Staff members at CREATE conduct presentations at GSU CEHD, providing students with an overview of the three-year residency program and inviting them to submit an application to become a resident. These presentations usually take place in the spring and summer, setting students up to begin their residency at the beginning of their final year of study at GSU CEHD. CREATE admits students into the program based on a variety of information in the applications, including the students' degree of interest in teaching in historically-underserved communities in Atlanta and how their goals and experiences relate to CREATE's mission. Through these recruitment efforts, the CREATE program team is dedicated to contributing to the diversification of the teacher workforce.

Professional Learning Opportunities for Experienced Educators

In addition to the residency program for new teachers, CREATE offers a series of PL opportunities for experienced educators. The majority of participants in these PL opportunities work in CREATE partner schools, but those who work in schools that are not partners of CREATE may also attend the PL opportunities. CREATE's partner schools are given a

limited number of complimentary seats for each PL opportunity, increasing accessibility for educators in the CREATE community.

Based on the understanding that educator efficacy benefits from the existence and accessibility of a like-minded community of colleagues, CREATE's PL offerings aim to support experienced educators in equity-centered and compassion-based work. CREATE offers PL opportunities that focus on developing teacher leaders within CREATE partner schools in the form of CTs and SBMs, training experienced educators to work one-on-one with CREATE residents as mentors in professionalism, instructional skills, classroom practice, collaboration, reflection skills, and other essential skills for teaching. CREATE also offers equity-centered and compassion-centered PL opportunities, which allow experienced educators to work and learn alongside other educators, to collaborate and develop their own teaching practices. CREATE's primary equity-centered PL opportunity through summer 2022 was the Equity-Centered Critical Friendship (ECCF) Institute, at which educators focused on developing understandings of individual and collective identity and liberatory practices, as they contribute to educational equity. CREATE's primary compassion-based PL opportunity through Summer 2022 was Presence, Power, Impact (PPI), which supported participants in strengthening self-awareness, self-compassion, and their contemplative practices that helped cultivate personal resilience during challenging teaching moments (CREATE personal communication, 2021). PPI drew upon practices taught at a previous iteration of PPI—Cognitively-Based Compassion Training© (CBCT)—at which CREATE engaged research-based practices born out of Emory University's Center for Contemplative Science and Compassion-Based Ethics to foster educator resilience through increased mindfulness (CREATE, personal communication, 2019).

Leadership Restructuring

The CREATE teacher residency program has continuously evolved since its inception, in response to shifts in needs, social context, and leadership members and structure. In the spring of 2022, CREATE initiated an organizational restructuring to a) formally recognize the Black leadership already evident throughout its program, and b) distribute expertise across more leadership roles, as the program had expanded significantly. CREATE moved from a two-leader (one executive director and one principal investigator) model—both White women—to a five-co-director model. Under the new model, each co-director manages a major component of the CREATE program: the residency program (two directors), research (one director), sustainability and operations (one director), and community engagement and continuing education (one director). The five new co-directors—all Black women—work closely with the original principal investigator and former program lead. Those who entered the new co-director roles were already part of the CREATE community, with demonstrated leadership. Alongside the change in leadership, a continual refocusing and refining of programming occurred to align with CREATE's evolving core values centered around uplifting and practicing abolitionist teaching principles. The new co-directors revised the CREATE curriculum to re-center Black brilliance and joy, incorporate freedom dreaming, and focus on culturally- and historically-relevant pedagogy (Muhammad, 2020; Muhammad 2023). The compassion-based portion of the programming moved away from being closely tied to CBCT, and towards an emphasis on rest and healing as a means to well-being. CREATE also acknowledged and worked to address the sense of moral injury that many educators face, resulting from exploitative practices embedded in education systems.

Local and National Context During the Study Years

As a structured source of support, CREATE inevitably works with residents as they navigate the development of their professional careers against local, national, and global events—and their aftermath. Specifically, during the grant years covered in this report, major events that impacted the CREATE community included the COVID-19 pandemic, as well as conversations and actions around racial and socioeconomic justice. During the 2020–21 school year (the first school year covered by this grant), educational and public health leaders had different ideas about how best to move forward to make

schools sites effective at learning while supporting the health of students and their families to varying degrees after the spring of the previous school year was abruptly disrupted by the global recognition of the COVID-19 pandemic and its dangers. In the summer of 2020, many American cities were experiencing social unrest and navigating conversations and demands for structural action towards racial justice after the killings of Ahmaud Arbery, George Floyd, and Breonna Taylor by police officers and former police officers. The disproportionate impact of COVID-19 on Black, Hispanic, and Native American people weaved itself into these conversations about injustice. Calls for the accountability of the police and others in positions of power and for the reform of systems that were created to harm poor people of color were in the spotlight nationally. Moreover, in the state of Georgia, House Bill 1084, which defines concepts relating to systemic racism and discrimination as "divisive" and bans them from being discussed in classrooms, was codified in the spring of 2022, directly impacting educators and organizations that work with educators, such as CREATE. An understanding of CREATE impact during these grant years requires an understanding of the context within which the organization's work occurred, and with which it grappled.

Appendix B. Participant Recruitment

Recruitment of study participants began in spring 2020 for Cohort 6, spring 2021 for Cohort 7, and spring 2022 for Cohort 8. To recruit participants each year, the research team delivered a presentation about the research study to students of GSU CEHD who were entering the final year of their teacher credentialing program in the Early Childhood and Elementary Education and Middle and Secondary Education departments.

Each year, we presented information about the background, eligibility requirements, and components of the research study to students who were in their final year of GSU CEHD's teacher credentialing program and we had identified as eligible for participation in the research. We recruited treatment cases from the pool of teacher candidates who had applied to participate in the CREATE teacher residency program, and who CREATE admitting into the program. We recruited comparison cases for the study from the pool of students who were eligible to participate in the CREATE teacher residency program, but who chose to not apply for the program.

In order for teacher candidates to be eligible for inclusion in the research study, they needed to:

- be enrolled in GSU CEHD during their first year of participation in the research study,
- be completing their student teaching requirement during their first year of participation in the research,
- plan to teach in a public school in Georgia once credentialed,
- plan to teach in a core subject area (math, ELA, science, and/or social studies) once credentialed,
- plan to teach in grades kindergarten through eighth grade once credentialed, and
- expect to complete the teacher certification requirements and graduate from GSU CEHD by the summer following their first year of participation in the research study.

To recruit participants for all cohorts, researchers held virtual recruitment events at which eligible GSU CEHD students learned about the research study and the expectations for participation in the study. The research team conducted presentations during courses or orientation events that GSU CEHD students attended, during which researchers provided eligible teacher candidates with information about the research study, data collection activities, and related stipends. Each recruitment presentation gave attendees an opportunity to ask questions, as well as instructions on how to complete a consent form if they were interested in participating in the study.

Researchers also recruited CREATE residents and comparison participants who were not in attendance at other virtual recruitment events by sending informative emails and posting fliers through contacts working for CREATE or for GSU CEHD who already had professional relationships with eligible teacher candidates. These emails included a link to a recorded presentation that, similar to the live presentations, provided information about the rationale for the study, eligibility requirements, and the data collection components study participants would engage in. These communications also included links to the online consent form, where interested teacher candidates could submit their interest and consent to participate in the study.

Once participants submitted consent forms for the study, members of the research team reviewed each consent form individually to confirm that interested participants met eligibility requirements, and then communicated with each interested individual about appropriate next steps.

Appendix C. Details about Study Participation

After agreeing to participate in the study, some participants left the study for a variety of reasons. The following section details attrition from the study: per cohort, each year, and for what reasons. Note that the sample for any given analysis varies based on availability of data. We continued to collect outcomes from the Georgia Department of Education (GaDOE), APS, or publicly available records after teachers left the study or stopped responding to surveys, as outlined in the study consent form.

Cohort 6 Summary of Attrition

Year 1: 2020–2021

During recruitment, the baseline sample for Cohort 6 consisted of 18 CREATE and 34 comparison group study participants. Six Cohort 6 participants—one CREATE participant and five comparison participants—became inactive during Year 1 of the study. The one CREATE participant did not complete the baseline survey and was non-responsive to data collection attempts. Of the five comparison participants, we made two inactive after we didn't receive principal permissions to collect data, one no longer wanted to participate in the research, and two others did not complete the baseline survey and were non-responsive to data collection attempts. Table C1 includes a summary of attrition for Cohort 6 Year 1.

TABLE C1. ATTRITION FROM STUDY FOR COHORT 6, YEAR 1 (2020–21)

Event/Reason for attrition	CREATE group	Comparison group
Baseline sample	18	34
(Did not receive research permissions)	(0)	(2)
(No longer wanted to participate in research/ Not responsive to data collection)	(1)	(3)
Remaining sample at end of Year 1	17	29

Year 2: 2021–2022

By the end of Year 2, an additional 24 Cohort 6 study participants had left the study. Seven were part of the CREATE group, and 17 were part of the comparison group. Two comparison participants rejoined the study in Year 2. We provide details below and in Table C2.

CREATE Group

Of the seven CREATE participants that left the study in Year 2, we made one participant inactive because they no longer wanted to participate in the study. An additional four participants were not responsive to our data collection attempts and were therefore made inactive. We made two participants inactive because they were no longer part of the expected sample due to no longer participating in the CREATE program: one was no longer living in Georgia and the other was not working in a classroom setting.

Comparison Group

Of the 17 comparison study participants who did not participate in the study in Year 2, we did not receive district permission to collect data from the school districts where three participants taught. Nine participants were non-responsive to our data collection attempts. Three participants became inactive due to no longer being part of the expected sample: one participant moved out of Georgia, and two participants were not working in K-8 classroom settings. We decided to make two additional comparison participants inactive and did not include them in the sample: one participant

was a full-time special education teacher, and the other did not student teach in Year 1 of the study. A factor in the decision to make these two participants inactive in Year 2 was that they were unlikely to have similar teaching experiences to a CREATE participant. Two comparison participants rejoined the study when we received permissions from district contacts to conduct research in their districts.

TABLE C2. ATTRITION FROM STUDY FOR COHORT 6, YEAR 2 (2021–22)

Event/Reason for attrition	CREATE group	Comparison group
Sample at beginning of Year 2	17	29
(Did not receive research permissions)	(0)	(3)
(No longer wanted to participate in research/ Not responsive to data collection)	(5)	(9)
(Not part of expected sample)	(2)	(3)
(Researcher chose not to include in sample)	(0)	(2)
Obtained district/school permission to survey participant	0	2
Remaining sample at end of Year 2	10	14

Year 3: 2022–2023

Seven study participants from Cohort 6 became inactive in Year 3. Two were part of the CREATE group, and five were part of the comparison group. We provide details below and in Table C3.

CREATE Group

Two CREATE participants became inactive in Year 3. Both participants were no longer part of the expected sample since they were no longer participating in the CREATE program: one participant was no longer teaching and the other was teaching outside of APS.

Comparison Group

Five comparison group study participants became inactive in Year 3. The research team did not receive permissions to collect data in the school district one participant was working in. One participant was nonresponsive to data-collection attempts. Three participants were no longer part of the expected sample since they were no longer working as teachers.

TABLE C3. ATTRITION FROM STUDY FOR COHORT 6, YEAR 3 (2022–23)

Event/Reason for attrition	CREATE group	Comparison group
Sample at beginning of Year 3	10	14
(Did not receive research permissions)	(0)	(1)
(No longer wanted to participate in research/ Not responsive to data collection)	(0)	(1)
(Not part of expected sample)	(2)	(3)
Remaining sample at end of Year 3	8	9

Summary of Cohort 6 Attrition

A total of 35 study participants (10 CREATE and 25 comparison) in Cohort 6 became (and remained) inactive between Year 1 and Year 3. See Table C for a summary table of attrition from the study sample in Cohort 6 during the three years of the study.

TABLE C4. SUMMARY OF COHORT 6 ATTRITION (2020–23)

Event/Reason for attrition	CREATE group	Comparison group
Baseline sample	18	34
(Did not receive research permissions)	(0)	(4)
(No longer wanted to participate in research/Not responsive to data collection)	(6)	(13)
(Not part of expected sample)	(4)	(6)
(Researchers chose not to include in sample)	(0)	(2)
Remaining sample at end of Year 3	8	9

COHORT 7 SUMMARY OF ATTRITION

Year 1: 2021–2022

For Cohort 7, CREATE offered two different residency pathways. Pathway 1 was a new one-year residency program. Pathway 3 was the three-year residency program that had been offered to previous cohorts. We have denoted which CREATE participants were in Pathway 1 vs. Pathway 3. During recruitment, the baseline sample consisted of 27 CREATE (five Pathway 1 and 22 Pathway 3) and 19 comparison group study participants. Nine Cohort 7 participants—four CREATE participant and five comparison participants—became inactive during Year 1 of the study. One Pathway 1 CREATE participant no longer wished to participate in the study. One Pathway 3 participant did not respond to data collection attempts. Two other Pathway 3 CREATE participants were no longer part of the expected sample since they were no longer participating in the CREATE program. Of the five comparison participants, we were unable to obtain district permissions to collect data in the school districts where two participants worked. One comparison participant no longer wanted to participate in the research, and two others were non-responsive to data collection attempts. Table C1 includes a summary of attrition for Cohort 7, Year1.

TABLE C1. ATTRITION FROM STUDY FOR COHORT 7, YEAR 1 (2021–22)

Event/Reason for attrition	CREATE group		Comparison group
Baseline sample	27		19
	Pathway 1	**Pathway 3**	
Baseline sample separated by pathway	5	22	
(Did not receive research permissions)	(0)	(0)	(2)
(No longer wanted to participate in research/ Not responsive to data collection)	(1)	(1)	(3)
(Not part of expected sample)	(0)	(2)	(0)
Remaining sample at end of Year 1 separated by pathway	4	19	
Remaining sample at end of Year 1	23		14

Year 2: 2022–2023

By the end of Year 2, an additional 20 study participants were no longer participating in the research. Of the 20 participants, 12 were part of the CREATE group (three of these participants were in Pathway 1, and 9 were in Pathway 3), and eight were part of the comparison group. By the start of Year 2, Pathway 1 CREATE residents had completed their one-year residency with CREATE and were no longer participating in the CREATE program. We continued to survey participants in both Pathways who were eligible and interested in participating in the study. We provide details below and in Table C2.

CREATE Group

Of the 12 CREATE participants that left the study in Year 2, two CREATE Pathway 1 participants no longer wanted to participate in the research. Another Pathway 1 participant was no longer part of the expected sample because they were teaching in high school. One Pathway 3 participant was non-responsive to data collection attempts. An additional eight Pathway 3 participants became inactive because they were no longer part of the expected sample. The following are reasons these eight participants stopped participating in the CREATE program.

- No longer living in Georgia (2 participants)
- Not teaching in K-8 (2 participants)
- No longer teaching (2 participants)
- Not teaching at a CREATE partner school (1 participant)
- Health reasons (1 participant)

Comparison Group

Of the eight comparison study participants who left the study in Year 2, we were unable to obtain research permissions to collect data from the school district where one participant taught. One participant no longer wanted to participate in the research, and three participants were unresponsive to our data collection attempts. Three participants were no longer part of the expected sample because two were no longer working as teachers, and the other was teaching pre-kindergarten.

TABLE C2. ATTRITION FROM STUDY FOR COHORT 7, YEAR 2 (2022–23)

Event/Reason for attrition	CREATE group		Comparison group
Sample at beginning of Year 2	23		14
	Pathway 1	**Pathway 3**	
Sample at beginning of Year 2, separated by pathway	4	19	
(Did not receive research permissions)	(0)	(0)	(1)
(No longer wanted to participate in research/Not responsive to data collection)	(2)	(1)	(4)
(Not part of expected sample)	(1)	(8)	(3)
Remaining sample at end of Year 2, separated by pathway	1	10	
Remaining sample at end of Year 2	11		6

Summary of Cohort 7 Attrition

A total of 29 Cohort 7 participants (16 CREATE and 13 comparison) became inactive over the course of Year 1 and Year 2. Of the CREATE participants, four were in Pathway 1, and 12 were in Pathway 3. See Table C3 for a summary table of attrition from the study sample in Cohort 7 during these two years of the study.

TABLE C3. SUMMARY OF COHORT 7 ATTRITION (2021–23)

Event/Reason for attrition	CREATE group		Comparison group
Baseline sample	27		19
	Pathway 1	**Pathway 3**	
Baseline sample, separated by pathway	5	22	
(Did not receive research permissions)	(0)	(0)	(3)
(No longer wanted to participate in research/Not responsive to data collection)	(3)	(2)	(7)
(Not part of expected sample)	(1)	(10)	(3)
Remaining sample at end of Year 2, separated by pathway	1	10	
Remaining sample at end of Year 2	11		6

COHORT 8 SUMMARY OF ATTRITION

Year 1: 2022–2023

During recruitment, 16 CREATE and 16 comparison group study participants agreed to be part of the study. Five Cohort 8 participants—three CREATE participants and two comparison participants—became inactive during Year 1 of the study. Two CREATE participants did not respond to our data collection attempts. One CREATE participant was no

longer part of the expected sample because they were no longer part of the CREATE program due to personal reasons. The two comparison participants were not responsive to our data collection attempts. Table C4 includes a summary of attrition for Cohort 8.

TABLE C4. ATTRITION FROM STUDY FOR COHORT 8, YEAR 1 (2022–23)

Event/Reason for attrition	CREATE group	Comparison group
Baseline sample	16	16
(No longer wanted to participate in research/Not responsive to data collection)	(2)	(2)
(Not part of expected sample)	(1)	(0)
Remaining sample at end of Year 1	13	14

Appendix D. Data Collection Sources

This study utilized multiple sources of data to measure desired outcomes, including participant surveys and well-being logs (WBL), administrative data from the GaDOE and school districts, publicly available data on teacher certification and teaching status, and CREATE program records. For both established and researcher-developed scales, a brief description is provided. However, only the items from researcher-developed scales are listed in the appendix; item-level details for established scales are not included. We describe these sources in more detail in this appendix.

Participant Surveys

Baseline Survey

We invited study participants from both the CREATE teacher residency program and the traditional credentialing program at GSU CEHD to complete the initial baseline survey after they agreed to participate in the research study. This baseline survey asked study participants questions about their teaching-related backgrounds, motivations, perspectives, and interests. Responses to this survey allowed us to confirm participants' eligibility for the research study, as well as inform the selection of the comparison group. We also used data from this survey in analysis: as variables for matching, as covariates in ANCOVA analysis, and as moderators in assessments of differential impacts. We used three baseline scales as variables for matching and to establish baseline equivalence for teacher survey outcomes. These scales include two researcher-developed scales—(1) Motivation for Entering Teaching, and (2) Confidence in General Teaching Skills—and an established scale, Connor-Davidson Resilience Scale (CD-RISC) (Davidson & Connor, 2016). We administered this survey to study participants one time only, when they joined the research study.

Motivation for Entering Teaching

Motivation for Entering Teaching is a researcher-developed scale which consists of the following five items.

1. I am interested in making a difference in the lives of children in Atlanta.
2. I believe that as a teacher I will contribute to social justice.
3. I am interested in how students learn.
4. I look at teaching as a process of self-discovery.
5. I am entering teaching because I want to eventually mentor teachers.

Each item allowed five response options from *strongly disagree* to *strongly agree*, which we coded as integers ranging from 1 to 5. We estimated the score for an individual by averaging that person's responses across the 5 items. A higher score on the scale means a participant self-reports that they have greater motivation to enter into teaching. In our previous evaluation, the scale had moderate reliability (α = .51; Jaciw et al., 2023).

Confidence in General Teaching Skills

Confidence in General Teaching Skills is a researcher-developed scale asking teachers to respond to how confident they are in their abilities within each of the following areas.

1. Collaborating with others
2. Reflecting on my actions
3. Planning and organization
4. Oral communication

5. Written communication
6. Working with children
7. Working with parents
8. Classroom management

Each item in the scale allowed five response options from *not at all confident* to *extremely confident*, which we coded as integers ranging from 1 to 5. We estimated the score for an individual by averaging that person's responses across the eight items. A higher score on the scale means a participant self-reports that they have greater confidence in general teaching skills. We found this scale to have good internal reliability (α = .81) in our previous evaluation (Jaciw et al., 2023).

Connor-Davidson Resilience Scale

The CD-RISC 10 is a 10-item scale that is based on a larger 25-item scale; it assesses resilience. The CD-RISC 10 has been validated across multiple populations, countries, stressor situations, and study designs; it has been used to "assess change during treatment with medication, psychotherapy, or from some other form of intervention" (Davidson & Connor, 2016). We asked participants to respond to the 10 (proprietary) items on a five-point Likert scale ranging from *not at all true* to *true nearly all of the time*. The CD-RISC 10 has a Cronbach's alpha of .87 (Davidson & Connor, 2016).

Quarterly Surveys

Participants in both conditions completed quarterly online surveys for their respective cohorts for the three years of the study. The design of these quarterly surveys, each taking an average of 20 minutes, considered the participants' teaching workloads. We planned deployment around the schedules of GSU CEHD and participating districts, avoiding breaks and testing windows. We asked teachers to complete all quarterly surveys outside of their work hours so as not to disrupt their teaching. Each semester, we deployed two surveys, for a total of four surveys per research year.

Quarterly surveys included questions related to support during participants' student teaching year and first two years of teaching, classroom experiences, and plans for continued teaching. Participants in the comparison group responded to items similar to those presented to CREATE residents regarding CREATE-specific supports. Comparison participants saw survey questions with more generic terminology regarding mentor teaching and professional learning/collaboration opportunities, in place of CREATE-specific language. Additionally, for both CREATE and comparison participants, we incorporated survey items in the quarterly surveys related to teacher well-being. These included scales to measure stress management and empathy related to teaching, self-compassion, racial equity and social justice in schools, and teacher social capital. All survey items were either researcher-developed or used components of established scales to investigate the proposed research questions regarding the different experiences, supports, and barriers between CREATE and comparison participants. See Appendix E for study participant response rates for each survey.

Stress Management and Empathy Related to Teaching

We assessed stress management and empathy related to teaching using six items from a researcher-developed scale. Each item in the scale allowed five response options ranging from *strongly disagree* to *strongly agree*. We estimated the outcome for an individual by averaging that person's responses across the following six items.

1. Since starting this school year, I have learned techniques that enable me to manage my reactions in a healthy, constructive way when faced with stressful situations.
2. Since starting this school year, I feel more confident that I will be able to handle the stress of being a teacher.
3. Since starting this school year, I feel more confident identifying and advocating for my own professional needs.

4. Since starting this school year, I have increased my commitment to helping my students learn and thrive.
5. Since starting this school year, I feel more able to understand the perspective of my students.
6. Since starting this school year, I feel more able to understand the perspective of my fellow teachers/colleagues.

A higher score on the scale indicates a participant self-report that they are more empathetic and effective at managing stress. The Cronbach's alpha (.92) for this scale, found in our previous evaluation, indicates excellent internal reliability (Jaciw et al., 2023).

Self-Compassion Scale

The Self-Compassion Scale (SCS) is a 26-item scale that assesses an individual's levels of self-compassion, according to three main components of self-compassion: 1) self-kindness (versus self-judgment), 2) a sense of common humanity (versus isolation), and 3) mindfulness (versus over-identification). We asked respondents a series of statements about their thoughts and feelings with regard to the above three components on a five-point scale of *almost never* to *almost always*. Total SCS scores have good internal reliability (Cronbach's alpha = .92; Neff, 2003). The six scales have Cronbach's alphas ranging from .75–.81. A large body of research also demonstrates strong construct validity; thus, SCS scores correlate with wellbeing (Neff, 2003).

Opportunities & Support to Think About Social Justice & Racial Equity in Schools

A core component of the CREATE teacher residency program is providing residents with the knowledge and tools to engage in social justice and racial equity practice with their school communities. We aimed to measure the opportunities and support given to both CREATE residents and comparison participants to think about social justice and racial equity issues at their schools. We measured this using an item that we co-developed with CREATE staff. During the 2021–22 school year (Cohort 6 Year 2), we asked participants to respond on a five-point Likert scale ranging from *almost never* to *almost always* about how often they are encouraged to think deeply about race- and equity-related issues in their school.

The following school year (that is, 2022–23, or Cohort 6 Year 3 and Cohort 7 Year 2), we combined this item with a revised set of items from the State of the American Teacher and State of the American Principal Surveys (Doan et al., 2022). We followed a formal process to obtain permission from the RAND team to incorporate these items into their surveys. For this new set of five adapted items, we asked participants to respond about how often their school has provided them opportunities or support for each item using a six-point Likert scale. The Likert scales measured frequency with the following options: *daily*, *weekly*, *monthly*, *at least once per semester*, *never*, or *other*. Using the adapted items, we asked participants to indicate how often their school has provided opportunities or supports to do the following.

1. Think deeply about social justice and/or racial equity issues
2. Incorporate antiracist teaching methods or content in the classroom (e.g., addressing historical events that reflect systemic racism or current events)
3. Help teachers to develop positive social identities based on their membership in one or more social groups (e.g., race, gender, sexuality, socioeconomic status, (dis)ability status)
4. Help teachers recognize and understand individual-level bias
5. Help teachers recognize and understand systemic oppression (e.g., racism, sexism, ableism)

Our analysis of the Cohort 7 Year 2 data indicates high internal reliability (Cronbach's alpha = .95).

Confidence in Understanding Concepts Related to Race & Equity in Schools

Similar to opportunities and supports to think about social justice and racial equity in schools, we also measured participants' confidence in their understanding of concepts related to race and equity in schools. To measure this, we created one researcher-developed item. CREATE residents and comparison participants were asked how confident they are in their understanding of concepts related to social justice and/or racial equity in schools on a five-point Likert scale that ranged from *not at all confident* to *extremely confident*.

Teacher Social Capital Scale

The Teacher Social Capital Scale (Minckler, 2011) is designed around the premise that relationships have value for helping teachers accomplish things they could not do alone (e.g., help students learn) and for satisfying a need to be a member of a community. The scale is organized around three primary domains: 1) Preconditions/Energizers (assesses opportunities for teachers to build and join communities at the schools), 2) Teacher Bonding Social Capital (assesses cohesiveness and strength of relationships among teachers within the school), and 3) Teacher Bridging Social Capital (assesses the links between teachers within the school to people or organizations outside of the school). We asked participants to respond on a five-point scale ranging from *strongly disagree* to *strongly agree*. Our analysis for the grant years focused on the subscales of teacher bonding and opportunities to collaborate. Across the subscales, the internal consistency reliability was considered high (Cronbach's alpha ranged from .94–.96; Lau et al., 2021).

Well-Being Logs

Well-Being Logs (WBLs) were an added component to the study during the 2021–22 study year (Year 2 for Cohort 6 and Year 1 for Cohort 7), and they continued in the years after. The decision to use WBLs emerged from discussion between Empirical and CREATE during which it was posited that effects of CREATE in earlier studies (Jaciw et al., 2021)—especially on retention in teaching—may be occurring through the insulating effects of TT meetings. In other words, impacts of PL may occur in the period after the meeting, in which case they might not be captured through our standard quarterly surveys.

Through the 2020 SEED grant, TT included practices designed to strengthen the resilience and compassion of teachers through CBCT, as well as to build skills that include mindfulness, communication, and a willingness to engage across differences that facilitate the building of meaningful relationships with students and colleagues. This, in part, led us to prioritizing measures of teacher well-being. To investigate this, we invited participants in both conditions to complete four online WBLs each study year. Deployment of the WBLs for CREATE residents and comparison participants centered around TT. During the 2021–22 school year, we randomly assigned participants to respond to one of four timepoints: (a) on the day of the TT meeting, (b) the day after the TT meeting, (c) two days after the TT meeting, or (d) 1 week after the TT meeting. During the next school year (2022–23), we randomly assigned participants to respond to WBLs to one of two timepoints: (a) on the day of the TT meeting or (b) four days after the TT meeting. We observed no decay between these time points, so we combined all data in the analysis. The WBLs took approximately 5 to 10 minutes to complete.

To address the need for momentary/ecological responses, the WBLs included items from momentary Positive Emotions, Engagement, Relationships, Meaning, and Accomplishment (mPERMA) and we added three additional items to create the Supported & Hopeful scale (Heshmati et al., 2023). Additionally, we incorporated items from the Maslach Burnout Inventory for Educators (Maslach et al., 2018) to capture participants' mindfulness, occupational health, well-being, and engagement. The Self-Regulation subscale from the Multidimensional Assessment of Interoceptive Awareness (MAIA-2; Mehling et al., 2018) provided information regarding participants lasting, rather than momentary, interoception. We included two items recommended by the external expert consultant (personal communication, January 12, 2022) to measure attention. Our team also co-developed with the external expert consultant the Focus & Awareness scale to

measure mindfulness through focusing attention. Together with the CREATE team, we co-developed the Sense of Equity in Role & Experience scale to measure the impact of racial equity work occurring during TT. We ended the WBLs with short open-ended questions asking respondents to reflect on stressors in the past few days and whether TT provided tools and conditioning for dealing with them. For additional details about the survey scales included on WBL, see Table D1 at the end of this section. See Appendix E for study participant response rates for each individual well-being log.

Focus & Awareness

Through collaboration with Dr. Robert Roeser, an expert on contemplative practices in education, we co-developed two items aimed to measure mindfulness through focusing attention and being aware of the present situation. We asked participants to respond how much they agreed with the following statements.

1. Today I was able to focus my attention.
2. Today I found myself very aware of the present moment.

Participants responded on a five-point Likert scale of *strongly disagree* to *strongly agree*. We calculated Cronbach's alpha from the Cohort 6 and Cohort 7 Year 2 data, resulting in a value of .98, indicating very high internal reliability.

Difficulty With Attention

Dr. Roeser also recommended we include two items to measure the level to which participants had difficulty concentrating to solve problems and deficits in attention due to distracting thoughts (personal communication, January 12, 2022). For both items, we asked participants to rate their level of agreement on a five-point Likert ranging from *strongly disagree* to *strongly agree*. We calculated Cronbach's alpha from the Cohort 6 and Cohort 7 Year 2 sample and found a very high internal reliability of .95 for this scale.

Momentary Positive Emotions, Engagement, Relationships, Meaning, Accomplishment

Additionally, Dr. Roeser recommended items from mPERMA (Heshmati et al., 2023). Based on the PERMA theoretical perspective of well-being, mPERMA adapted the PERMA-Profiler (Butler & Kern, 2016) and Ecological Momentary Assessment paradigm (Stone & Shiffman, 1994) to create a momentary and multidimensional measure of well-being (Heshmati et al., 2023). mPERMA consists of three scale items for each of the five elements (positive emotions, engagement, relationships, meaning, and accomplishment), resulting in 15 items total. The original scale's response options ranged from *not at all* to *extremely*, using a scale of 0 to 100. To be consistent with other scales on the well-being log, we asked participants to respond about how much they agree with each statement by measured responses on a five-point Likert scale ranging from *strongly disagree* to *strongly agree*. The Cronbach's alpha coefficients for each element reported by Heshmati et al. (2023) were .88 for positive emotions, .85 for engagement, .91 for relationship, .92 for meaning, and .92 for accomplishment, which all support good internal reliability.

Supported & Hopeful

An essential component of the CREATE residency program is providing a network of support for early career teachers. We aimed to measure participants' feelings of support in daily activities and their sense of hope by adding three researcher-developed items to WBLs. We asked participants to respond on a five-point Likert scale ranging from *strongly disagree* to *strongly agree* about how much they agreed with the following statements.

1. I am feeling stressed.
2. I am feeling supported.
3. I am feeling hopeful.

The Cronbach's alpha for this scale was .83, indicating the items of the scale are highly related to each other. We calculated the Cronbach's alpha using the Cohort 6 and Cohort 7 Year 2 sample.

Sense of Equity in Role & Experience

Through partnership with the CREATE team, we developed five items to better understand the impact on CREATE residents of the racial equity work that was occurring during TT, and if there was a noticeable difference compared to the comparison group. These items measured participants' sense of their ability to connect with students of various backgrounds, capacity to strive for equity goals, comfort in discussing equity-related issues, being valued in the school, and being respected by colleagues. Using a Likert scale ranging from *strongly disagree* to *strongly agree*, participants rated the extent to which they agreed with the following statements.

1. I am equipped to build relationships with students whose backgrounds are different from mine.
2. I have coping mechanisms that help me stay energized in serving students while striving for equity.
3. I feel comfortable discussing equity-related issues in my school.
4. I feel valued in my school.
5. I feel respected by my colleagues.

We calculated the internal reliability for this scale using the Cohort 6 and Cohort 7 Year 2 sample, resulting in a Cronbach's alpha of .83, indicating good internal reliability.

Multidimensional Assessment of Interoceptive Awareness

Through the logs, we sought to gain a deeper understanding of the enduring aspects of well-being, particularly interoceptive awareness. Interoceptive awareness, as defined by Mehling et al. (2018), is the conscious recognition by the nervous system of the body's signals in real-time. To measure this, we employed the Self-Regulation subscale of the MAIA-2. We asked participants to report how often each of 4 statements applied to their general daily life using a five-point Likert scale ranging from *never* to *always*. The Cronbach alpha for the Self-Regulation subscale was .79 (Mehling et al., 2018), demonstrating good internal consistency.

Maslach Burnout Inventory for Educators

Teachers, among others who frequently work with people, are at risk for burnout. The Maslach Burnout Inventory for Educators assesses the three components of burnout using a 22-item scale: 1) emotional exhaustion (depletion of emotional resources), 2) depersonalization (negative feelings about one's students), and 3) reduced personal accomplishment (tendency to evaluate oneself negatively with regard to work). Respondents read a series of statements and responded with how frequently they experienced the stated feelings or attitudes on a 6-point scale ranging from *never* to *everyday*. Cronbach's alpha is .90 for emotional exhaustion, .79 for depersonalization, and .71 for personal accomplishment (Maslach et al., 2018), reflecting acceptable internal reliability.

TABLE D1. SURVEY SCALES USED TO MEASURE TEACHER WELL-BEING

Scale labels	Subconstructs	Alpha	Scale range
Focus & Awareness	Mindfulness through focusing attention and being aware of the present situation	.98	5-point scale (1 = strongly disagree to 5 = strongly agree)
Difficulty With Attention	Difficulty concentrating to solve problems, and deficits in attention due to distracting thoughts	.95	
mPERMA: Positive Emotions Subscale	Sense of joy, positivity and contentment in daily activity	.88[a]	
mPERMA: Meaning Subscale	Sense of purpose, value and direction in life	.92[a]	
mPERMA: Accomplishment Subscale	Progress in accomplishing goals and in handling responsibilities	.92[a]	
Supported & Hopeful	Feeling support in daily activities, and a sense of hope	.83	
Sense of Equity in Role & Experience	Sense of ability to connect with students of various backgrounds, capacity to strive for equity goals, comfort in discussing equity-related issues, being valued in the school, being respected by colleagues	.83	
MAIA: Self-Regulation	Perceptions of having coping strategies to achieve calm in stressful and taxing situations	.79[a]	
MBI: Emotional Exhaustion	In the school context, sense of feeling drained, fatigued, burned-out and frustrated	.90[a]	7-point scale (1 = never, 2 = a few times a year or less, 3 = once a month or less, 4 = a few times a month, 5 = once a week, 6 = a few times a week, 7 = every day)
MBI: Depersonalization	In the school context, treating others impersonally, feeling hardened emotionally, feeling apathetic	.79[a]	
MBI: Personal Accomplishment	In the school context, feeling a sense of accomplishment, exhilaration in working with students, positive influence over others, understanding of students' feelings, being effective in addressing problems students experience, achieving a relaxed atmosphere with students. accomplishing worthwhile achievement in work	.71[a]	

[a] Cronbach's alphas for the mPERMA scales (Positive Emotions, Meaning, and Accomplishment) are from Heshmati et al. (2023), for the MAIA scale (Self-Regulation) are from Mehling et al. (2018), and for the MBI subscales (Emotional Exhaustion, Depersonalization, and Personal Accomplishment) are from Maslach et al. (2018). All other alpha calculations used data from the Cohort 6 and Cohort 7 Year 2 sample.

Note. MAIA = Multidimensional Assessment of Interoceptive Awareness; mPERMA = Momentary Positive Emotions, Engagement, Relationships, Meaning, Accomplishment; and MBI = Maslach Burnout Inventory for Educators.

Administrative Data

Teacher Level Data

We collected teacher-level data from GaDOE and APS, which included employment information, gender, race, ethnicity, and termination information, if applicable and available to share. Additionally, we planned to collect Teacher Assessment on Performance Standards (TAPS) ratings as a measure of teacher effectiveness. TAPS is a rubrics-based evaluation method used to measure Georgia public school teachers' performance on a set of designated performance standards. TAPS measures teacher effectiveness consistently throughout the state. There are ten performance standards that TAPS

uses to rate teachers on a scale of 0 to 3: Level 0 is Emerging, Level I is Developing, Level II is Proficient, and Level III is Advanced. Through the programming and support it offers, CREATE aims primarily to improve teacher efficacy in two of the ten performance standards measured by TAPS, both of which we planned to analyze for this project.

1. Instructional strategies (The teacher promotes student learning by using research-based instructional strategies relevant to the content area to engage students in active learning and to facilitate the students' acquisition of key knowledge and skills.)
2. Positive learning environment (The teacher provides a well-managed, safe, and orderly environment that is conducive to learning and encourages respect for all.) (GaDOE, 2020)

However, due to COVID-19 interruptions, TAPS scores were unavailable for spring 2022 and spring 2023. GaDOE did not calculate these data at all for the 2021–22 school year. Since submissions for TAPS data from districts to the state were optional for 2022–23, the state did not calculate these data for the 2022-23 school year. APS could not share the TAPS data with researchers.

Student Level Data

Student level data requested from GaDOE or school districts include gender, age, grade level, race, ethnicity, special education status, limited English proficiency status, and Georgia Milestones ELA and mathematics scores. The Georgia Milestones assesses student achievement in ELA and mathematics for students in grades 3–8, according to state-adopted content standards. The Georgia Milestones is a valid and reliable measure for student achievement in Georgia. Cronbach's alpha reliability coefficient for the Georgia Milestones ranges from .89 to .94 across all subjects, which is an adequate level of reliability for the stated goals of the assessment (GaDOE, 2019). As described in Chapter 2, at the time of publication, we were unable to obtain all of the necessary data from APS to evaluate the impact of CREATE on student achievement using the Georgia Milestones. If we obtain the remaining data and have the resources to analyze them, we will append a supplemental chapter with those findings.

Publicly Available Data on Teacher Certification and Teaching Status

Certification Data

The Educator Certification Division of the Georgia Professional Standards Commission provides a publicly available database to confirm certification status for Georgia educators (Georgia Professional Standards Commission, n.d.). The database includes certification type, level, field, and issue and validity dates. We used this database to triangulate self-reported data (if needed) or fill in missing values for teacher preparation program completion.

Teaching Status

The State of Georgia provides a publicly available database of information on state expenditures (Open Georgia, 2008). The database includes annual salaries and travel expenses for employees of Local Boards of Education, including teachers. We used this information, along with a variety of other data sources, to determine teachers' teaching status.

CREATE Program Data

We collected various program data from the CREATE program team in order to corroborate resident self-report survey data on FOI measures (and to report on FOI measures not addressed by resident survey data). Program data for residents include classroom placement rosters, TT meeting attendance, logs for mentor meetings and observation cycles, and summer internship/academy attendance. We also collected program data for experienced educators at CREATE schools participating in CREATE activities.

Appendix E. Data Collection Activity Response Rates

Table E1 through Table E6 provide response rates to the quarterly surveys and WBLs from participants in Cohort 6 Year 1 through Year 3, Cohort 7 Year 1 and Year 2, and Cohort 8 Year 1.

TABLE E1. COHORT 6 YEAR 1 SURVEY RESPONSE RATES (2020–21)

Survey	Date	Total response rate (no. responded / no. active)	CREATE group response rate	Comparison group response rate
Baseline survey	April – September 2020	46/46 (100%)	17/17 (100%)	29/29 (100%)
Quarterly Survey 1	December 2020	37/39 (95%)	16/17 (94%)	21/22 (95%)
Quarterly Survey 2	January 2021	35/39 (90%)	15/17 (88%)	20/22 (91%)
Quarterly Survey 3	March 2021	43/46 (93%)	15/17 (88%)	28/29[a] (97%)
Quarterly Survey 4	May 2021	42/46 (91%)	15/17 (88%)	27/29[a] (93%)

[a] Our research team was able to obtain research permissions for seven comparison participants in the middle of the year. We only invited these participants to complete Quarterly Survey 3 and Quarterly Survey 4.

TABLE E2. COHORT 6 YEAR 2 SURVEY AND WELL-BEING LOG RESPONSE RATES (2021–22)

Survey	Date	Total response rate (no. responded / no. active)	CREATE group response rate	Comparison group response rate
Quarterly Survey 1	December 2021	23/24 (96%)	10/10 (100%)	13/14 (93%)
Well-Being Log 1[a]	January 2022	20/22 (91%)	9/10 (90%)	11/12 (92%)
Quarterly Survey 2	January 2022	23/24 (96%)	10/10 (100%)	13/14 (93%)
Well-Being Log 2[a]	February–March 2022	20/22 (91%)	9/10 (90%)	11/12 (92%)
Quarterly Survey 3	March 2022	23/24 (96%)	10/10 (100%)	13/14 (93%)
Well-Being Log 3[a]	March–April 2022	17/21 (81%)	8/10 (80%)	9/11[b] (82%)

TABLE E2. COHORT 6 YEAR 2 SURVEY AND WELL-BEING LOG RESPONSE RATES (2021–22)

Survey	Date	Total response rate (no. responded / no. active)	CREATE group response rate	Comparison group response rate
Well-Being Log 4[a]	April–May 2022	22/23 (96%)	9/10 (90%)	13/13[c] (100%)
Quarterly Survey 4	May 2022	23/24 (96%)	10/10 (100%)	13/14 (93%)

[a] In Year 2, we asked participants if they would like to participate in the well-being log aspect of the study. Only 22 out of the 24 participants agreed to complete well-being logs.

[b] Well-Being Log 3 didn't include one comparison participant due to their district's "no survey" window.

[c] For Well-Being Log 4, one comparison participant agreed to participate in this aspect of the study in April. Because of this timing, we only invited that participant to complete Well-Being Log 4.

TABLE E3. COHORT 6 YEAR 3 SURVEY AND WELL-BEING LOG RESPONSE RATES (2022–23)

Survey	Date	Total response rate (no. responded / no. active)	CREATE group response rate	Comparison group response rate
Well-Being Log 1	October 2022	13/17 (76%)	6/8 (75%)	7/9 (78%)
Quarterly Survey 1	November 2022	17/17 (100%)	8/8 (100%)	9/9 (100%)
Well-Being Log 2	January 2023	13/17 (76%)	6/8 (75%)	7/9 (78%)
Quarterly Survey 2	January 2023	17/17 (100%)	8/8 (100%)	9/9 (100%)
Well-Being Log 3	February 2023	15/17 (88%)	7/8 (88%)	8/9 (89%)
Well-Being Log 4	March 2023	14/17 (82%)	7/8 (88%)	7/9 (78%)
Quarterly Survey 3	March 2023	17/17 (100%)	8/8 (100%)	9/9 (100%)
Quarterly Survey 4	May 2023	17/17 (100%)	8/8 (100%)	9/9 (100%)

TABLE E4. COHORT 7 YEAR 1 SURVEY AND WELL-BEING LOG RESPONSE RATES (2021–22)

Survey	Date	Total response rate (no. responded / no. active)	CREATE group response rate [a]	Comparison group response rate
Baseline survey	April–August 2021	37/37 (100%)	23/23 (100%)	14/14 (100%)
Quarterly Survey 1	December 2021	31/33 (94%)	23/23 (100%)	8/10 (80%)
Well-Being Log 1	January 2022	29/33 (88%)	22/23 (96%)	7/10 (70%)
Quarterly Survey 2	January 2022	31/33 (94%)	23/23 (100%)	8/10 (80%)
Well-Being Log 2	February 2022	30/33 (91%)	23/23 (100%)	7/10 (70%)
Well-Being Log 3	March 2022	30/33 (91%)	22/23 (96%)	8/10 (80%)
Quarterly Survey 3	March 2022	35/37[a] (95%)	23/23 (100%)	12/14[b] (86%)
Well-Being Log 4	April–May 2022	34/37 (92%)	23/23 (100%)	11/14 (79%)
Quarterly Survey 4	May 2022	34/37 (92%)	23/23 (100%)	11/14 (79%)

[a] Four participants out of 23 in the CREATE group were part of CREATE Pathway 1.

[b] Four comparison participants entered into the sample in spring 2022. They are included in the baseline sample for Cohort 7, but did not start data collection until March 2022.

TABLE E5. COHORT 7 YEAR 2 SURVEY AND WELL-BEING LOG RESPONSE RATES (2022–23)

Survey	Date	Total response rate (no. responded / no. active)	CREATE group response rate[a]	Comparison group response rate
Well-Being Log 1	October 2022	14/17 (82%)	8/11 (73%)	6/6 (100%)
Quarterly Survey 1	November 2022	17/17 (100%)	11/11 (100%)	6/6 (100%)

TABLE E5. COHORT 7 YEAR 2 SURVEY AND WELL-BEING LOG RESPONSE RATES (2022–23)

Survey	Date	Total response rate (no. responded / no. active)	CREATE group response rate[a]	Comparison group response rate
Well-Being Log 2	January 2023	15/17 (88%)	9/11 (82%)	6/6 (100%)
Quarterly Survey 2	January 2023	17/17 (100%)	11/11 (100%)	6/6 (100%)
Well-Being Log 3	February 2023	14/17 (82%)	8/11 (73%)	6/6 (100%)
Well-Being Log 4	March 2023	16/17 (94%)	10/11 (91%)	6/6 (100%)
Quarterly Survey 3	March 2023	16/17 (94%)	10/11 (91%)	6/6 (100%)
Quarterly Survey 4	May 2023	16/17 (94%)	10/11 (91%)	6/6 (100%)

[a] Two participants out of 11 in the CREATE group were part of CREATE Pathway 1.

TABLE E6. COHORT 8 YEAR 1 SURVEY AND WELL-BEING LOG RESPONSE RATES (2022–23)

Survey	Date	Total response rate (no. responded / no. active)	CREATE group response rate	Comparison group response rate
Baseline survey	April–November 2022	27/27 (100%)	13/13 (100%)	14/14 (100%)
Well-Being Log 1	October 2022	24/27 (89%)	13/13 (100%)	11/14 (79%)
Quarterly Survey 1	November 2022	27/27 (100%)	13/13 (100%)	14/14 (100%)
Well-Being Log 2	January 2023	26/27 (96%)	13/13 (100%)	13/14 (93%)
Quarterly Survey 2	January 2023	26/27 (96%)	12/13 (92%)	14/14 (100%)
Well-Being Log 3	February 2023	26/27 (96%)	12/13 (92%)	14/14 (100%)
Quarterly Survey 3	March 2023	25/27 (93%)	12/13 (92%)	13/14 (93%)

TABLE E6. COHORT 8 YEAR 1 SURVEY AND WELL-BEING LOG RESPONSE RATES (2022–23)

Survey	Date	Total response rate (no. responded / no. active)	CREATE group response rate	Comparison group response rate
Well-Being Log 4	March 2023	25/27 (93%)	12/13 (92%)	13/14 (93%)
Quarterly Survey 4	May 2023	25/27 (93%)	12/13 (92%)	13/14 (93%)

Appendix F. Fidelity of Implementation Matrix

Table F1 through Table F6 present the Fidelity of Implementation Matrix for each key component.

TABLE F1. KEY COMPONENT 1: PROGRESSIVE CORE CLASSROOM ROLES

Indicators	Definition	Unit of implementation	Data sources	Data collection (who, when)	Score for levels of implementation at unit level	Threshold for adequate implementation at unit level	Score and threshold for adequate implementation at sample level	Expected years of fidelity measurement
Indicator 1: Y1	(Y1) Resident is placed in the same building as another Y1 resident for fall semester	Resident	CREATE program rosters, Resident surveys	Rosters emailed to evaluators from CREATE; Surveys administered quarterly by evaluators	0 = Resident is not placed in the same building as another Y1 resident; 1 = Resident is placed in the same building as another Y1 resident for the fall semester	Adequate implementation at resident level = score of 1		Cohort 6: 2020-2021 Cohort 7: 2021-2022 Cohort 8: 2022-2023
Indicator 2: Y1	(Y1) Resident is placed in classroom of experienced educator trained in mentoring	Resident	CREATE program rosters, Resident surveys	Rosters emailed to evaluators from CREATE; Surveys administered quarterly by evaluators	0 = Resident is not placed with an experienced educator trained in mentoring; 1 = Resident is placed with an experienced educator trained in mentoring	Adequate implementation at resident level = score of 1		Cohort 6: 2020-2021 Cohort 7: 2021-2022 Cohort 8: 2022-2023
Indicator 3: Y1	(Y1) Resident is placed in a CREATE partner school	Resident	CREATE program rosters, Resident surveys	Rosters emailed to evaluators from CREATE; Surveys administered quarterly by evaluators	0 = Resident is placed in a non-CREATE partner school; 1 = Resident is placed in a CREATE partner school	Adequate implementation at resident level = score of 1		Cohort 6: 2020-2021 Cohort 7: 2021-2022 Cohort 8: 2022-2023
Indicator 1: Y2	(Y2) Resident is in a paired teaching dyad with another Y2 resident or an experienced educator, as a teacher-of-record	Resident	CREATE program rosters, Resident surveys	Rosters emailed to evaluators from CREATE; Surveys administered quarterly by evaluators	0 = Resident is not paired with co-resident or experienced educator; 1 = Resident is paired with co-resident or experienced educator	Adequate implementation at resident level = score of 1		Cohort 6: 2021-2022 Cohort 7: 2022-2023

TABLE F1. KEY COMPONENT 1: PROGRESSIVE CORE CLASSROOM ROLES

Indicators	Definition	Unit of implementation	Data sources	Data collection (who, when)	Score for levels of implementation at unit level	Threshold for adequate implementation at unit level	Score and threshold for adequate implementation at sample level	Expected years of fidelity measurement
Indicator 2: Y2	(Y2) Resident is teaching in subject and grade level for which resident is certified	Resident	CREATE program rosters, Resident surveys	Rosters emailed to evaluators from CREATE; Surveys administered quarterly by evaluators	0 = Resident is not teaching in subject and grade level for which resident is certified; 1 = Resident is teaching in subject and grade level for which resident is certified	Adequate implementation at resident level = score of 1		Cohort 6: 2021-2022 Cohort 7: 2022-2023
Indicator 3: Y2	(Y2) Resident is teaching at a CREATE partner school	Resident	CREATE program rosters, Resident surveys	Rosters emailed to evaluators from CREATE; Surveys administered quarterly by evaluators	0 = Resident is not teaching at a CREATE partner school; 1 = Resident is teaching at a CREATE partner school	Adequate implementation at resident level = score of 1		Cohort 6: 2021-2022 Cohort 7: 2022-2023
Indicator 1: Y3	(Y3) Resident is teaching solo as teacher-of-record	Resident	CREATE program rosters, Resident surveys	Rosters emailed to evaluators from CREATE; Surveys administered quarterly by evaluators	0 = Resident is not teaching solo as teacher-of-record; 1 = Resident is teaching solo as teacher-of-record	Adequate implementation at resident level = score of 1		Cohort 6: 2022-2023
Indicator 2: Y3	(Y3) Resident is teaching in subject and grade level for which resident is certified	Resident	CREATE program rosters, Resident surveys	Rosters emailed to evaluators from CREATE; Surveys administered quarterly by evaluators	0 = Resident is not teaching in subject area and grade level for which resident is certified; 1 = Resident is teaching in subject and grade level for which resident is certified	Adequate implementation at resident level = score of 1		Cohort 6: 2022-2023

TABLE F1. KEY COMPONENT 1: PROGRESSIVE CORE CLASSROOM ROLES

Indicators	Definition	Unit of implementation	Data sources	Data collection (who, when)	Score for levels of implementation at unit level	Threshold for adequate implementation at unit level	Score and threshold for adequate implementation at sample level	Expected years of fidelity measurement
Indicator 3: Y3	(Y3) Resident is teaching at a CREATE partner school	Resident	CREATE program rosters, Resident surveys	Rosters emailed to evaluators from CREATE; Surveys administered quarterly by evaluators	0 = Resident is not teaching at a CREATE partner school; 1= Resident is teaching at a CREATE partner school	Adequate implementation at resident level = score of 1		Cohort 6: 2022-2023
All indicators					Resident-level component score: Number of indicators implemented with fidelity per year. Possible range (per year) = 0 - 3	Threshold = 2 (Resident meets fidelity on at least 2 indicators)	Component score ranges from 0-1: Fidelity not met: Y1: Less than 90% of residents meet fidelity on 2+ indicators Y2: Less than 80% of residents meet fidelity on 2+ indicators Y3: Less than 85% of residents meet fidelity on 2+ indicators Fidelity met = Y1: 90% or more of residents meet fidelity on 2+ indicators; Y2: 80% or more of residents meet fidelity on 2+ indicators; Y3: 85% or more of residents meet fidelity on 2+ indicators	Cohort 6: 2020-2021 (Y1), 2021-2022 (Y2), 2022-2023 (Y3); Cohort 7: 2021-2022 (Y1), 2022-2023 (Y2); Cohort 8: 2022-2023 (Y1)

Note. Y1 = Year 1 of the residency program, Y2 = Year 2 of the residency program, and Y3 = Year 3 of the residency program.

TABLE F2. KEY COMPONENT 2: RECRUITMENT AND TRAINING OF COOPERATING TEACHERS

Indicators	Definition	Unit of implementation	Data sources	Data collection (who, when)	Score for levels of implementation at unit level	Threshold for adequate implementation at unit level	Score and threshold for adequate implementation at sample level	Expected years of fidelity measurement
Indicator 1	Experienced educators are recruited to attend CREATE CT training and matched with CREATE residents	CT	CREATE recruitment records	Records emailed to evaluators from CREATE	2021-2022: 0 = CREATE does not recruit 12 CTs 1 = CREATE recruits 12 CTs 2022-2023: 0 = CREATE does not recruit 24 CTs 1 = CREATE recruits 24 CTs	Adequate implementation = score of 1	0 = Fidelity not met at unit level 1 = Fidelity met at unit level	2021-2022, 2022-2023
Indicator 2	CTs attend 2 days of summer CT training	CT	CREATE attendance rosters Participant surveys	Rosters emailed to evaluators from CREATE Surveys administered quarterly by evaluators	0 = CT attends less than 2 days of summer training 1= CT attends 2 days of summer training	Adequate implementation = score of 1	0 = Less than 90% of CTs meet fidelity at the unit level 1 = 90% or more of CTs meet fidelity at the unit level	2021-2022, 2022-2023
Indicator 3	CTs attend 3 days of CT training during the mentoring year	CT	CREATE attendance rosters Participant surveys	Rosters emailed to evaluators from CREATE Surveys administered quarterly by evaluators	0 = CT attends less than 3 days of CT training during the mentoring year 1 = CT attends 3 days of CT training during the mentoring year	Adequate implementation = score of 1	0 = Less than 90% of CTs meet fidelity at the unit level 1 = 90% or more of CTs meet fidelity at the unit level	2021-22, 2022-23
All indicators							Component score ranges from 0–3: 0 = Fidelity not met: Less than 3 indicators met 1= Fidelity Met: 3 indicators met	2021-22, 2022-23

Note. CT = Cooperating Teacher.

TABLE F3. KEY COMPONENT 3: MULTIPLE FORMS OF MENTORING (Y2 & Y3 ONLY)

Indicators	Definition	Unit of implementation	Data sources	Data collection (who, when)	Score for levels of implementation at unit level	Threshold for adequate implementation at unit level	Score and threshold for adequate implementation at sample level	Expected years of fidelity measurement
Indicator 1	Residents have mentors (SBM) who attended training prior to mentoring	Mentor	Rosters from CREATE	Rosters emailed to evaluators from CREATE	0 = Resident has a mentor who did not attend training 1 = Resident has a mentor who attended training	Adequate implementation = score of 1	0 = Less than 90% of residents have mentors that meet fidelity at the unit level 1 = 90% or greater of residents have mentors that meet fidelity at the unit level	Cohort 6: 2021-2022, 2022-2023 Cohort 7: 2022-2023
Indicator 2	Residents have mentors who attended training during their mentoring year (2 meetings)	Mentor	Rosters from CREATE	Rosters emailed to evaluators from CREATE	0 = Resident has a mentor who attends 0–1 meetings 1 = Resident has a mentor who attends 2 or more meetings	Adequate implementation = score of 1	0 = Less than 90% of residents have mentors that meet fidelity at the unit level 1 = 90% or more of mentors have residents that meet fidelity at the unit level	Cohort 6: 2021-2022, 2022-2023 Cohort 7: 2022-2023
Indicator 3	Residents attend semi–monthly meetings with their mentors (SBM or IM)	Resident	Resident surveys, CREATE mentor logs (and spreadsheet)	Mentor logs emailed to evaluators from CREATE, Surveys administered quarterly by evaluators	Y2 Residents: 0 = Resident attends 0–20 meetings; 1 = Resident attends 20–24 meetings; 2 = Resident attends 25 or more meetings Y3 Residents: 0 = Resident attends 0–8 meetings; 1 = Resident attends 9–11 meetings; 2 = Resident attends 12 or more meetings	Adequate implementation = score of 2 in Y2 and Y3	0 = Less than 90% of residents earn a score of 2 at the unit level *or* anyone earns a score of 0 1 = 90% of residents earn a score of 2 at the unit level *and* no residents earn a score of 0	Cohort 6: 2021-2022, 2022-2023 Cohort 7: 2022-2023

TABLE F3. KEY COMPONENT 3: MULTIPLE FORMS OF MENTORING (Y2 & Y3 ONLY)

Indicators	Definition	Unit of implementation	Data sources	Data collection (who, when)	Score for levels of implementation at unit level	Threshold for adequate implementation at unit level	Score and threshold for adequate implementation at sample level	Expected years of fidelity measurement
Indicator 4	Resident participates in mentor-resident observation cycles with their CREATE IM	Resident	Resident surveys CREATE mentor logs (and spreadsheet)		0 = Resident participates in 0–1 cycles 1 = Resident participates in 2–3 cycles	Adequate implementation = score of 1	Y2: 0 = Less than 90% of residents meet fidelity at the unit level 1 = 90% or more of residents meet fidelity at the unit level Y3: 0 = Less than 80% of residents meet fidelity at the unit level 1 = 80% or more of residents meet fidelity at the unit level	Cohort 6: 2021-2022, 2022-2023 Cohort 7: 2022-2023
All indicators							Sum of indicator scores Component score ranges from 0–4: Fidelity not met = score of 0–3 Fidelity met = score of 4	Cohort 6: 2021-2022, 2022-2023 Cohort 7: 2022-2023

Note. SBM = School-Based Mentor, IM = Instructional Mentor, Y2 = Year 2 of the residency program, and Y3 = Year 3 of the residency program.

TABLE F4. KEY COMPONENT 4: SUMMER RESIDENT ACADEMY

Indicators	Definition	Unit of implementation	Data sources	Data collection (who, when)	Score for levels of implementation at unit level	Threshold for adequate implementation at unit level	Score and threshold for adequate implementation at sample level	Expected years of fidelity measurement
Indicator 1 (calculated in Y2)	Residents attend Summer Resident Academy between Y1 and Y2	Resident	CREATE attendance rosters Resident surveys	Rosters emailed to evaluators from CREATE Surveys administered quarterly by evaluators	0 = Resident attends less than 85% of Summer Resident Academy meetings 1= Resident attends 85% or more of Summer Resident Academy meetings	Adequate implementation = score of 1		Cohort 6: Summer 2021 Cohort 7: Summer 2022
Indicator 2 (calculated in Y3)	Residents attend Summer Resident Academy between Y2 and Y3	Resident	CREATE attendance rosters Resident surveys	Rosters emailed to evaluators from CREATE Surveys administered quarterly by evaluators	0 = Resident attends less than 85% of Summer Resident Academy meetings 1= Resident attends 85% or more of Summer Resident Academy meetings	Adequate implementation = score of 1		Cohort 6: Summer 2022
All indicators							Component score ranges from 0–1: 0 = Fidelity not met: Less than 80% of residents earn a score of 1 1= Fidelity Met: 80% or more of residents earn a score of 1	Cohort 6: Summer 2021, Summer 2022 Cohort 7: Summer 2022

Note. Y1 = Year 1 of the residency program, Y2 = Year 2 of the residency program, and Y3 = Year 3 of the residency program.

TABLE F5. KEY COMPONENT 5: EQUITY-CENTERED PROFESSIONAL LEARNING

Indicators	Definition	Unit of implementation	Data sources	Data collection (who, when)	Score for levels of implementation at unit level	Threshold for adequate implementation at unit level	Score and threshold for adequate implementation at sample level	Expected years of fidelity measurement
Indicator 1	CREATE administrators host 2 or more options for experienced educators to attend equity-centered PL each year.	Program	CREATE attendance rosters	Rosters emailed to evaluators from CREATE	0 = CREATE administrators host 0-1 institute; 1= CREATE administrators host 2+ institutes	Adequate implementation at program level = score of 1	0 = Fidelity not met at unit level 1 = Fidelity met at unit level	2020-2021, 2021-2022, 2022-2023
Indicator 2	Experienced educators attend equity-centered PL	Non-resident educators from CREATE partner schools	CREATE attendance rosters	Rosters emailed to evaluators from CREATE	0 = Educator attends 50% or fewer of days; 1 = Educator attends greater than 50% of days	Adequate implementation = score of 1	0 = Of the experienced educators who attend, less than 85% meet fidelity at the unit level 1 = Of the experienced educators who attend, 85% or more meet fidelity at the unit level	2020-2021, 2021-2022, 2022-2023
Indicator 3	Residents attend a CREATE Professional Learning opportunity (other than Together Time meetings) by the end of Y3 of their program. Residents may attend the CREATE PL in Y2 or Y3.	Residents	CREATE attendance rosters Resident surveys	Rosters emailed to evaluators from CREATE Surveys administered quarterly by evaluators	0 = Resident attends 50% or fewer of PL days 1 = Resident attends more than 50% of PL days	Adequate implementation = score of 1	0 = Less than 30% of residents meet fidelity at the unit level 1 = 30% or more of residents meet fidelity at the unit level	Cohort 6: Summer 2022; Summer 2023

TABLE F5. KEY COMPONENT 5: EQUITY-CENTERED PROFESSIONAL LEARNING

Indicators	Definition	Unit of implementation	Data sources	Data collection (who, when)	Score for levels of implementation at unit level	Threshold for adequate implementation at unit level	Score and threshold for adequate implementation at sample level	Expected years of fidelity measurement
Indicator 4	CREATE residents attend monthly Together Time meetings that include training in equity-centered practices	Residents	CREATE attendance rosters Resident surveys	Rosters emailed to evaluators from CREATE Surveys administered quarterly by evaluators	Y1 residents: 0 = Attend 0–7 meetings; 1 = Attend 8–10 meetings Y2 residents: 0 = Attend 0–5 meetings, 1 = Attend 6+ meetings Y3 residents: 0 = Attend 0–4 meetings; 1 = Attend 5+ meetings	Adequate implementation = score of 1	0 = Less than 95% of residents meet fidelity 1 = 95% or more of residents meet fidelity at the unit level	Cohort 6: 2020-2021, 2021-2022, 2022-2023 Cohort 7: 2021-2022, 2022-2023 Cohort 8: 2022-2023
All indicators							Y1 and Y2 (indicator 3 not measured) 0 = Fidelity not met= Fidelity not met for indicator 1, or fidelity met for Indicator 1 but no other indicator 1 = Fidelity met = Fidelity was met for Indicator 1 and at least one other indicator Y3 0 = Fidelity not met: fidelity not met for indicators 1 or 3, OR fidelity met for indicators 1 and 3 but not for Indicators 2 and 4 1 = Fidelity met: Fidelity was met for Indicator 1, Indicator 3 and at least one other indicator	Cohort 6: 2020-2021 (Y1), 2021–2022 (Y2), 2022–2023 (Y3) Cohort 7: 2021–2022 (Y1), 2022–2023 (Y2) Cohort 8: 2022–2023 (Y1)

Note. PL = professional learning, Y1 = Year 1 of the residency program, Y2 = Year 2 of the residency program, and Y3 = Year 3 of the residency program.

TABLE F6. KEY COMPONENT 6: COMPASSION-BASED PROFESSIONAL LEARNING

Indicators	Definition	Unit of implementation	Data sources	Data collection (who, when)	Score for levels of implementation at unit level	Threshold for adequate implementation at unit level	Score and threshold for adequate implementation at sample level	Expected years of fidelity measurement
Indicator 1	Program administrators offer at least one compassion-based PL per year for general population of teachers at CREATE partner schools (and other educators)	Program	Email communication with CREATE administrator	Annual email communication with CREATE administrators	0 = CREATE administrators offer 0 courses 1 = CREATE administrators offer 1 or more courses	Adequate implementation = score of 1	0 = Fidelity not met at unit level 1 = Fidelity met at unit level	2020–2021, 2021–2022, 2022-2023
Indicator 2	Residents attend monthly Together Time meetings that include training on compassion-based practices	Resident	CREATE attendance rosters Resident surveys	Rosters emailed to evaluators from CREATE Surveys administered quarterly by evaluators	Y1 residents: 0 = Attend 0–7 meetings; 1 = Attend 8–10 meetings Y2 residents: 0 = Attend 0–5 meetings, 1 = Attend 6+ meetings Y3 residents: 0 = Attend 0–4 meetings; 1 = Attend 5+ meetings	Adequate implementation = score of 1	0 = Less than 95% of residents meet fidelity at the unit level 1 = 95% or more of residents meet fidelity at the unit level	Cohort 6: 2020-2021, 2021-2022, 2022-2023 Cohort 7: 2021-2022, 2022-2023 Cohort 8: 2022-2023
All indicators							Component score is sum of indicator scores. Range is 0–2 Fidelity not met: score of 0–1 Fidelity met: score of 2	Cohort 6: 2020-2021, 2021-2022, 2022-2023 Cohort 7: 2021-2022, 2022-2023 Cohort 8: 2022-2023

Note. PL = professional learning, Y1 = Year 1 of the residency program, Y2 = Year 2 of the residency program, and Y3 = Year 3 of the residency program.

Appendix G. Fidelity of Implementation Results

In each section below, we provide a detailed description of each of CREATE's six key components. Then, we provide the Fidelity of Implementation (FOI) results for Cohort 6 Year 1 through Year 3, Cohort 7 Year 1 and Year 2, and Cohort 8 Year 1.

Key Component 1: Progressive Core Classroom Roles

As CREATE residents move through the three-year residency model, their role within the classroom changes. The progressive core classroom roles provide supports for residents while also providing space for increased autonomy, agency, and independence each year.

Residents enter the CREATE residency during the final year of their GSU CEHD teacher certification program. In this first year, Year 1 residents are placed in a classroom with a highly-skilled veteran teacher—known as a CT—who serves as their mentor at a CREATE partner school. CREATE aims to have multiple CREATE residents placed in each school.

Upon graduating from the GSU CEHD teacher certification program, residents in Year 2 continue on at a CREATE partner school, paired with each other as co-lead teachers-of-record. A small proportion of Year 2 residents were paired with experienced educators who have attended CREATE trainings. The lighter load and flexibility of having two teachers in one classroom allows more time for mentor-resident reflection and observations of other teachers. This arrangement is also intended to address the sense of overwhelm with new responsibilities that new teachers often cite as the reason they leave teaching. In addition, each resident should be teaching in the subject and grade level for which they are certified. CREATE recognizes that some teachers prefer to work as the sole teacher-of-record, and that some educators work more successfully when not sharing a classroom. Because CREATE has a close working relationship with each of their residents, CREATE administrators consider the preferences of residents when determining whether working with a co-teacher is the best choice for each resident. We have taken this into account in our calculation of FOI for the associated indicator.

In Year 3, all residents become lead teachers in their own classrooms after having had the aforementioned scaffolded teaching experiences during the previous two years. The CREATE program expects residents to teach at a CREATE partner school in the subject and grade for which they are certified.

We present data on the following indicators related to progressive core classroom roles.

- Indicator 1
    - Year 1: Residents are placed in the same building as another Year 1 resident.
    - Year 2: Residents co-teach with another Year 2 resident or experienced educator, and serve as a teacher-of-record.
    - Year 3: Residents teach solo as a teacher-of-record.
- Indicator 2
    - Year 1: Residents are placed with a CREATE-trained CT.
    - Year 2 and Year 3: Residents teach in their certified subject(s) and grade level(s).
- Indicator 3

- Year 1: Residents student teach in a CREATE partner school.
- Year 2 and Year 3: Residents are employed in a CREATE partner school.

The following tables of results display data from left to right. Column 1 describes the indicators of each CREATE program component that should be implemented with fidelity. Column 2 indicates the threshold that CREATE has set in order to meet fidelity on the respective indicator. The third column indicates how many residents in Cohort 6 met the threshold for fidelity. Column 4 indicates how many residents in Cohort 7 met the threshold for fidelity. Column 5 indicates how many residents in Cohort 8 met the threshold for fidelity. Column 6 indicates the combined score of each cohort for the corresponding indicator. The bottom row of each table is the final result, showing whether or not CREATE met FOI for each component at the program level.

Results for Key Component 1 in Year 1

CREATE administrators expressed that they intended to have 90% or more of residents meet fidelity on at least two of three indicators listed above in Year 1. The records from CREATE attendance rosters and survey data show that 95% of residents in Cohort 6 and 100% of residents in Cohort 7 and Cohort 8 met fidelity for this component in Year 1; all cohorts met fidelity for this component at the program level. More details about the number of residents who met each specific indicator are in the Table G1 below.

TABLE G1. DETAILS OF COMPONENT 1 IN YEAR 1

Indicator	Fidelity threshold	Cohort 6 met fidelity?	Cohort 7 met fidelity?	Cohort 8 met fidelity?	Combined score
1: Residents are placed in the same building as another Y1 resident	0 = Resident is not placed in the same building as another Y1 resident 1 = Resident is placed in the same building as another Y1 resident	17/21 (81%)	26/28 (93%)	20/22 (91%)	63/71 (89%)
2: Residents are placed with a CREATE-trained Cooperating Teacher	0 = Resident is not placed with CREATE-trained Cooperating Teacher 1 = Resident is placed with a Cooperating Teacher	20/21 (95%)	28/28 (100%)	22/22 (100%)	70/71 (99%)
3: Residents student teach in a CREATE partner school	0 = Resident is student teaching in a non-CREATE partner school 1 = Resident is student teaching in a CREATE partner school	21/21 (100%)	28/28 (100%)	22/22 (100%)	71/71 (100%)
Program level	90% or more of residents meet fidelity on 2+ indicators	20/21 (95%) Fidelity Met	28/28 (100%) Fidelity Met	22/22 (100%) Fidelity Met	70/71 (99%) Fidelity Met

Note. Y1 = Year 1 of the residency program, Y2 = Year 2 of the residency program, and Y3 = Year 3 of the residency program. Green = met program-level fidelity threshold.

Results for Key Component 1 in Year 2

CREATE administrators intended to have 80% or more of residents meet fidelity on at least two of the three indicators of Component 1 in Year 2. The CREATE attendance rosters and survey data show that 100% of Cohort 6 and Cohort 7 residents met fidelity in Year 2; all cohorts met fidelity for this component overall. More details about the number of residents who met each specific indicator are in the Table G2 below.

TABLE G2. DETAILS OF COMPONENT 1 IN YEAR 2

Indicator	Fidelity threshold	Cohort 6 met fidelity?	Cohort 7 met fidelity?	Combined score
1: Residents co-teach with another Y2 resident or CREATE-trained experienced educator, as a teacher-of-record	0 = Resident is not co-teaching as a teacher-of-record 1 = Resident is co-teaching as a teacher-of-record	15/17 (88%)	8/11 (73%)	23/28 (82%)
2: Residents teach in their certified subject(s) and grade level(s)	0 = Resident is not teaching in their certified subject(s) and grade level(s) 1 = Resident is teaching in their certified subject(s) and grade level(s)	16/17 (94%)	11/11 (100%)	27/28 (96%)
3: Residents are employed in a CREATE partner school	0 = Resident is not employed in a CREATE partner school 1 = Resident is employed in a CREATE partner school	17/17 (100%)	11/11 (100%)	28/28 (100%)
Program level	80% or more of residents meet fidelity on 2+ indicators	17/17 (100%) Fidelity Met	11/11 (100%) Fidelity Met	28/28 (100%) Fidelity Met

Note. Y2 = Year 2 of the residency program. Green = met program-level fidelity threshold.

Results for Key Component 1 in Year 3

CREATE administrators expressed that they intended to have 85% or more of residents meet fidelity on at least two of the three indicators in Component 1 in Year 3. The records from CREATE attendance rosters and survey data show that 100% of Cohort 6 residents met fidelity in Year 3. More details about the number of residents who met each specific indicator are in the Table G3 below.

TABLE G3. DETAILS OF COMPONENT 1 IN YEAR 3

Indicator	Fidelity threshold	Cohort 6 met fidelity?
1: Year 3: Residents teach as the solo teacher-of-record	0 = Resident is not teaching as the solo teacher-of-record 1 = Resident is teaching as the solo teacher-of-record	12/12 (100%)
2: Residents teach in their certified subject(s) and grade level(s)	0 = Resident is not teaching in their certified subject(s) and grade level(s) 1 = Resident is teaching in their certified subject(s) and grade level(s)	11/12 (92%)
3: Residents are employed at a CREATE partner school	0 = Resident is not employed at a CREATE partner school 1 = Resident is employed at a CREATE partner school	10/12 (83%)
Program level	85% or more of residents meet fidelity on 2+ indicators	12/12 (100%) Fidelity Met

Note. Green = met program-level fidelity threshold.

Overall Results for Component 1

The CREATE program met fidelity on Key Component 1 for Cohorts 6, 7, and 8 in all three years of their participation. We can, therefore, conclude that CREATE successfully implemented the core progressive classroom roles from the 2020–23 school years for Cohort 6 residents, the 2021–23 school years for Cohort 7 residents, and the 2022–23 school year for Cohort 8 residents.

Key Component 2: Recruitment and Training of Cooperating Teachers

CREATE trains experienced educators working in CREATE partner schools to serve as CTs, who host a Year 1 resident as a student teacher in their classroom. Component 2, which was added to FOI in the 2021–22 school year and measured only in Year 1, looks at the recruitment and training of CTs. We measured Component 2 for Cohort 7 and Cohort 8, but not for Cohort 6 (because Cohort 6 completed Year 1 before the introduction of this component).

For Cohorts 7 and 8 combined, the CREATE team met fidelity for Indicator 1, which measures how many experienced educators they recruited as CTs and matched with CREATE Year 1 residents. They also met fidelity for Indicator 2 (for Cohorts 7 and 8 combined), which aims to ensure CTs undergo training in mentorship prior to beginning the school year during which they mentor residents. To meet fidelity, CTs must attend two days of mentor training in the summer before the school year begins. Some CTs fulfilled this requirement through attendance at CREATE mentor trainings in prior years because they were returning mentors who were previously matched with a Year 1 resident. At the program level for Indicator 3, which measures CT attendance at CREATE mentor training during the mentoring year, CREATE did not

meet fidelity. In both the 2021–22 and 2022–23 academic years, CREATE offered seven trainings to CTs in the form of Mentor Teacher Workshops and Coaching Calls. As hiring cycles and processes at school sites can take time due to recruitment and administrative requirements, it is possible that some CTs did not join or match with student teachers until after the summer CT trainings occurred. CTs that matched with residents were to attend at least three days of CT training during the school year, which could be done through a combination of Mentor Teacher Workshops and Coaching Calls. We found that 79% of CTs of Cohort 7 residents met this threshold, and 100% of CTs of Cohort 8 residents met this threshold. The target for meeting fidelity at the program level is that 90% or more of CTs meet fidelity at the unit level. We only included training attendance during the measured school years (2021–22 and 2022–23) in score calculations for this indicator.

- Indicator 1: Recruit new CTs each academic year
- Indicator 2: CTs attend 2+ days of summer training
- Indicator 3: CTs attend 3+ days of training during the school year

TABLE G4. DETAILS OF COMPONENT 2 IN YEAR 1

Indicator	Fidelity threshold	Cohort 7 (2021–22) met fidelity?	Cohort 8 (2022–23) met fidelity?	Combined score
1: Recruit new CTs each academic year	2021–22: 0 = CREATE does not recruit 12 CTs 1 = CREATE recruits 12 CTs 2022–23: 0 = CREATE does not recruit 24 CTs 1 = CREATE recruits 24 CTs	CREATE recruited 37 CTs Fidelity met	CREATE recruited 29 CTs Fidelity met	Fidelity met
2: CTs attend 2+ days of summer training	0 = CT attends less than 2 days of summer training 1 = CT attends 2 days of summer CT training 90% or more CTs meet fidelity	26/28 (93%) of CTs met fidelity at the unit level Fidelity met	22/22 (100%) of CTs met fidelity at the unit level Fidelity met	48/50 (96%) met fidelity Fidelity met
3: CTs attend 3+ days of training during the school year	0 = CT attends less than 3 days of training during the school year 1 = CT attends 3 days of CT training during the school year 90% or more CTs meet fidelity	22/28 (79%) of CTs met fidelity at the unit level Fidelity not met	22/22 (100%) of CTs met fidelity at the unit level Fidelity met	44/50 (88%) met fidelity Fidelity not met
Program level	Component score ranges from 0-3: Fidelity met = 3 indicators met	Fidelity not met	Fidelity met	Fidelity not met

Note. CT = Cooperating Teacher. Red = did NOT meet program-level fidelity threshold.

Overall Results of Key Component 2

Cohort 7 did not meet fidelity thresholds at the program level for this component, while Cohort 8 did meet fidelity for Component 2. Cohort 7 failed to meet fidelity thresholds for Indicator 3, which measures mentoring trainings CTs attended during the mentoring year. This may have been due to accommodations regarding training requirements for those who have served as mentors with CREATE before, or to CREATE's renewed focus on the CT aspect of the program beginning this year. Scores listed in **Error! Reference source not found.** count Coaching Calls as CT trainings. Cohort 8 met fidelity for all indicators for Component 6, therefore meeting fidelity at the program level.

Key Component 3: Multiple Forms of Mentoring

CREATE equips experienced educators with the skills needed to guide CREATE residents through their first two years as full-time teachers. SBMs—who are veteran educators working at the same school as their CREATE resident mentee—receive training from CREATE in how to best coach their residents in professional practice techniques, the development of a growth mindset, and skills for maintaining their well-being during the high stress situations that occur in the early days of a teacher's career. The training offers several meetings beginning the summer prior to the start of the school and continuing throughout the year. Residents also receive IMs. IMs are members of the CREATE staff who are trained in supporting residents during their development of compassion-based, equitable, effective teaching practices. Component 3 only applies to years 2 and 3, the years in which residents are paired with SBMs and IMs. CREATE residents do not receive SBMs in Year 1, as they are still student teaching through GSU CEHD and paired with CTs, in whose classrooms they are placed.

CREATE Key Component 3 covers the multiple forms of mentoring offered to CREATE residents. We present data on the following indicators related to mentorship.

- Indicator 1 (Years 2 and 3): SBMs attend 1+ training before mentoring
- Indicator 2 (Years 2 and 3): SBMs attend 2+ days of training during their mentoring year
- Indicator 3 (Years 2 and 3): Residents attend meetings with IMs and SBMs (25+ in Year 2, 12+ in Year 3)
- Indicator 4 (Years 2 and 3): Residents complete 2+ observation cycles with their IM

Results of Key Component 3 for Year 2

CREATE expected that at least 90% of SBMs would attend a summer training prior to beginning their work as a mentor and that at least 90% of them would attend two or more trainings, which included coaching calls with CREATE, during the school year, while they served as mentors. Additionally, CREATE expected that 90% of residents would meet with their mentors at least 25 times during the year, and that 90% of them would complete at least 2 observation cycles with their IM during the year. CREATE did not meet the implementation goals for Indicator 1 for Cohort 6 or Cohort 7 in Year 2, and did meet the implementation goals for both Cohort 6 and Cohorts 7 for Indicator 2. Cohort 6 also met fidelity for Indicators 3 and 4, while Cohort 7 met fidelity thresholds for Indicator 3, but not Indicator 4. Cases in which CREATE did not meet fidelity thresholds may have been due to changing expectations for mentors who continued to serve as CREATE mentors for multiple school years (after they had already received CREATE training in previous mentoring years). Because all indicators must be met in order for Component 3 to reach fidelity at the program level, neither Cohort 6 nor Cohort 7 met fidelity on this component for Year 2. See Table G5 below.

TABLE G5. DETAILS OF COMPONENT 3 IN YEAR 2

Indicator	Fidelity threshold	Cohort 6 met fidelity?	Cohort 7 met fidelity?	Combined score
1: School-Based Mentors attend 1+ training before mentoring	0 = School-Based Mentor did not attend a training before mentoring 1 = School-Based Mentor attended at least one training before mentoring 90% of residents have mentors who meet fidelity	15/17 (88%) Fidelity not met	8/11 (73%) Fidelity not met	23/28 (82%) Fidelity not met
2: School-Based Mentors attend 2+ days of training during their mentoring year	0 = School-Based Mentor attends 0–1 training day 1 = School-Based Mentor attends 2 or more training days 90% of residents have mentors who meet fidelity	17/17 (100%) Fidelity met	11/11 (100%) Fidelity met	28/28 (100%) Fidelity met
3: Residents attend meetings with Instructional Mentors and School-Based Mentors (25+ in Year 2, 12+ in Year 3)	0 = Resident attends 0–20 meetings 1 = Resident attends 20–24 meetings 2 = Resident attends 25 or more meetings 90% of residents earn a score of 2 at the unit level and no residents earn a score of 0	17/17 (100%) receive score of 2 None receive score of 1 None receive score of 0 Fidelity met	11/11 (100%) receive score of 2 None receive score of 1 None receive a score of 0 Fidelity met	28/28 (100%) receive score of 2 None receive score of 1 None receive a score of 0 Fidelity met
4: Residents complete 2+ observation cycles with their Instructional Mentor	0 = Resident completes 0–1 cycle 1 = Resident completes 2–3 cycles 90% or more of residents meet fidelity	17/17 (100%) Fidelity met	9/11 (82%) Fidelity not met	26/28 (93%) Fidelity met
Program level	All indicators meet fidelity	Fidelity not met	Fidelity not met	Fidelity not met

Note. Red = did NOT meet program-level fidelity threshold.

Results of Key Component 3 for Year 3

As mentioned above, CREATE expected that 90% of mentors would attend a summer training prior to beginning their work as a mentor and that at least 90% of them would attend two or more trainings during the school year in which they serve as a mentor. In Year 3, CREATE expected that 90% of residents would meet with their mentors at least 12 times during the year, and that 80% of them would complete at least 2–3 observation cycles with their IM during the year. Cohort 6 did not meet fidelity for Indicator 1 in Year 3. Cohort 6 did meet fidelity for Indicator 2 during Year 3. As Indicators 1 and 2 address SBM trainings, program staff may have adjusted expectations for the repeat mentors who had already received training in mentoring from CREATE in previous mentoring years. CREATE met the goals for Indicator 3, but not for Indicator 4—which measured resident-mentor meetings and observation cycles—for Cohort 6 in Year 3. See Table G6.

TABLE G6. DETAILS OF COMPONENT 3 IN YEAR 3

Indicator	Fidelity threshold	Cohort 6 met fidelity?
1: School-Based Mentors attend 1+ training before mentoring	0 = School-Based Mentor did not attend a training before mentoring 1 = School-Based Mentor attended a training before mentoring 90% of residents have mentors who meet fidelity	8/12 (67%) Fidelity not met
2: School-Based Mentors attend 2+ days of training during their mentoring year	0 = School-Based Mentor attends 0–1 training day 1 = Resident has a mentor that attends 2 or more training days 90% residents have mentors who meet fidelity	11/12 (92%) Fidelity met
3: Residents attend meetings with Instructional Mentors and School-Based Mentors (25+ in Year 2, 12+ in Year 3)	0 = Resident attends 0–8 meetings 1 = Resident attends 9–11 meetings 2 = Resident attends 12 or more meetings 90% of residents earn a score of 2 at the unit level and no residents earn a score of 0	12/12 (100%) receive a score of 2 Fidelity met
4: Residents complete 2+ observation cycles with their Instructional Mentor	0 = Resident completes 0–1 cycle 1 = Resident completes 2–3 cycles 80% or more of residents meet fidelity	6/12 (50%) Fidelity not met
Program level	All indicators meet fidelity	Fidelity not met

Note. Red = did NOT meet program-level fidelity threshold.

Overall Results of Key Component 3

For Year 2, Indicator 1, Cohort 6 mentors did not meet fidelity for attending training in the summer prior to the mentoring year. Cohort 7, in Year 2, did not meet fidelity for Indicators 1 and 4, which relate to mentors attending summer trainings and completing observation cycles. For Year 3, Cohort 6 did not meet Indicators 1 and 4. CREATE Cohorts 6 and 7 consistently met fidelity for all years of residency for Indicator 2, which measure mentors attending mentor trainings during the mentoring year.

Key Component 4: Summer Resident Academy

SRA is an intensive training for new teachers that occurs during the summers during the CREATE residency, after residents graduate from GSU CEHD and before they begin their first year as full-time teachers. We measured Key Component 4 in Year 2 and Year 3, after residents graduated from GSU CEHD and had begun teaching as full-time credentialed teachers (see Table G7 and Table G8). We do not measure Component 4 in Year 1 of residency, because CREATE offered an SRA for residents entering their first year of residency for the first time in summer 2021. In addition to inviting residents who were entering Year 1 of residency, CREATE also invited those entering Year 2 of residency to participate, although rising Year 1 residents (Cohort 7) had priority. Attendance at this SRA was optional for all participants. We measured FOI for this component in the year after residents attended each SRA. Activities that residents participated in during the academy included preparing lesson plans that would contribute to a safe and culturally responsive classroom, learning research-based instructional skills, practicing strategies to maintain personal and physical heath, among other skills and competencies they would need as full-time teachers. We present data on the following indicator related to SRA.

- Indicator 1: Residents attend at least 85% of SRA days. At least 80% of residents would need to attend 85% of each SRA to meet fidelity at the unit level.

TABLE G7. DETAILS OF COMPONENT 4 IN YEAR 2

Indicator	Fidelity threshold	Cohort 6 met fidelity?	Cohort 7 met fidelity?	Combined score
1: Residents attend at least 85% of Summer Resident Academy days	0 = Resident attends less than 85% of Summer Resident Academy days 1 = Resident attends 85% or more of Summer Resident Academy days 80% or more of residents earn a score of 1 at the unit level	17/17 (100%) of active residents receive a score of 1 Fidelity met	10/11 (91%) of active residents receive a score of 1 Fidelity met	27/28 (96%) Fidelity met
Program level	80% or more of residents earn a score of 1 at the unit level	Fidelity met	Fidelity met	Fidelity met

Note. Green = met program-level fidelity threshold.

TABLE G8. DETAILS OF COMPONENT 4 IN YEAR 3

Indicator	Fidelity threshold	Cohort 6 met fidelity?
1: Residents attend at least 85% of Summer Resident Academy days	0 = Resident attends less than 85% of Summer Resident Academy days 1 = Resident attends 85% or more of Summer Resident Academy days 80% or more of residents earn a score of 1 at the unit level	6/12 (50%) of active residents receive a score of 1 Fidelity not met
Program level	80% or more of residents earn a score of 1 at the unit level	Fidelity not met

Note. Red = did NOT meet program-level fidelity threshold.

Overall Results of Key Component 4

For Year 2, Component 4 was implemented with fidelity, as 96% of Cohort 6 and Cohort 7 residents attended SRA between Year 1 and Year 2. For Year 3, this component did not meet fidelity, as fewer than 80% of residents received a score of 1 at the individual level. Half of Cohort 6 residents did not attend at least 85% of the days that constituted the SRA between Year 2 and Year 3.

Key Component 5: Equity-Centered Professional Learning

CREATE designed the equity-centered PL opportunities to enhance pedagogical skills that interrupt the inequities that teachers encounter at their schools. Through its equity-focused PL, CREATE aims to equip educators to closely examine issues of oppression, power, and privilege present in school structures, and to address inequitable practices in education. While CREATE offers equity-centered PL for all educators at CREATE partner schools, as well as other educators in the community, CREATE residents experience learning around equity-based practices at monthly TT meetings with other residents in their cohort. Each resident should attend a CREATE PL opportunity, other than TT, by the end of Year 3 of their residency.

We present data on the following indicators related to equity-centered PL.

- Indicator 1 (Years 1, 2, and 3): CREATE hosts 2+ equity-centered PL meetings per year
- Indicator 2 (Years 1, 2, and 3): Experienced educators attend 50% of equity-centered PL days
- Indicator 3 (Year 3 only): Residents attend at least 50% of a CREATE PL opportunities (other than TT) in Year 2 or Year 3 of residency
- Indicator 4 (Years 1, 2, and 3): Residents attend TT meetings (8+ in Year 1, 6+ in Year 2, 5+ in Year 3)

Results of Key Component 5 for Year 1

CREATE administrators expect that Indicator 1 and at least one other indicator listed above will reach fidelity in Year 1. Cohorts 6, 7, and 8 met fidelity at the program level for this component. At the indicator level, Cohorts 6 through 8 met or exceeded fidelity thresholds for all indicators, with the exception of Cohort 7 for Indicator 4; fewer than the expected 95% attended monthly TT meetings, which included training in equity-centered classroom practices. More details about the number of residents who met each specific indicator are in the table below.

The first two indicators in the Table G9 refer to experienced educators, who are not part of a particular cohort of residents. For this reason, columns 3, 4, and 5 correspond to the proper school year, rather than the residency cohort.

TABLE G9. DETAILS OF COMPONENT 5 IN YEAR 1

Indicator	Fidelity threshold	2020–21 school year	2021–22 school year	2022–23 school year	Combined score
1: CREATE hosts 2+ equity-centered PLs per year	0 = CREATE hosts 0–1 equity-centered PL 1= CREATE hosts 2+ equity-centered PLs	CREATE hosted 3 equity-centered PLs Fidelity met	CREATE hosted 3 equity-centered PLs Fidelity met	CREATE hosted 3 equity-centered PLs Fidelity met	All met fidelity
2: Experienced educators attend 50% of equity-centered PL days	0 = Educator attends 50% of equity-centered PL days 1 = Educator attends more than 50% of equity-centered PL days Of the experienced educators who attend, 85% or more meet fidelity	45/45 (100%) Fidelity met	52/53 (98%) Fidelity met	30/33 (91%) Fidelity met	127/131 (97%) Fidelity met

Indicator	Fidelity threshold	Cohort 6 met fidelity?	Cohort 7 met fidelity?	Cohort 8 met fidelity?	Combined score
4: Residents attend Together Time (8+ in Year 1, 6+ in Year 2, 5+ in Year 3)	0 = Attend 0–7 meetings 1 = Attend 8–10 meetings 95% or more of residents meet fidelity	21/21 (100%) Fidelity met	23/28 (82%) Fidelity not met	21/22 (95%) Fidelity met	65/71 (92%) Fidelity met
Program level	0 = Fidelity not met for any indicator, or fidelity met for Indicator 1 but no other indicator 1 = Fidelity was met for Indicator 1 and at least one other indicator	Fidelity met	Fidelity met	Fidelity met	Fidelity met

Note. PL = professional learning. Indicator 3 (residents attend a CREATE PL opportunity outside of Together Time by the end of Year 3 of their program) is not in this table because it does not apply to Year 1 residents. Green = met program-level fidelity threshold.

Results of Key Component 5 for Year 2

CREATE administrators expected that Indicator 1 and at least one other indicator would be met in Year 2, in order to reach FOI. The records from CREATE attendance rosters and survey data show that Cohorts 6 and 7 met fidelity at the program level in Year 2 for this component. All indicators met fidelity, with the exception of Indicator 4 for Cohort 7, when 91% of Cohort 7 residents attended at least six TT meetings during the 2021–22 school year. This caused FOI calculations to fall below the fidelity threshold of 95% residents attending at least six TT meetings throughout the school year. When interpreting these results, it is important to acknowledge that the small sample size of Cohort 7 residents in Year 2 combined with the high threshold percentage for this indicator would have required 100% of residents to meet the threshold in order for this indicator to meet fidelity. Though these two factors combined led to a failure to meet fidelity for this indicator, the fact remains that CREATE saw TT attendance rates that met or exceeded their expectations from ten of the eleven active Cohort 7 residents.

More details about the number of residents who met each specific indicator are in the Table G10 below.

TABLE G10. DETAILS OF COMPONENT 5 IN YEAR 2

Indicator	Fidelity threshold	2021–22 school year	2022–23 school year	Combined score
1: CREATE hosts 2+ equity-centered PLs per year	0 = CREATE hosts 0–1 equity-centered PL 1= CREATE hosts 2+ equity-centered PLs	CREATE hosted 3 equity-centered PLs Fidelity met	CREATE hosted 3 equity-centered PLs Fidelity met	Fidelity met
2: Experienced educators attend 50% of equity-centered PL days	0 = Educator attends 50% of equity-centered PL days 1 = Educator attends more than 50% of equity-centered PL days 85% or more meet fidelity	52/53 (98%) Fidelity met	30/33 (91%) Fidelity met	82/86 (95%) Fidelity met
Indicator	**Fidelity threshold**	**Cohort 6 met fidelity?**	**Cohort 7 met fidelity?**	**Combined score**
4: Residents attend Together Time meetings (8+ in Year 1, 6+ in Year 2, 5+ in Year 3)	0 = Attend 0–5 meetings 1 = Attend 6+ meetings 95% or more of residents meet fidelity	17/17 (100%) Fidelity met	10/11 (91%) Fidelity not met	27/28 (96%) Fidelity met
Program level	0 = Fidelity not met for any indicator, or fidelity met for Indicator 1 but no other indicator 1 = Fidelity was met for Indicator 1 and at least one other indicator	Fidelity met	Fidelity met	Fidelity met

Note. PL = professional learning. Indicator 3 (Residents attend a CREATE PL outside of Together Time by the end of Year 3 of their program) is not in this table because it does not apply to Year 1 residents. Note. Green = met program-level fidelity threshold.

Results of Key Component 5 for Year 3

CREATE administrators expected that Indicator 1, Indicator 3, and at least one other indicator in Year 3 would reach FOI. The records from CREATE attendance rosters and survey data show that Cohort 6 did not meet fidelity for Component 5 in Year 3. Cohort 6 did not meet expected thresholds for resident attendance at TT meetings during which training in equity-centered classroom practices was provided; 25% of Cohort 6 residents did not attend the minimum expectation of five TT meetings. Each of these three Cohort 6 residents attended four TT meetings during Year 3, making them one meeting short of meeting the fidelity threshold on the individual level. According to CREATE attendance rosters, a number of these absences had justifiable excuses. Cohort 6 also did not meet the fidelity threshold for Indicator 3, which measures the attendance of residents at CREATE PL outside of TT meetings. CREATE expected at least 30% of residents to attend a CREATE PL meeting by the end of Year 3 of their residency. In 2021–22, CREATE hosted three PPI institutes and two ECCF Institutes. In 2022–23, CREATE hosted two ECCF institutes, one PPI institute, and one CREATE 101: Community Engagement and Continuing Education institute. Resident attendance at any of these PLs, which occurred both during the summers between school years and throughout the school years themselves, counted towards implementation of Indicator 3 with fidelity. Only 17% of Cohort 6 residents attended at least one PL during their residency, so Cohort 6 did not meet fidelity for this indicator and, therefore, the entire component for Cohort 6 in Year 3.

More details about the number of residents who met each specific indicator are in the Table G11 below.

TABLE G11. DETAILS OF COMPONENT 5 IN YEAR 3

Indicator	Fidelity threshold	2022–23 school year
1: CREATE hosts 2+ equity-centered PLs per year	0 = CREATE hosts 0–1 equity-centered PL 1 = hosts 2+ equity-centered PLs	CREATE hosted 3 equity-centered PLs Fidelity met
2: Experienced educators attend 50% of equity-centered PL days	0 = Educator attends 50% of equity-centered PL days 1 = Educator attends more than 50% of equity-centered PL days 85% or more meet fidelity	30/33 (91%) Fidelity met
Indicator	**Fidelity threshold**	**Cohort 6 met fidelity?**
3: Residents attend at least 50% of a CREATE PL (other than Together Time) in Year 2 or Year 3 of residency	0 = Resident attends 50% of PL days 1 = Resident attends more than 50% of PL days 30% or more of residents meet fidelity	2/12 (17%) Fidelity not met

TABLE G11. DETAILS OF COMPONENT 5 IN YEAR 3

Indicator	Fidelity threshold	2022–23 school year
4: Residents attend Together Time (8+ in Year 1, 6+ in Year 2, 5+ in Year 3)	0 = Attend 0–4 meetings 1 = Attend 5+ meetings 95% or more of residents meet fidelity	9/12 (75%) Fidelity not met
Program level	0 = Fidelity not met for indicators 1 or 3, OR fidelity met for indicators 1 and 3 but not for Indicator 2 or 4. 1 = Fidelity was met for Indicator 1, Indicator 3 and at least one other indicator	Fidelity not met

Note. PL = professional learning. Red = did NOT meet program-level fidelity threshold.

Overall Results of Key Component 5

CREATE consistently met or exceeded the fidelity threshold for offering equity-centered PL, at which they achieved a high level of consistent attendance by experienced educators. Indicator 4 measures CREATE residents' attendance at TT meetings, during which they receive equity-centered training. Cohort 6 did not meet Indicator 4 in Year 3; Cohort 7 did not meet Indicator 4 in Year 1 or Year 2. This was the 2021–22 and 2022–23 school years during which the effects from the COVID-19 pandemic, sociopolitical unrest locally and on a national level, and changes to the CREATE program continued to affect the schooling, work, and personal lives of residents.

At the program level, Cohorts 6, 7, and 8 met fidelity for Year 1 and 2, but Cohort 6 did not meet fidelity for Indicator 3 in Year 3.

Key Component 6: Compassion-Based Professional Learning

Acknowledging that workplace and relationship stressors diminish a teacher's ability to meet the demands of teaching, all CREATE residents engaged in regularly scheduled compassion-based trainings throughout their residency. During the grant years, CREATE's compassion-based PL opportunity aimed to teach techniques designed to help educators develop a greater awareness of their own attitudes and behaviors and how these attitudes and behaviors impact their relationships with others. The primary goal of CREATE's compassion-based PLs is to teach educators how to respond to stressful situations in a healthy manner and strengthen their ability to conjure empathy and compassion for themselves and others during stressful situations via practicing increased awareness of their thoughts and feelings as they occur in each moment.

We present data on the following indicators related to compassion-based PL.

- Indicator 1 (Years 1, 2, and 3): CREATE hosts 1+ compassion-based PL meetings per year
- Indicator 2 (Years 1, 2, and 3): Residents attend TT meetings (8+ in Year 1, 6+ in Year 2, 5+ in Year 3)

Results of Key Component 6 for Year 1

In Year 1, CREATE expected to hold at least one compassion-based PL opportunity for the general population of teachers at CREATE partner schools and for other educators. CREATE exceeded this goal by offering two compassion-based institutes during the 2020–21 school year, three compassion-based institutes during the 2021–22 school year, and one compassion-based institute during the 2022–23 school year. Additionally, they expected 95% or more of the residents to receive compassion-based training by attending at least eight TT meetings during Year 1 of their residency. They met this goal for Cohorts 6 and Cohort 8. Cohort 7 did not meet fidelity for this indicator, with 82% of residents having attended at least eight TT meetings. Because they did not meet fidelity for both indicators, Cohort 7 did not meet fidelity at the program level for this component. **Despite the score not meeting fidelity, we must note that the combined score for all three cohorts show that 92% (65/71) of all Year 1 residents in Cohorts 6 through 8 met TT attendance expectations. The efforts and commitment to professional educator development of both CREATE and its residents are evident in these numbers.** Cohort 6 and Cohort 8 met fidelity overall for Component 6 in Year 1. See Table G12 below.

TABLE G12. DETAILS OF COMPONENT 6 IN YEAR 1

Indicator	Fidelity threshold	2020–21 school year	2021–22 school year	2022–23 school year	Combined score
1: CREATE hosts 1+ compassion-based PLs per year	0 = CREATE offers 0 PLs 1 = CREATE offers 1 or more PL	CREATE hosted 2 compassion-based PLs Fidelity met	CREATE hosted 3 compassion-based PLs Fidelity met	CREATE hosted 1 compassion-based PL Fidelity met	Fidelity met
Indicator	**Fidelity threshold**	**Cohort 6 met fidelity?**	**Cohort 7 met fidelity?**	**Cohort 8 met fidelity?**	**Combined score**
2: Residents attend Together Time meetings (8+ in Year 1, 6+ in Year 2, 5+ in Year 3)	0 = Attend 0–7 meetings 1 = Attend 8–10 meetings 95% or more of residents meet fidelity	21/21 (100%) Fidelity met	23/28 (82%) Fidelity not met	21/22 (95%) Fidelity met	65/71 (92%) Fidelity not met
Program level	Component score is sum of indicator scores. Fidelity not met = score of 0–1 Fidelity met = score of 2	Fidelity met	Fidelity not met	Fidelity met	Fidelity not met

Note. PL = professional learning. Red = did NOT meet program-level fidelity threshold.

Results of Key Component 6 for Year 2

In Year 2, CREATE expected to hold at least one compassion-based course for the general population of teachers at CREATE partner schools and for other educators. CREATE exceeded this goal by offering three compassion-based PPI institutes during the 2021–22 school year, and one compassion-based PPI institute during the 2022–23 year, in June 2022. Additionally, they expected 95% or more of the residents to receive compassion-based training through attending at least

six TT meetings in Year 2. CREATE met this goal for Cohort 6, but did not meet this goal for Cohort 7. At the program level, Cohort 6 met fidelity, but Cohort 7 did not meet fidelity for Component 6 in Year 2. See Table G13 below.

TABLE G13. DETAILS OF COMPONENT 6 IN YEAR 2

Indicator	Fidelity threshold	2021-22 school year	2022–23 school year	Combined score
1: CREATE hosts 1+ compassion-based PLs per year	0 = CREATE offers 0 PLs 1 = CREATE offers 1 or more PL	CREATE hosted 3 compassion-based PLs Fidelity met	CREATE hosted 1 compassion-based PLs Fidelity met	Fidelity met
Indicator	**Fidelity threshold**	**Cohort 6 met fidelity?**	**Cohort 7 met fidelity?**	**Combined score**
2: Residents attend Together Time meetings (8+ in Year 1, 6+ in Year 2, 5+ in Year 3)	0 = Attend 0–5 meetings 1 = Attend 6+ meetings 95% or more of residents meet fidelity	17/17 (100%) Fidelity met	10/11 (91%) Fidelity not met	27/28 (96%) Fidelity met
Program level	Component score is sum of indicator scores. Fidelity not met = score of 0-1 Fidelity met = score of 2	Fidelity met	Fidelity not met	Fidelity met

Note. PL = professional learning. Green = met program-level fidelity threshold.

Results of Key Component 6 for Year 3

In Year 3, CREATE expected to hold at least one compassion-based PL opportunity for the general population of teachers at CREATE partner schools and for other educators. CREATE met this goal by offering one compassion-based PL during the 2022–23 school year. Additionally, they expected 95% or more of the residents to receive compassion-based training by attending at least five TT meetings in Year 3. CREATE did not meet this goal for Cohort 6. At the program level, Cohort 6 did not meet fidelity for Component 6 in Year 3. See Table G14 below.

TABLE G14. DETAILS OF COMPONENT 6 IN YEAR 3

Indicator	Fidelity threshold	2022–23 school year
1: CREATE hosts 1+ compassion-based PLs per year	0 = CREATE offers 0 PLs 1 = CREATE offers 1 or more PL	CREATE hosted 1 compassion-based PL Fidelity met

TABLE G14. DETAILS OF COMPONENT 6 IN YEAR 3

Indicator	Fidelity threshold	Cohort 6 met fidelity?
2: Residents attend Together Time meetings (8+ in Year 1, 6+ in Year 2, 5+ in Year 3)	0 = Attend 0–4 meetings 1 = Attend 5+ meetings 95% or more of residents meet fidelity	9/12 (75%) Fidelity not met
Program level	Component score is sum of indicator scores. Fidelity not met = score of 0–1 Fidelity met = score of 2	Fidelity not met

Note. PL = professional learning. Red = did NOT meet program-level fidelity threshold.

Overall Results of Key Component 6

CREATE regularly offered more than one compassion-based PL opportunity per year for all years during which this indicator was measured. CREATE did not meet fidelity for Cohort 6 during the 2022–23 school year (Year 3) or for Cohort 7 during the 2021–22 (Year 1) and the 2022–23 (Year 2) school years for Indicator 2. This indicator measures residents' attendance at TT meetings, at which they receive compassion-based training. The effects from the COVID-19 pandemic and other local and national movements catalyzed many changes to schooling, the world of work, and personal life during these school years. We suspect that these changes and the need for adjustment directly affected the low attendance rates at TT meetings. We calculated this indicator using the same attendance data as Indicator 4 for Component 2, so fidelity results for these two indicators will always be the same. Cohort 8 met fidelity for this indicator in Year 1 of their residency. A combination of schools fully returning to in-person learning for the first time since the COVID-19 pandemic and CREATE's adjustments to its programming—in response to challenges with resident attendance and engagement in the preceding years—may have supported Cohort 8.

At the program level, CREATE met fidelity for the compassion-based PL component for Cohorts 6 and 8 during Year 1 and Cohorts 6 during Year 2.

Appendix H. Technical Details of Teacher Well-Being Analysis

Table H1 through Table H24 present the WBL outcomes and Table H25 through Table H39 present survey outcomes.

Well-Being Log Outcomes

TABLE H1. SUMMARY OF BASELINE EQUIVALENCE TESTS FOR ANALYSIS OF IMPACTS ON WELL-BEING LOG OUTCOMES

Scale	Sample size (CREATE)	Raw score mean (CREATE)	Raw score standard deviation (CREATE)	Sample size (comparison)	Raw score mean (comparison)	Raw score standard deviation (comparison)	Regression adjusted difference estimate	Standardized mean difference (Effect Size)
Motivation for Entering Teaching	17	4.64	0.27	17	4.55	0.44	0.08	0.23
Confidence in General Teaching Skills	17	4.04	0.55	17	3.92	0.55	0.12	0.21
CD-RISC	17	4.18	0.59	17	4.06	0.80	0.12	0.17

Note. N = 34. The standardized mean difference is the regression-adjusted difference estimate based on a model with dummy variables to indicate cohort membership and a variable indicating the treatment condition, divided by the pooled standard deviation of the pretest distribution. CD-RISC = Connor-Davidson Resilience Scale.

TABLE H2. ANALYTIC SAMPLE SIZES BY COHORT

	CREATE sample	Comparison sample	Total
Cohort 6	8	11	19
Cohort 7	9	6	15
Total	17	17	34

TABLE H3. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON FOCUS & AWARENESS

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	2.230	2.059	28	1.08	0.288
Impact of CREATE	0.592	0.300	28	1.98	0.058
Belongs to Cohort 7	0.126	0.301	28	0.42	0.680
Confidence in teaching	-0.156	0.304	28	-0.51	0.613
Motivation for teaching	0.143	0.427	28	0.33	0.741
CD-RISC	0.336	0.228	28	1.47	0.151

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale).

TABLE H4. EFFECT SIZE TABLE FOR FOCUS & AWARENESS

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.68	0.95	17	0.78	.028
	CREATE	4.32	0.68	17		
Adjusted effect size[b]	Comparison	3.68	0.95	17	0.71	.058
	CREATE	4.27	0.68	17		

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.

TABLE H5. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON DIFFCULTY WITH ATTENTION

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	2.354	3.226	28	0.73	0.472
Impact of CREATE	-0.304	0.469	28	-0.65	0.523
Belongs to Cohort 7	0.142	0.472	28	0.30	0.766
Confidence in teaching	0.112	0.476	28	0.24	0.816
Motivation for teaching	-0.359	0.670	28	-0.54	0.596
CD-RISC	0.507	0.357	28	1.42	0.166

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale).

TABLE H6. EFFECT SIZE TABLE FOR DIFFICULTY WITH ATTENTION

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.26	1.23	17	-0.18	.600
	CREATE	3.03	1.37	17		
Adjusted effect size[b]	Comparison	3.26	1.23	17	-0.23	.523
	CREATE	2.96	1.37	17		

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.

TABLE H7. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON SUPPORTED & HOPEFUL

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	1.805	1.387	28	1.30	0.204
Impact of CREATE	0.064	0.202	28	0.32	0.754
Belongs to Cohort 7	-0.316	0.203	28	-1.56	0.131
Confidence in teaching	-0.248	0.205	28	-1.21	0.235
Motivation for teaching	0.130	0.288	28	0.45	0.656
CD-RISC	0.506	0.153	28	3.30	0.003

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale).

TABLE H8. EFFECT SIZE TABLE FOR SUPPORTED & HOPEFUL

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.36	0.71	17	0.08	.816
	CREATE	3.41	0.58	17		
Adjusted effect size[b]	Comparison	3.36	0.71	17	0.10	.754
	CREATE	3.43	0.58	17		

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.

TABLE H9. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON MAIA: SELF-REGULATION

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	3.812	2.250	28	1.69	0.101
Impact of CREATE	0.567	0.327	28	1.73	0.094
Belongs to Cohort 7	0.158	0.329	28	0.48	0.635
Confidence in teaching	-0.008	0.332	28	-0.03	0.980
Motivation for teaching	-0.394	0.467	28	-0.84	0.406
CD-RISC	0.353	0.249	28	1.42	0.167

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale). MAIA = Multidimensional Assessment of Interoceptive Awareness.

TABLE H10. EFFECT SIZE TABLE FOR MAIA: SELF-REGULATION

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.47	1.04	17	0.66	.061
	CREATE	4.07	0.76	17		
Adjusted effect size[b]	Comparison	3.47	1.04	17	0.62	.094
	CREATE	4.04	0.76	17		

Note. The *p* values are for the corresponding impact estimates in the regression model. MAIA = Multidimensional Assessment of Interoceptive Awareness.

[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.

TABLE H11. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON MPERMA: POSITIVE EMOTIONS SUBSCALE

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	3.626	2.047	28	1.77	0.087
Impact of CREATE	0.490	0.298	28	1.64	0.111
Belongs to Cohort 7	-0.043	0.300	28	-0.14	0.886
Confidence in teaching	-0.348	0.302	28	-1.15	0.259
Motivation for teaching	-0.226	0.425	28	-0.53	0.593
CD-RISC	0.575	0.226	28	2.54	0.017

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale). mPERMA = Momentary Positive Emotions, Engagement, Relationships, Meaning, Accomplishment.

TABLE H12. EFFECT SIZE TABLE FOR MPERMA: POSITIVE EMOTIONS SUBSCALE

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.55	0.80	17	0.56	.872
	CREATE	4.04	0.96	17		
Adjusted effect size[b]	Comparison	3.55	0.80	17	0.56	.111
	CREATE	4.04	0.96	17		

Note. The *p* values are for the corresponding impact estimates in the regression model. mPERMA = Momentary Positive Emotions, Engagement, Relationships, Meaning, Accomplishment.

[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.

TABLE H13. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON MPERMA: MEANING SUBSCALE

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	3.899	1.676	28	2.33	0.027
Impact of CREATE	0.621	0.244	28	2.55	0.017
Belongs to Cohort 7	-0.128	0.245	28	-0.52	0.605
Confidence in teaching	-0.092	0.247	28	-0.37	0.713
Motivation for teaching	-0.211	0.348	28	-0.61	0.550
CD-RISC	0.318	0.185	28	1.72	0.097

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale). mPERMA = Momentary Positive Emotions, Engagement, Relationships, Meaning, Accomplishment.

TABLE H14. EFFECT SIZE TABLE FOR MPERMA: MEANING SUBSCALE

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.82	0.54	17	0.89	.013
	CREATE	4.43	0.81	17		
Adjusted effect size[b]	Comparison	3.82	0.54	17	0.90	.017
	CREATE	4.44	0.81	17		

Note. The *p* values are for the corresponding impact estimates in the regression model. mPERMA = Momentary Positive Emotions, Engagement, Relationships, Meaning, Accomplishment.

[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.

TABLE H15. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON MPERMA: ACCOMPLISHMENT SUBSCALE

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	2.925	1.508	28	1.94	0.063
Impact of CREATE	0.373	0.219	28	1.70	0.101
Belongs to Cohort 7	-0.202	0.221	28	-0.91	0.368
Confidence in teaching	-0.102	0.223	28	-0.46	0.649
Motivation for teaching	0.071	0.313	28	0.23	0.821
CD-RISC	0.272	0.167	28	1.63	0.114

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale). mPERMA = Momentary Positive Emotions, Engagement, Relationships, Meaning, Accomplishment.

TABLE H16. EFFECT SIZE TABLE FOR MPERMA: ACCOMPLISHMENT SUBSCALE

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.88	0.37	17	0.59	.092
	CREATE	4.25	0.79	17		
Adjusted effect size[b]	Comparison	3.88	0.37	17	0.60	.101
	CREATE	4.25	0.79	17		

Note. The *p* values are for the corresponding impact estimates in the regression model. mPERMA = Momentary Positive Emotions, Engagement, Relationships, Meaning, Accomplishment.

[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.

TABLE H17. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON MBI: EMOTIONAL EXHAUSTION

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	3.220	3.321	28	0.97	0.341
Impact of CREATE	-0.724	0.483	28	-1.50	0.145
Belongs to Cohort 7	0.117	0.486	28	0.24	0.812
Confidence in teaching	0.660	0.490	28	1.35	0.189
Motivation for teaching	0.487	0.689	28	0.71	0.485
CD-RISC	-0.931	0.367	28	-2.54	0.017

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale). MBI = Maslach Burnout Inventory for Educators.

TABLE H18. EFFECT SIZE TABLE FOR MBI: EMOTIONAL EXHAUSTION SUBSCALE

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	4.29	1.60	17	-0.48	.168
	CREATE	3.59	1.26	17		
Adjusted effect size[b]	Comparison	4.29	1.60	17	-0.50	.145
	CREATE	3.56	1.26	17		

Note. The *p* values are for the corresponding impact estimates in the regression model. MBI = Maslach Burnout Inventory for Educators.

[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.

TABLE H19. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON MBI: DEPERSONALIZATION

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	-1.227	3.359	28	-0.37	0.718
Impact of CREATE	-0.214	0.489	28	-0.44	0.666
Belongs to Cohort 7	-0.305	0.492	28	-0.62	0.540
Confidence in teaching	0.249	0.496	28	0.50	0.619
Motivation for teaching	1.156	0.697	28	1.66	0.108
CD-RISC	-0.586	0.371	28	-1.58	0.126

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale). MBI = Maslach Burnout Inventory for Educators.

TABLE H20. EFFECT SIZE TABLE FOR MBI: DEPERSONALIZATION

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	2.53	1.37	17	-0.15	.662
	CREATE	2.32	1.47	17		
Adjusted effect size[b]	Comparison	2.53	1.37	17	-0.15	.666
	CREATE	2.32	1.47	17		

Note. The *p* values are for the corresponding impact estimates in the regression model. MBI = Maslach Burnout Inventory for Educators.

[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.

TABLE H21. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON MBI: PERSONAL ACCOMPLISHMENT

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	5.901	2.175	28	2.71	0.011
Impact of CREATE	0.292	0.316	28	0.92	0.364
Belongs to Cohort 7	0.260	0.318	28	0.82	0.420
Confidence in teaching	0.349	0.321	28	1.09	0.286
Motivation for teaching	-0.192	0.451	28	-0.42	0.674
CD-RISC	-0.230	0.241	28	-0.96	0.346

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale). MBI = Maslach Burnout Inventory for Educators.

TABLE H22. EFFECT SIZE TABLE FOR MBI: PERSONAL ACCOMPLISHMENT

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	5.55	0.88	17	0.39	.258
	CREATE	5.89	0.85	17		
Adjusted effect size[b]	Comparison	5.55	0.88	17	0.34	.364
	CREATE	5.85	0.85	17		

Note. The *p* values are for the corresponding impact estimates in the regression model. MBI = Maslach Burnout Inventory for Educators.

[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.

TABLE H23. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON SENSE OF EQUITY IN ROLE & EXPERIENCE

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	2.058	1.502	28	1.37	0.181
Impact of CREATE	0.492	0.219	28	2.25	0.032
Belongs to Cohort 7	0.138	0.220	28	0.63	0.535
Confidence in teaching	-0.183	0.222	28	-0.83	0.415
Motivation for teaching	0.162	0.312	28	0.52	0.608
CD-RISC	0.378	0.166	28	2.27	0.031

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale).

TABLE H24. EFFECT SIZE TABLE FOR SENSE OF EQUITY IN ROLE & EXPERIENCE

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.66	0.59	17	0.87	.016
	CREATE	4.21	0.68	17		
Adjusted effect size[b]	Comparison	3.66	0.59	17	0.77	.032
	CREATE	4.15	0.68	17		

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.

Survey Outcomes

TABLE H25. SUMMARY OF BASELINE EQUIVALENCE TESTS FOR ANALYSIS OF IMPACTS ON SURVEY OUTCOMES

Scale	Sample size (CREATE)	Raw score mean (CREATE)	Raw score standard deviation (CREATE)	Sample size (comparison)	Raw score mean (comparison)	Raw score standard deviation (comparison)	Regression adjusted difference estimate	Standardized mean difference
Motivation for Entering Teaching	18	4.67	0.25	18	4.60	0.37	0.07	0.21
Confidence in General Teaching Skills	18	4.06	0.49	18	3.97	0.54	0.10	0.19
CD-RISC	18	4.11	0.54	18	4.09	0.80	0.02	0.03

Note. N = 36. The standardized mean difference is the regression-adjusted difference estimate based on a model with dummy variables to indicate cohort membership and a variable indicating the treatment condition, divided by the pooled standard deviation of the pretest distribution. CD-RISC = Connor-Davidson Resilience Scale.

TABLE H26. SUMMARY OF BASELINE EQUIVALENCE TESTS FOR ANALYSIS OF OPPORTUNITIES & SUPPORT TO THINK ABOUT SOCIAL JUSTIC & EQUITY IN SCHOOL FOR COHORT 6 YEAR 2

Scale	Sample size (CREATE)	Raw score mean (CREATE)	Raw score standard deviation (CREATE)	Sample size (comparison)	Raw score mean (comparison)	Raw score standard deviation (comparison)	Regression adjusted difference estimate	Standardized mean difference (Effect Size)
Motivation for Entering Teaching	10	4.60	0.27	10	4.68	0.29	-0.08	-0.29
Confidence in General Teaching Skills	10	4.16	0.49	10	3.92	0.48	0.24	0.49
CD-RISC	10	4.03	0.46	10	4.20	0.65	-0.17	-0.30

Note. The Opportunities & Support to Think About Social Justice & Racial Equity in Schools scale was only administered during Cohort 6, Year 2; therefore, our analysis includes only the Cohort 6 sample.

N = 20. The standardized mean difference is the regression-adjusted difference estimate based on a model with dummy variables to indicate cohort membership and a variable indicating the treatment condition, divided by the pooled standard deviation of the pretest distribution. CD-RISC = Connor-Davidson Resilience Scale.

TABLE H27. ANALYTIC SAMPLE SIZES BY COHORT

	CREATE sample	Comparison sample	Total
Cohort 6	10	10	20
Cohort 7	8	8	16
Total	18	18	36

Note. The Opportunities & Support to Think About Social Justice & Racial Equity in Schools scale was only administered during Cohort 6, Year 2; therefore, our analysis includes only the Cohort 6 sample.

TABLE H28. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON SELF-COMPASSION SCALE

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	2.591	1.374	30	1.89	0.069
Impact of CREATE	0.163	0.177	30	0.92	0.366
Belongs to Cohort 7	-0.449	0.177	30	-2.54	0.017
Confidence in teaching	-0.114	0.193	30	-0.59	0.558
Motivation for teaching	0.124	0.306	30	0.41	0.687
CD-RISC	0.151	0.142	30	1.06	0.297

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale).

TABLE H29. EFFECT SIZE TABLE FOR SELF-COMPASSION SCALE

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.13	0.46	18	0.29	.400
	CREATE	3.29	0.64	18		
Adjusted effect size[b]	Comparison	3.13	0.46	18	0.29	.366
	CREATE	3.29	0.64	18		

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.

TABLE H30. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON STRESS MANAGEMENT & EMPATHY RELATED TO TEACHING

Fixed effects	Estimate	Standard error	*df*	t--value	Pr > |t|
Intercept	2.728	1.497	30	1.82	0.078
Impact of CREATE	0.245	0.193	30	1.27	0.214
Belongs to Cohort 7	0.430	0.193	30	2.23	0.034
Confidence in teaching	-0.013	0.210	30	-0.06	0.951
Motivation for teaching	0.113	0.334	30	0.34	0.736
CD-RISC	0.137	0.155	30	0.88	0.384

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale).

TABLE H31. EFFECT SIZE TABLE FOR STRESS MANAGEMENT & EMPATHY RELATED TO TEACHING

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.95	0.57	18	0.43	.215
	CREATE	4.20	0.61	18		
Adjusted effect size[b]	Comparison	3.95	0.57	18	0.41	.214
	CREATE	4.19	0.61	18		

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.

TABLE H32. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON TEACHER SOCIAL CAPITAL SCALE: TEACHER BONDING

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	4.608	2.235	30	2.06	0.048
Impact of CREATE	0.226	0.288	30	0.79	0.438
Belongs to Cohort 7	0.613	0.288	30	2.13	0.041
Confidence in teaching	-0.220	0.313	30	-0.70	0.487
Motivation for teaching	-0.346	0.498	30	-0.69	0.493
CD-RISC	0.263	0.231	30	1.14	0.265

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale).

TABLE H33. EFFECT SIZE TABLE FOR TEACHER SOCIAL CAPITAL SCALE: TEACHER BONDING

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.49	0.74	18	0.21	.541
	CREATE	3.68	1.02	18		
Adjusted effect size[b]	Comparison	3.49	0.74	18	0.25	.438
	CREATE	3.72	1.02	18		

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.

TABLE H34. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON TEACHER SOCIAL CAPITAL SCALE: OPPORTUNITIES TO COLLABORATE

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	3.118	1.706	30	1.83	0.078
Impact of CREATE	0.337	0.220	30	1.53	0.136
Belongs to Cohort 7	0.603	0.220	30	2.74	0.010
Confidence in teaching	0.128	0.239	30	0.53	0.598
Motivation for teaching	-0.395	0.380	30	-1.04	0.308
CD-RISC	0.301	0.177	30	1.71	0.098

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale).

TABLE H35. EFFECT SIZE TABLE FOR TEACHER SOCIAL CAPITAL SCALE: OPPORTUNITIES TO COLLABORATE

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.31	0.77	18	0.45	.194
	CREATE	3.64	0.68	18		
Adjusted effect size[b]	Comparison	3.31	0.77	18	0.47	.136
	CREATE	3.65	0.68	18		

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.

TABLE H36. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON OPPORTUNITIES & SUPPORT TO THINK ABOUT SOCIAL JUSTICE & RACIAL EQUITY ISSUES IN SCHOOLS

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	9.784	5.438	15	1.80	0.092.
Impact of CREATE	1.694	0.551	15	3.07	0.008
Confidence in teaching	-0.446	0.568	15	-0.78	0.445
Motivation for teaching	-1.057	1.010	15	-1.05	0.312
CD-RISC	-0.164	0.498	15	-0.33	0.747

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale). We only asked The Opportunities & Support to Think About Social Justice & Racial Equity in Schools scale during Cohort 6 Year 2, so there is no row for the effect of membership in Cohort 7 (relative to Cohort 6). There is also a different *df* value due to the difference in sample size.

TABLE H37. EFFECT SIZE TABLE FOR OPPORTUNITIES & SUPPORT TO THINK ABOUT SOCIAL JUSTICE & RACIAL EQUITY ISSUES IN SCHOOLS

	Condition	Means	Standard deviations	No. of teachers[c]	Effect size	*p* value
Unadjusted effect size[a]	Comparison	2.40	1.07	10	1.49	<.001
	CREATE	4.10	1.20	10		
Adjusted effect size[b]	Comparison	2.40	1.07	10	1.49	.008
	CREATE	4.09	1.20	10		

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.

[c] This scale was only administered during Cohort 6, Year 2; therefore, the number of teachers for this item is 10 comparison participants and 10 CREATE participants.

TABLE H38. DETAILED RESULTS OF REGRESSION-ADJUSTED ANALYSIS OF THE FIXED EFFECTS OF CREATE ON CONFIDENCE IN UNDERSTANDING CONCEPTS RELATED TO RACE & EQUITY IN SCHOOLS

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	6.248	2.214	30	2.82	0.008
Impact of CREATE	0.675	0.285	30	2.37	0.025
Belongs to Cohort 7	0.185	0.285	30	0.65	0.521
Confidence in teaching	0.662	0.310	30	2.13	0.041
Motivation for teaching	-1.043	0.493	30	-2.11	0.043
CD-RISC	-0.188	0.229	30	-0.82	0.418

Note. The estimate for the intervention is the average impact of CREATE on the outcome. Analysis includes the remaining baseline covariates to increase the precision of the impact estimate. They include the effects of membership in Cohort 7 (relative to Cohort 6), Confidence in Teaching, Motivation to Teach, and Resilience (assessed through the Connor-Davidson Resilience scale).

TABLE H39. EFFECT SIZE TABLE FOR CONFIDENCE IN UNDERSTANDING CONCEPTS RELATED TO RACE & EQUITY IN SCHOOLS

	Condition	Means	Standard deviations	No. of teachers	Effect size	*p* value
Unadjusted effect size[a]	Comparison	3.39	0.98	18	0.75	.034
	CREATE	4.06	0.80	18		
Adjusted effect size[b]	Comparison	3.39	0.98	18	0.75	.025
	CREATE	4.06	0.80	18		

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.

Correlation Among Responses to Well-Being Log and Survey Responses

Table H40 and Table H41 display the correlations between well-being log and survey responses. We combined the responses across conditions.

TABLE H40. CORRELATIONS BETWEEN WELL-BEING LOG AND SURVEY RESPONSES (COHORT 6, CREATE AND COMPARISON, N = 22)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 Self-Compassion Scale[b]																
2 Opportunities & Support to Think About Social Justice & Racial Equity in Schools[b]	-0.10															
3 Confidence in Understanding Concepts Related to Race & Equity in Schools[b]	0.03	0.65														
4 Teacher Social Capital Scale: Teacher Bonding[b]	0.05	0.50	0.20													
5 Teacher Social Capital Scale: Opportunities to Collaborate[b]	-0.08	0.48	0.46	0.63												
6 Focus & Awareness[a]	0.58	-0.01	0.30	0.22	0.37											
7 Difficulty with attention[a, c]	0.34	-0.33	-0.18	-0.48	-0.06	0.11										
8 mPERMA: Positive Emotions Subscale[a]	0.58	-0.24	0.19	0.02	0.25	0.73	0.36									
9 mPERMA: Meaning Subscale[a]	0.78	-0.04	0.23	0.06	0.16	0.76	0.39	0.78								
10 mPERMA: Accomplishment Subscale[a]	0.76	0.05	0.22	0.12	0.27	0.66	0.39	0.69	0.89							

TABLE H40. CORRELATIONS BETWEEN WELL-BEING LOG AND SURVEY RESPONSES (COHORT 6, CREATE AND COMPARISON, N = 22)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
11 Supported & Hopeful[a]	0.72	-0.42	-0.15	-0.16	-0.10	0.54	0.63	0.76	0.75	0.69						
12 Sense of Equity in Role & Experience[a]	0.65	0.16	0.22	0.29	0.31	0.50	0.25	0.58	0.69	0.75	0.67					
13 MAIA: Self-Regulation[a]	0.55	0.08	0.41	0.14	0.27	0.53	0.27	0.46	0.69	0.65	0.56	0.72				
14 MBI: Emotional Exhaustion[a, c]	-0.62	-0.04	0.08	-0.17	-0.11	-0.67	-0.15	-0.43	-0.65	-0.62	-0.56	-0.54	-0.54			
15 MBI: Depersonalization[a, c]	-0.42	0.00	0.11	-0.16	-0.20	-0.64	-0.22	-0.45	-0.55	-0.45	-0.56	-0.59	-0.48	0.87		
16 MBI: Personal Accomplishment[a]	0.33	-0.05	0.32	-0.06	0.29	0.21	0.52	0.66	0.50	0.45	0.47	0.41	0.29	0.03	-0.05	

Note. The row headers display both a number and the outcome name, while the column headers show only the number. Each number refers to the same outcome across rows and columns. Shading indicates the magnitude of the correlation, regardless of direction; both positive and negative values are included. Turquoise represents correlations of .80–.89, green = .70–.79, and yellow = .60–.69. MAIA = Multidimensional Assessment of Interoceptive Awareness; mPERMA = Momentary Positive Emotions, Engagement, Relationships, Meaning, Accomplishment; and MBI = Maslach Burnout Inventory for Educators.

[a] Indicates a well-being log outcome

[b] Indicates a survey outcome

[c] Indicates scales where a lower score is a favorable outcome (e.g. a lower score on MBI: Emotional Exhaustion is a favorable outcome).

TABLE H41. CORRELATIONS BETWEEN WELL-BEING LOG AND SURVEY RESPONSES (COHORT 7, CREATE AND COMPARISON, N=16)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 Self-Compassion Scale[b]															
2 Confidence in Understanding Concepts Related to Race & Equity in Schools[b]	-0.04														
3 Teacher Social Capital Scale: Teacher Bonding[b]	0.41	0.14													
4 Teacher Social Capital Scale: Opportunities to Collaborate[b]	0.37	0.22	0.77												
5 Focus & Awareness[a]	0.15	0.30	0.29	0.56											
6 Difficulty with Attention[a, c]	0.15	0.02	0.43	0.45	0.71										
7 mPERMA: Positive Emotions Subscale[a]	0.26	0.07	0.15	0.42	0.68	0.63									
8 mPERMA: Meaning Subscale[a]	0.18	-0.09	0.10	0.46	0.61	0.57	0.64								
9 mPERMA: Accomplishment Subscale[a]	-0.07	0.24	-0.06	0.10	0.57	0.54	0.57	0.42							
10 Supported & Hopeful[a]	-0.06	-0.28	0.07	0.35	0.56	0.48	0.68	0.69	0.40						
11 Sense of Equity in Role & Experience[a]	0.55	0.49	0.69	0.82	0.62	0.46	0.44	0.52	0.23	0.23					

TABLE H41. CORRELATIONS BETWEEN WELL-BEING LOG AND SURVEY RESPONSES (COHORT 7, CREATE AND COMPARISON, N=16)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
12 MAIA: Self-Regulation[a]	0.42	0.22	0.32	0.54	0.82	0.58	0.64	0.73	0.39	0.43	0.70				
13 MBI: Emotional Exhaustion[a, c]	-0.44	-0.26	-0.65	-0.73	-0.77	-0.73	-0.59	-0.61	-0.43	-0.51	-0.83	-0.78			
14 MBI: Depersonalization[a, c]	-0.19	-0.33	-0.66	-0.55	-0.42	-0.68	-0.22	-0.19	-0.13	-0.08	-0.48	-0.20	0.56		
15 MBI: Personal Accomplishment[a]	0.12	0.06	0.56	0.67	0.21	0.36	0.17	0.50	0.12	0.23	0.58	0.32	-0.50	-0.18	

Note. The row headers display both a number and the outcome name, while the column headers show only the number. Each number refers to the same outcome across rows and columns. Shading indicates the magnitude of the correlation, regardless of direction; both positive and negative values are included. Turquoise represents correlations of .80–.89, green = .70–.79, and yellow = .60–.69. MAIA = Multidimensional Assessment of Interoceptive Awareness; mPERMA = Momentary Positive Emotions, Engagement, Relationships, Meaning, Accomplishment; and MBI = Maslach Burnout Inventory for Educators.

[a] Indicates a well-being log outcome

[b] Indicates a survey outcome

[c] Indicates scales where a lower score is a favorable outcome (e.g. a lower score on MBI: Emotional Exhaustion is a favorable outcome).

Appendix I. Detailed Analysis of Emotional Exhaustion Impacts (to Support WWC Review)

Emotional Exhaustion is the only teacher-level outcome that aligns to a WWC-reviewable (Version 5.0) domain. That domain is Teacher Well-Being, which, in this study, we evaluated at baseline through the CD-RISC resilience scale. The outcome, Emotional Exhaustion, is a component of Teacher Well-Being.

In this section, we provide information needed to support a WWC review of the results. The results are extensive—across multiple years and cohorts—and provide all information required to conduct a review for the multi-year design.

Information About Measures

CD-RISC is the baseline measure of teacher general well-being. We evaluated impacts on Emotional Exhaustion, which is a subdomain of teacher Well-Being. We describe the two scales in Appendix D. Data Collection Sources.

Information About the Study Design

Table I1 shows the survey occasions for each of four cohorts. For completeness, we include all cohorts for which we have outcomes and that would be eligible for review under WWC. This includes impacts for Cohort 5, which we assessed under a prior grant (Jaciw et al., 2023).

CREATE implementation takes places over a three-year period for each cohort. Teachers in Cohorts 5 and 6 experienced the full length of the intervention during the course of the study. Teachers in Cohort 7 experienced the first two years of the intervention, but outcomes are available for only the first year. (Second-year outcomes were invalid because of a coding error that we could not correct.) Cohort 8 experienced only the first year of the intervention, and outcomes are available for that year.

We administered the CD-RISC at baseline (B), and the Emotional Exhaustion scale four times during each of the three years of implementation.

TABLE I1. SURVEY OCCASIONS FOR WHICH OUTCOMES ARE AVAILABLE ACROSS COHORTS 5–8

	2019–2020	2020–2021	2021–2022	2022–2023
Cohort 5	B, S1, S2, S3, S4	S1, S2, S3, S4	S1, S2, S3, S4	
Cohort 6		B, S1, S2, S3, S4	S1, S2, S3, S4	S1, S2, S3, S4
Cohort 7			B, S1, S2, S3, S4	MISSING
Cohort 8				B, S1, S2, S3, S4

Note. B = baseline survey, S = post-baseline survey

Analysis and Results

To allow maximum flexibility for interpreting impact findings and to support methods of data synthesis, we analyzed impacts on Emotional Exhaustion separately for each of the posttest survey occasions (12 for Cohort 5, 12 for Cohort 6, 4 for Cohort 7, and 4 for Cohort 8).

We matched teachers separately for each of the 32 posttest survey administrations. On all 32 occasions, the pretests achieved baseline equivalence (with a standardized mean difference on the pretest of less than .25, and impact analysis adjusted for the effect of the pretest for each of the 32 analyses). We report baseline equivalence for each of the 32 blocks as the point estimate for the effect associated with the $condition$ variable (coded 1 for CREATE and 0 for control) in the following model, divided by the pooled standard deviation of the outcome variable (CD-RISC).

$$CDRISC_i = \alpha + \beta condition_i + e_i \quad (1)$$

We report impact for each block as the point estimate for the effect associated with the $condition$ variable in the following model, divided by the pooled standard deviation of the outcome variable ($EmotExhaust$).

$$EmotExhaust_i = \alpha + \beta condition_i + \gamma CDRISC_i + e_i \quad (2)$$

We report sample sizes, baseline equivalence, and impacts for each block.

Results also report the correlation matrices between outcomes for each pairwise set of surveys within each cohort. (We assume samples are independent across cohorts; however, within a cohort, the same teachers often provide responses across multiple survey occasions.) The correlations we provided allow calculation of the dependencies in outcomes across survey occasions, which WWC may use during synthesis of impact findings. One can use correlations to adjust the standard error of the impact estimates when synthesizing the results. We state two types of correlations for each cohort.

1. Correlations between outcome $EmotExhaust_i$ across survey occasions
2. Correlations between residualized outcomes (e_i in Equation 2) across survey occasions

In each case, we adjusted correlations for non-overlap in the sample. For example, after matching in every block, the number of teachers in two consecutive blocks in a cohort may be 12 and 14. The samples of 12 and 14 have 5 teachers in common (7 appear only in the sample of 12, and 8 appear only in the sample of 14). We expect non-zero correlations among only the 5 common teachers. The correlations that we reported factor in (a) non-zero correlations for repeated measures for the same persons across occasions, and (b) zero correlations between different persons across occasions. That is, the correlations are weighted to reflect the share of dependent and share of independent observations between samples on any two occasions. Please see Figure I1–Figure I8 and Table I2–Table I9.

Correlation Coefficients for Cohort 5

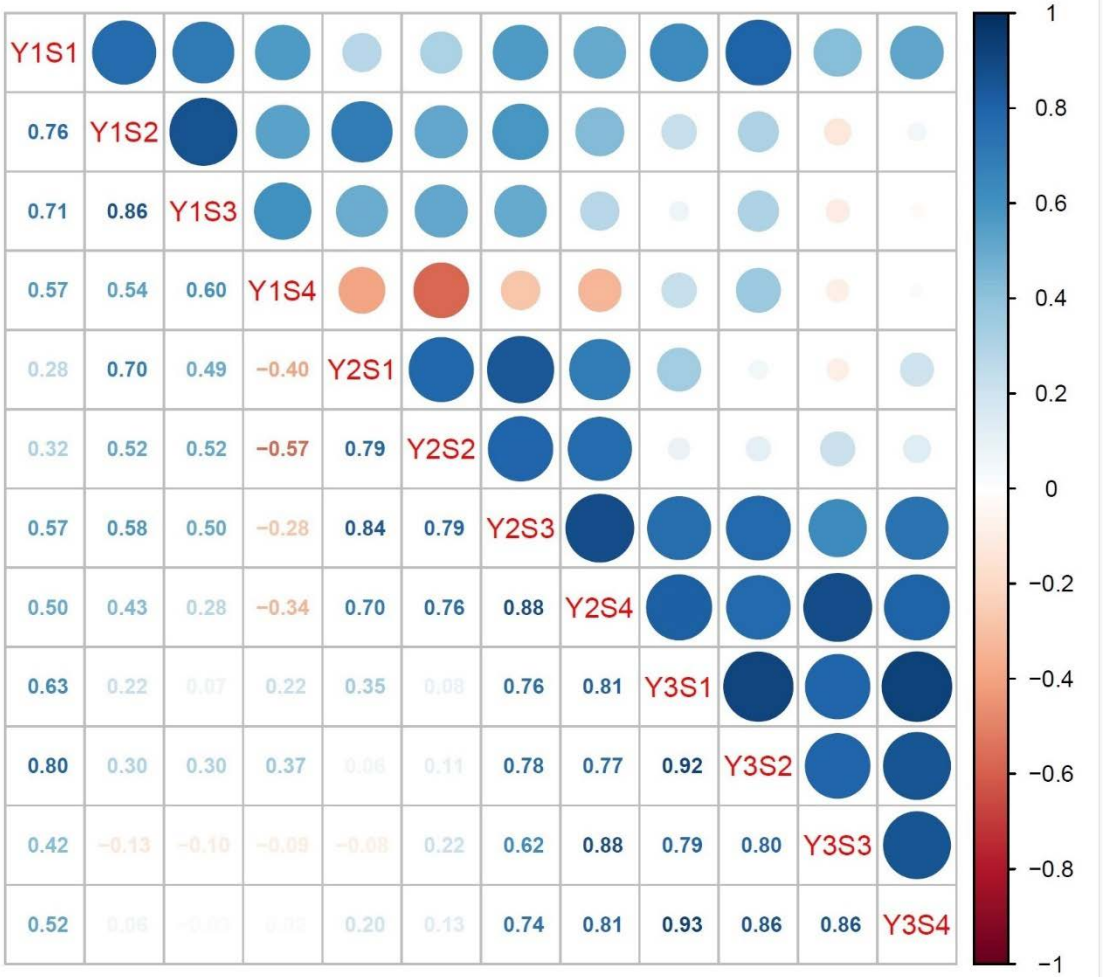


FIGURE I1. CORRELATION OF COHORT 5 OUTCOMES AMONG MATCHED CASES

Note. Y1S1 indicates "Year 1, Survey 1". The remaining diagonal terms follow the same format. Colors indicate larger positive correlations (darker blue) or stronger negative correlations (darker red).

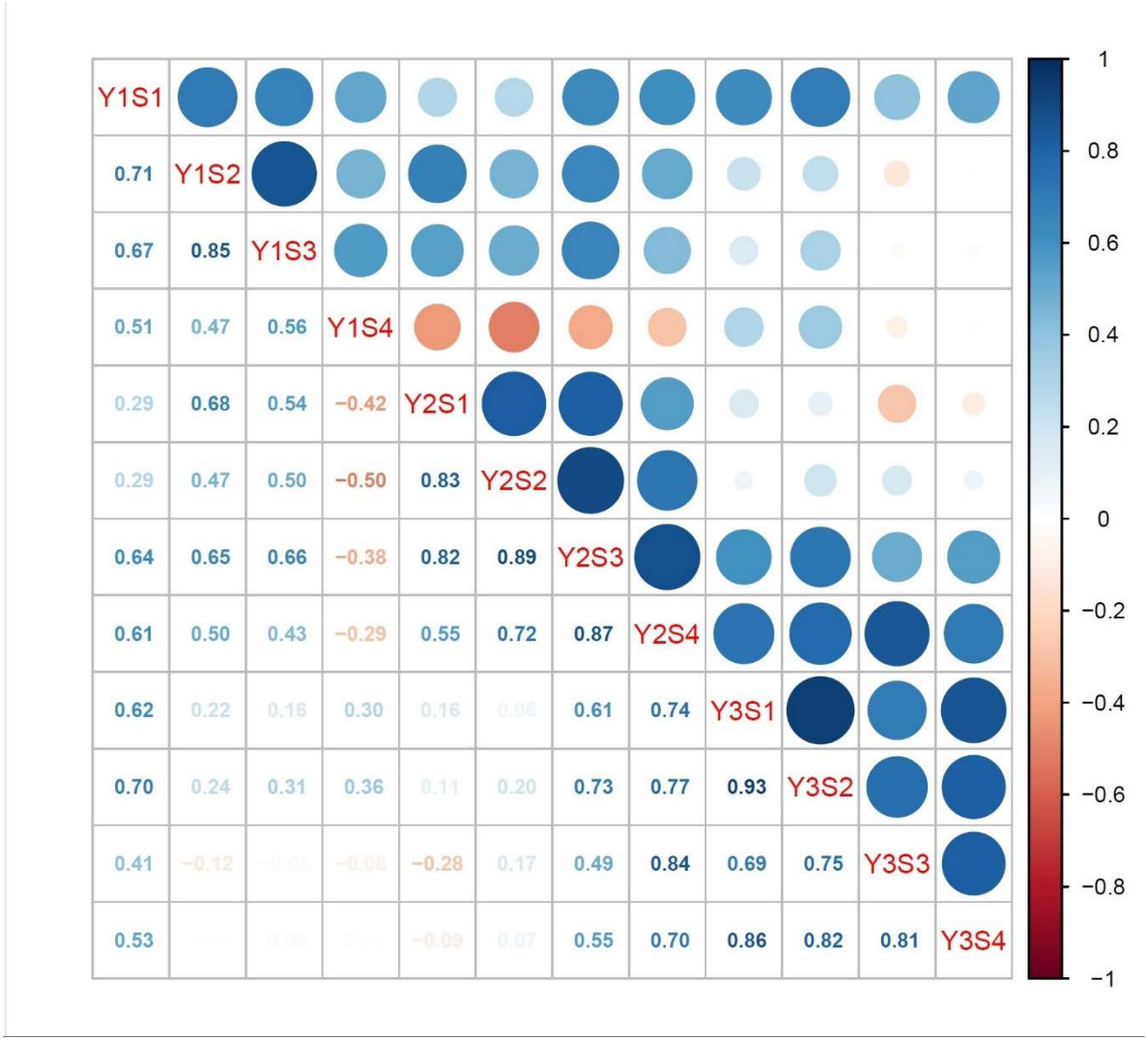


FIGURE I2. CORRELATION OF COHORT 5 RESIDUALIZED OUTCOMES AMONG MATCHED CASES

Note. Y1S1 indicates "Year 1, Survey 1". The remaining diagonal terms follow the same format. Colors are used to indicate larger positive correlations (darker blue) or stronger negative correlations (darker red).

TABLE I2. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

	Data point 1		Data point 2		Data point 3		Data point 4	
	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE
Baseline outcome								
No. of teachers	12	16	12	16	12	16	12	16
Mean	4.44	4.50	4.44	4.50	4.44	4.50	4.44	4.50
Standard deviations	0.50	0.56	0.50	0.56	0.50	0.56	0.50	0.56
Pooled standard deviations	0.53		0.53		0.53		0.53	
Baseline difference	0.06		0.06		0.06		0.06	
Effect size	0.10		0.10		0.10		0.10	
Outcomes								
No. of teachers	12	16	12	16	12	16	12	16
Mean	3.96	3.13	4.18	3.17	3.97	3.36	3.53	3.04
Standard deviations	1.57	1.22	1.46	1.40	1.62	1.28	1.65	1.51
Pooled standard deviations	1.38		1.43		1.43		1.57	
Impact	-0.79		-0.95		-0.57		-0.44	
Effect size	-0.57		-0.66		-0.40		-0.28	

TABLE I3. RESULTS FOR COHORT 5 YEAR 2 MATCHED CASES OF TEACHERS

	Data point 1		Data point 2		Data point 3		Data point 4	
	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE
Baseline outcome								
No. of teachers	2	9	2	9	3	9	3	9
Mean	4.67	4.63	4.67	4.63	4.67	4.63	4.67	4.62
Standard deviations	0.47	0.31	0.47	0.31	0.33	0.31	0.33	0.33
Pooled standard deviations	0.33		0.33		0.31		0.33	
Baseline difference	-0.04		-0.04		-0.04		-0.04	
Effect size	-0.11		-0.11		-0.12		-0.13	
Outcomes								
No. of teachers	2	9	2	9	3	9	3	8
Mean	3.22	2.77	2.89	2.83	4.00	2.72	4.09	2.29
Standard deviations	1.89	0.98	1.73	0.81	2.21	0.96	1.96	0.62
Pooled standard deviations	1.12		0.96		1.31		1.07	
Impact	-0.43		-0.07		-1.22		-1.77	
Effect size	-0.38		-0.07		-0.93		-1.65	

TABLE I4. RESULTS FOR COHORT 5 YEAR 3 MATCHED CASES OF TEACHERS

	Data point 1		Data point 2		Data point 3		Data point 4	
	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE
Baseline outcome								
No. of teachers	5	8	4	8	4	8	4	8
Mean	4.60	4.58	4.50	4.58	4.50	4.58	4.50	4.58
Standard deviations	0.43	0.30	0.43	0.30	0.43	0.30	0.43	0.30
Pooled standard deviations	0.35		0.34		0.34		0.34	
Baseline difference	-0.02		0.08		0.08		0.08	
Effect size	-0.05		0.24		0.24		0.24	
Outcomes								
No. of teachers	5	8	4	8	4	8	4	8
Mean	4.49	3.61	4.58	4.22	4.86	3.76	5.06	3.51
Standard deviations	1.43	1.07	1.79	1.03	1.65	1.01	1.65	1.47
Pooled standard deviations	1.21		1.31		1.24		1.53	
Impact	-0.86		-0.39		-1.19		-1.71	
Effect size	-0.71		-0.30		-0.96		-1.12	

Correlation Coefficients for Cohort 6

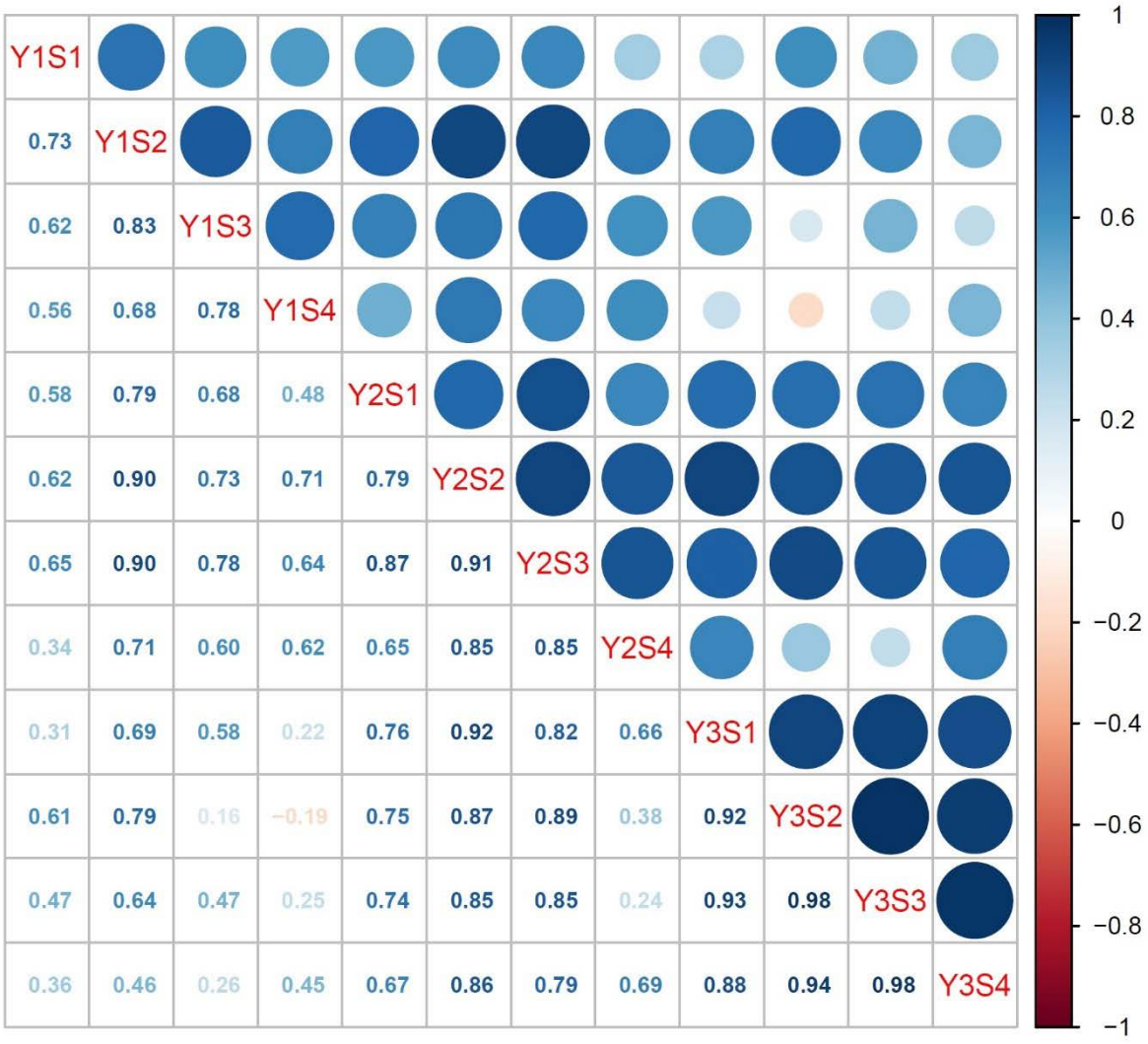


FIGURE I3. CORRELATIONS OF COHORT 6 OUTCOMES AMONG MATCHED CASES

Note. Y1S1 indicates "Year 1, Survey 1". The remaining diagonal terms follow the same format. Colors are used to indicate larger positive correlations (darker blue) or stronger negative correlations (darker red).

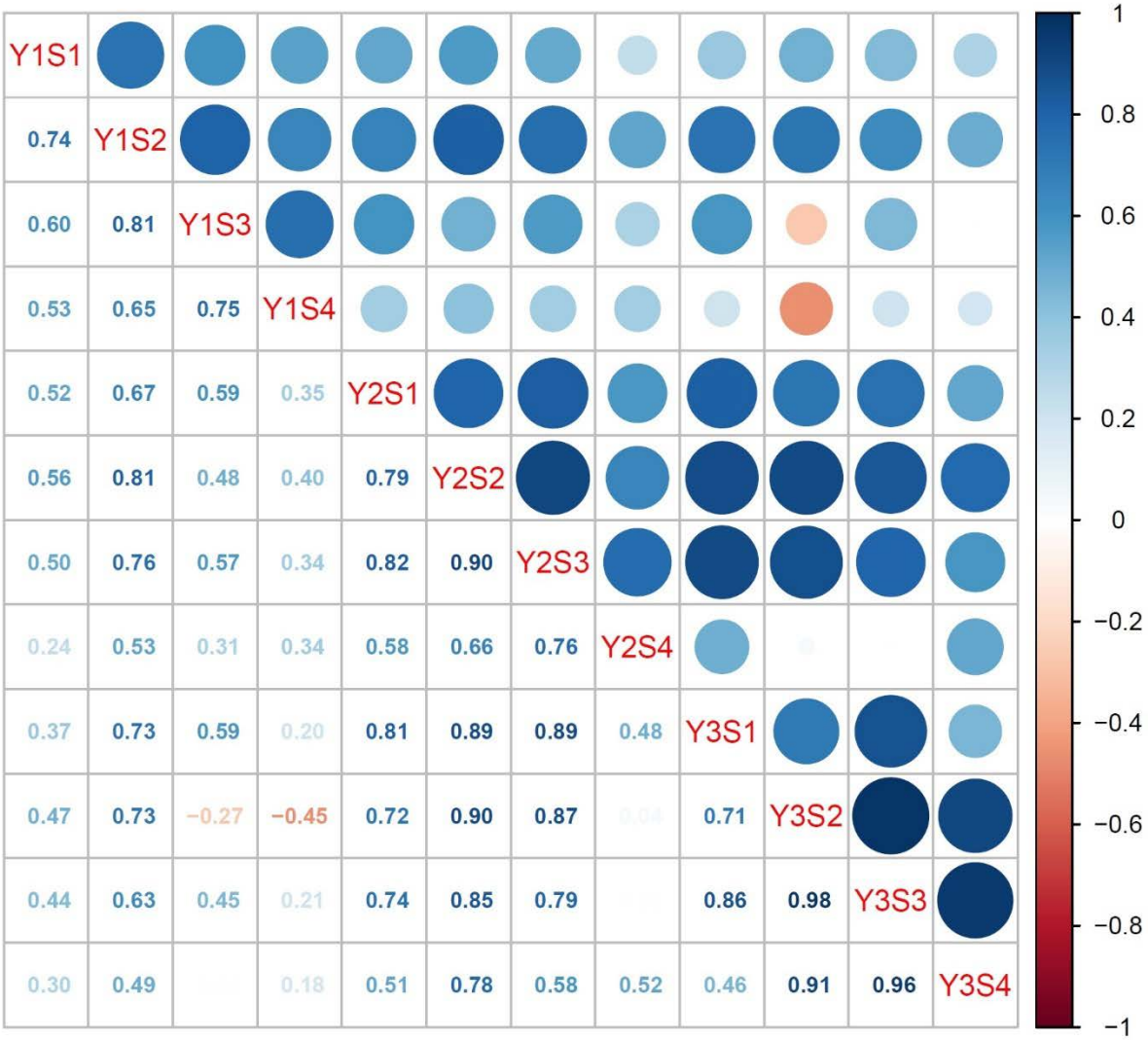


FIGURE I4. CORRELATION OF COHORT 6 RESIDUALIZED OUTCOMES AMONG MATCHED CASES

Note. Y1S1 indicates "Year 1, Survey 1". The remaining diagonal terms follow the same format. Colors are used to indicate larger positive correlations (darker blue) or stronger negative correlations (darker red).

TABLE I5. RESULTS FOR COHORT 6 YEAR 1 MATCHED CASES OF TEACHERS

	Data point 1		Data point 2		Data point 3		Data point 4	
	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE
Baseline outcome								
No. of teachers	18	16	17	15	23	15	22	15
Mean	4.09	4.12	4.12	4.13	4.06	4.13	4.08	4.13
Standard deviations	0.61	0.44	0.62	0.45	0.62	0.45	0.63	0.45
Pooled standard deviations	0.54		0.55		0.56		0.57	
Baseline difference	0.03		0.02		0.08		0.06	
Effect size	0.06		0.03		0.13		0.10	
Outcomes								
No. of teachers	18	16	17	15	23	15	22	15
Mean	3.25	2.53	2.93	2.59	2.95	2.51	2.79	2.44
Standard deviations	1.32	0.80	1.29	1.16	1.21	0.99	1.21	0.94
Pooled standard deviations	1.10		1.23		1.13		1.11	
Impact	-0.71		-0.33		-0.39		-0.32	
Effect size	-0.64		-0.27		-0.34		-0.29	

TABLE I6. RESULTS FOR COHORT 6 YEAR 2 MATCHED CASES OF TEACHERS

	Data point 1		Data point 2		Data point 3		Data point 4	
	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE
Baseline outcome								
No. of teachers	9	8	9	9	7	8	11	8
Mean	4.00	3.96	4.04	4.04	4.00	3.96	4.06	4.00
Standard deviations	0.73	0.45	0.72	0.48	0.82	0.45	0.68	0.50
Pooled standard deviations	0.61		0.61		0.65		0.61	
Baseline difference	-0.04		-0.00		-0.04		-0.06	
Effect size	-0.07		-0.00		-0.06		-0.10	
Outcomes								
No. of teachers	9	8	9	9	7	8	11	8
Mean	4.75	3.26	4.94	3.65	5.63	3.69	4.63	3.49
Standard deviations	1.44	1.32	1.45	1.65	1.62	1.98	1.79	1.74
Pooled standard deviations	1.39		1.56		1.82		1.77	
Impact	-1.51		-1.28		-1.99		-1.24	
Effect size	-1.09		-0.82		-1.09		-0.70	

TABLE I7. RESULTS FOR COHORT 6 YEAR 3 MATCHED CASES OF TEACHERS

	Data point 1		Data point 2		Data point 3		Data point 4	
	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE
Baseline outcome								
No. of teachers	6	5	2	4	4	6	3	5
Mean	4.28	4.20	4.17	4.17	4.33	4.28	4.33	4.27
Standard deviations	0.39	0.30	0.24	0.33	0.27	0.33	0.33	0.37
Pooled standard deviations	0.35		0.31		0.31		0.35	
Baseline difference	-0.08		0.00		-0.06		-0.07	
Effect size	-0.22		0.00		-0.18		-0.19	
Outcomes								
No. of teachers	6	5	2	4	4	6	3	5
Mean	4.19	4.42	4.72	3.98	4.44	4.15	4.41	3.69
Standard deviations	1.07	1.53	0.08	1.64	0.48	1.82	0.68	1.35
Pooled standard deviations	1.29		1.42		1.47		1.17	
Impact	0.26		-0.74		-0.33		-0.82	
Effect size	0.20		-0.52		-0.22		-0.70	

Correlation Coefficients for Cohort 7



FIGURE I5. CORRELATIONS OF COHORT 7 OUTCOMES AMONG MATCHED CASES

Note. Due to technical limitations unrelated to treatment status, correlations involving Year 2 outcomes are not available.



FIGURE I6. CORRELATIONS OF COHORT 7 RESIDUALIZED OUTCOMES AMONG MATCHED CASES

Note. Due to technical limitations unrelated to treatment status, correlations involving Year 2 outcomes are not available.

TABLE I8. RESULTS FOR COHORT 7 YEAR 1 MATCHED CASES OF TEACHERS

	Data point 1		Data point 2		Data point 3		Data point 4	
	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE
Baseline outcome								
No. of teachers	6	20	5	21	6	20	9	21
Mean	3.94	4.07	4.07	4.03	3.94	4.00	4.04	4.03
Standard deviations	0.39	0.72	0.28	0.72	0.39	0.73	0.54	0.72
Pooled standard deviations	0.67		0.67		0.67		0.67	
Baseline difference	0.12		-0.03		0.06		-0.01	
Effect size	0.18		-0.05		0.08		-0.01	
Outcomes								
No. of teachers	6	20	5	21	6	20	9	21
Mean	3.50	3.14	3.11	3.20	3.59	3.42	3.16	2.87
Standard deviations	1.94	1.32	1.92	1.40	1.53	1.63	1.49	1.56
Pooled standard deviations	1.47		1.50		1.61		1.54	
Impact	-0.29		0.08		-0.16		-0.29	
Effect size	-0.20		0.05		-0.10		-0.19	

Correlation Coefficients for Cohort 8

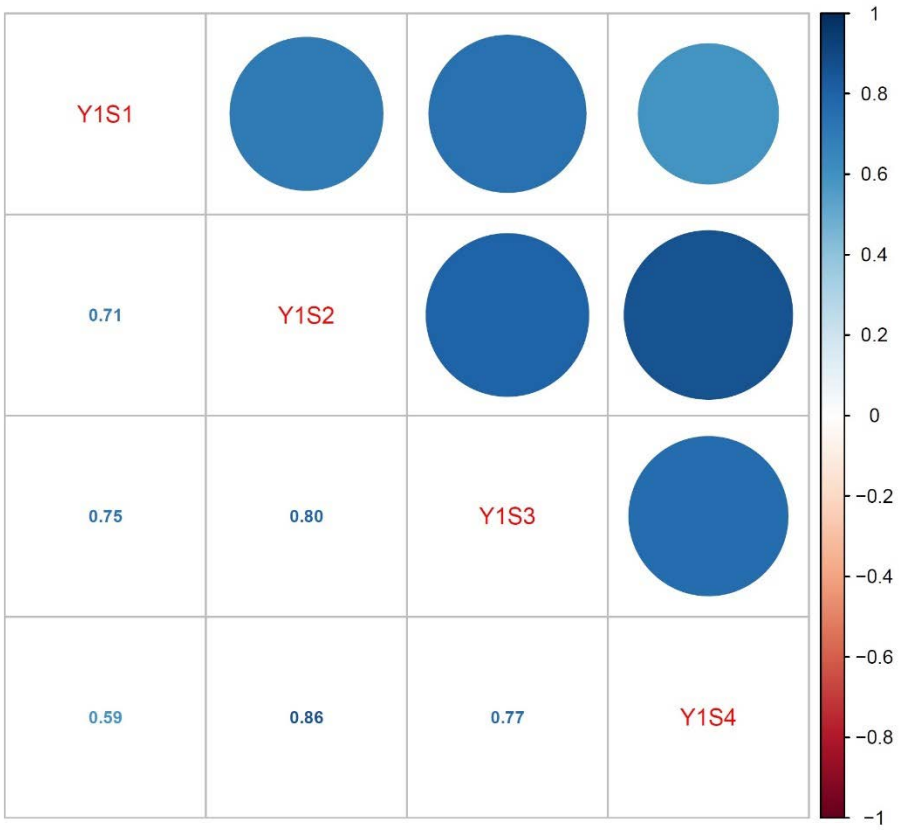


FIGURE I7. CORRELATIONS OF COHORT 8 OUTCOMES AMONG MATCHED CASES

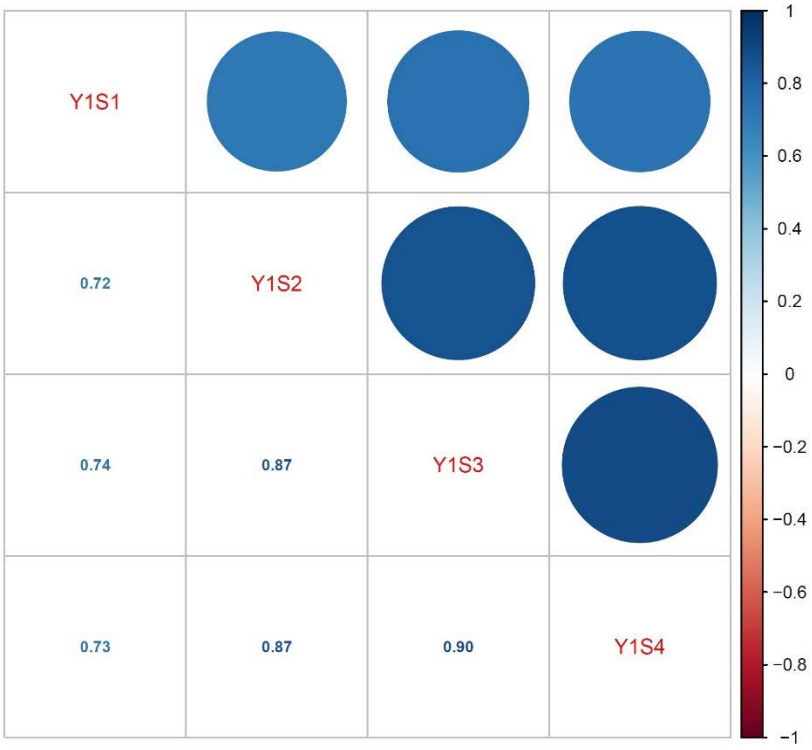


FIGURE I8. CORRELATIONS OF COHORT 8 RESIDUALIZED OUTCOMES AMONG MATCHED CASES

TABLE I9. RESULTS FOR COHORT 8 YEAR 1 MATCHED CASES OF TEACHERS

	Data point 1		Data point 2		Data point 3		Data point 4	
	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE	Comparison	CREATE
Baseline outcome								
No. of teachers	12	11	13	11	14	9	13	9
Mean	3.75	3.91	3.77	3.91	3.81	3.93	3.79	3.93
Standard deviations	0.57	0.72	0.55	0.72	0.55	0.55	0.57	0.55
Pooled standard deviations	0.64		0.63		0.55		0.56	
Baseline difference	0.16		0.14		0.12		0.13	
Effect size	0.25		0.22		0.21		0.23	
Outcomes								
No. of teachers	12	11	13	11	14	9	13	9
Mean	3.83	3.76	3.75	3.59	3.96	3.11	3.51	3.04
Standard deviations	1.47	1.64	1.42	1.28	1.43	1.61	1.24	1.34
Pooled standard deviations	1.55		1.35		1.50		1.28	
Impact	-0.02		-0.09		-0.87		-0.33	
Effect size	-0.01		-0.07		-0.58		-0.26	

Appendix J. Retention Data Sources, Coding, and Decision Tree

Appendix J provides a visual representation of the data sources used in the retention analysis, along with the codes applied to each participant case and the decision tree to determine if a participant was retained at each of the three timepoints: (1) graduation from GSU CEHD, (2) employment as a first-year teacher-of-record, and (3) retention into the second year of teaching. Figure J1 summarizes the data sources for the retention analysis.

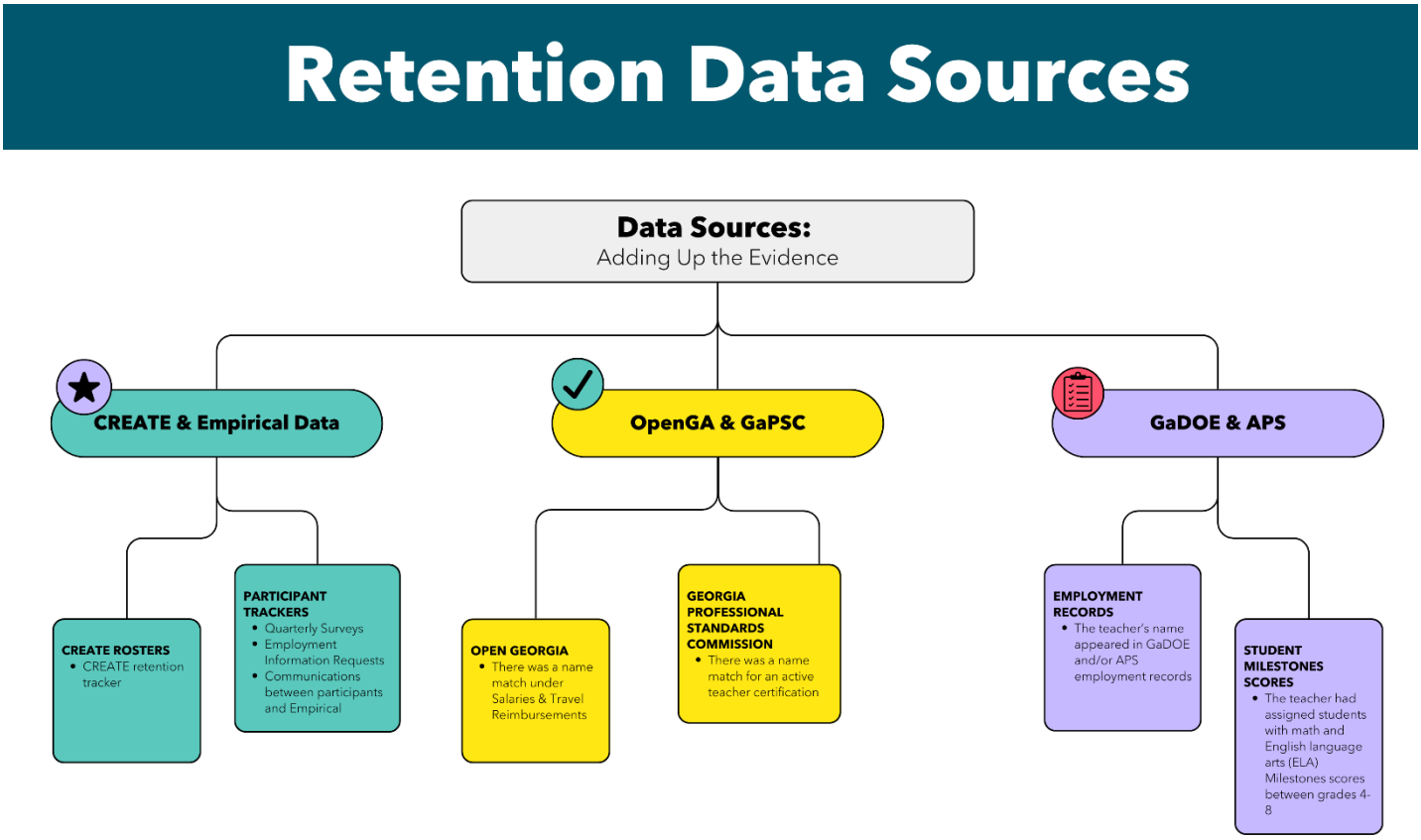


FIGURE J1. DATA SOURCES FOR THE RETENTION ANALYSIS

Next, Figure J2 summarizes the codes applied to cases in the treatment and comparison groups.

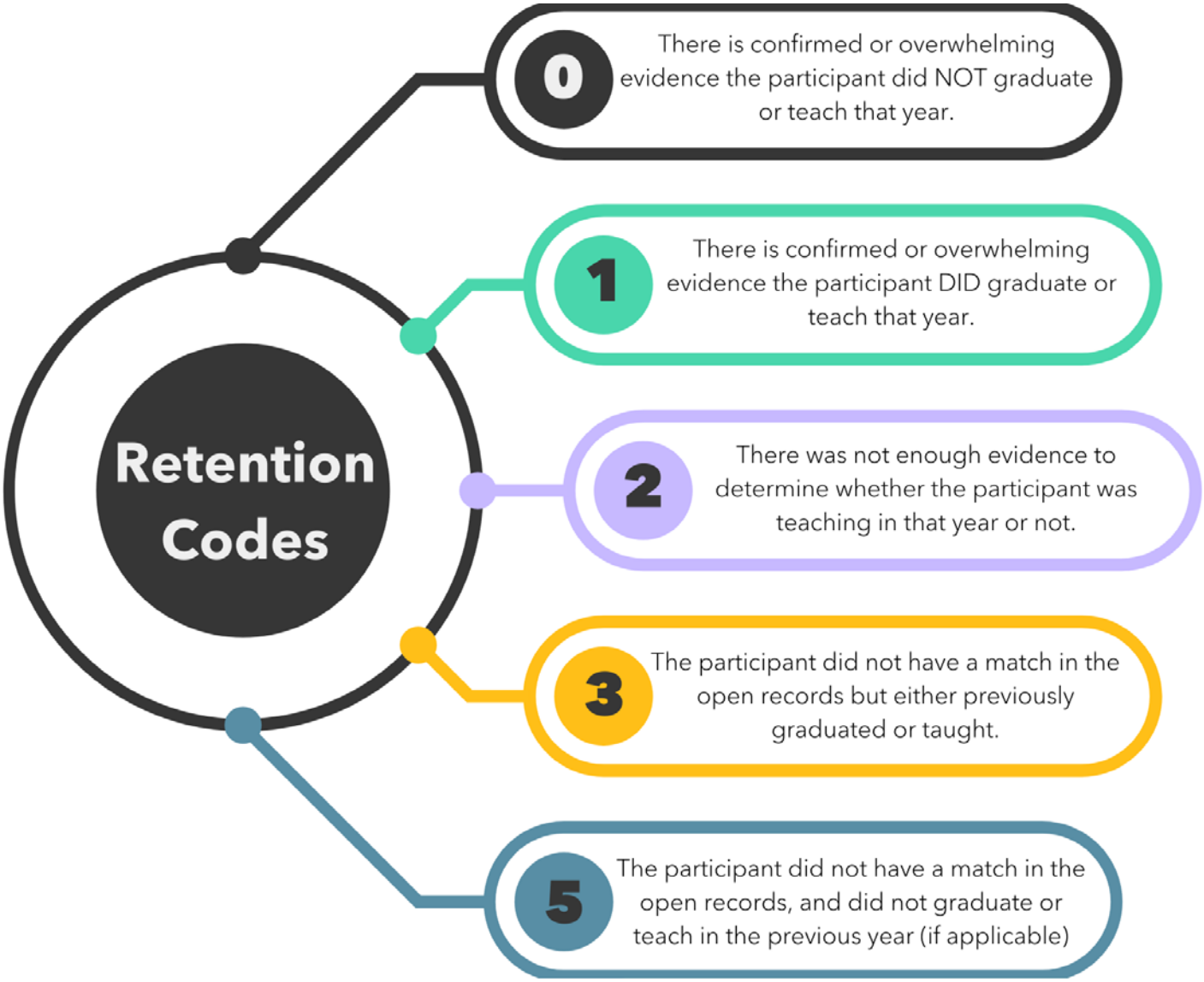


FIGURE J2. CODES APPLIED TO CASES IN THE RETENTION ANALYSIS

Next, Figure J3 shows the decision tree used to determine if a participant graduated from GSU CEHD.

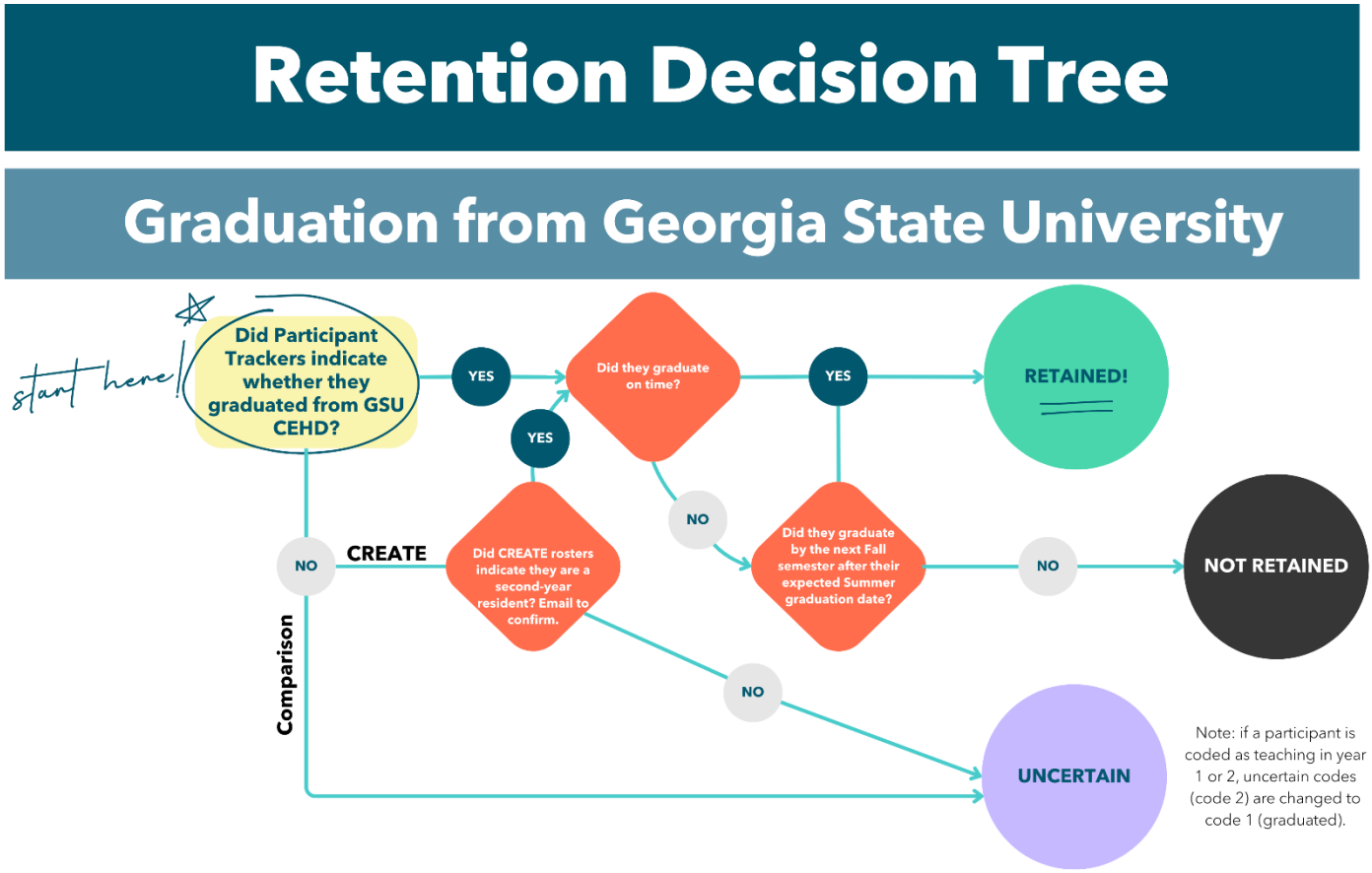


FIGURE J3. DECISION TREE TO DETERMINE IF A PARTICIPANT GRADUATED FROM GEORGIA STATE UNIVERSITY

Last, Figure J4 illustrates the decision tree to determine if a participant was teaching in year 1 and year 2.

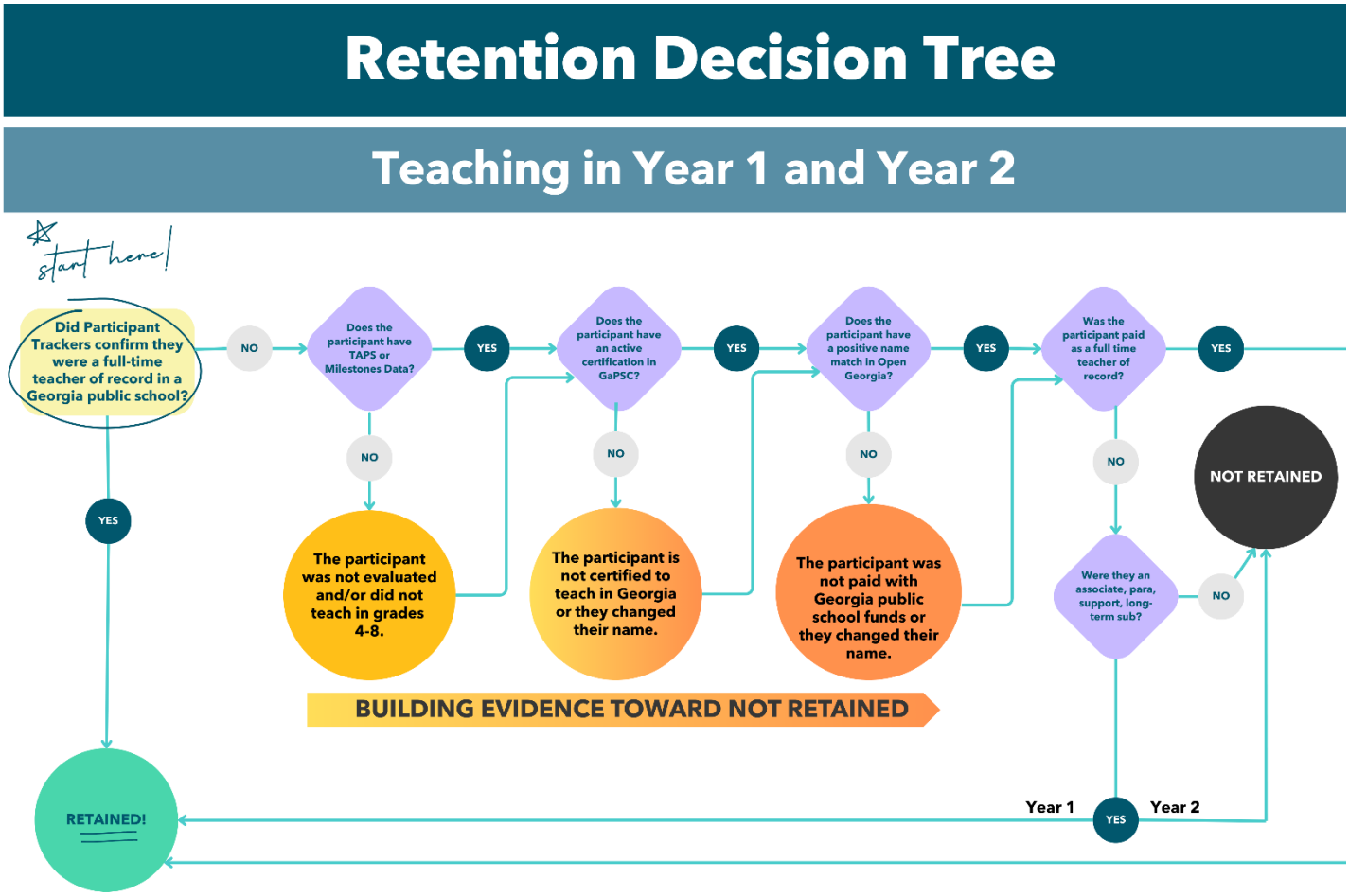


FIGURE J4. DECISION TREE TO DETERMINE IF A PARTICIPANT WAS TEACHING IN YEARS 1 AND 2

Appendix K. Technical Details on Retention Analysis

Survey 1 Matching Method

We matched on propensity scores informed by the following covariates.

- Motivation for Entering Teaching
- Confidence in General Teaching Skills
- CD-RISC (resilience)
- Cohort

We estimated propensity scores using a logistic regression. We used the default matching method (nearest neighbor, 1:1 matching without replacement) to match each CREATE participant to a similar comparison group member. After matching, we assessed balance diagnostics (e.g., standardized mean differences, variance ratios) to verify adequate reduction of imbalance on covariates between conditions. See Table K1.

Survey 2 Baseline Equivalence

TABLE K1. SUMMARY OF BASELINE EQUIVALENCE TESTS FOR ANALYSIS OF IMPACTS ON RETENTION OUTCOMES

Scale	Sample size (CREATE)	Raw score mean (CREATE)	Raw score standard deviation (CREATE)	Sample size (comparison)	Raw score mean (comparison)	Raw score standard deviation (comparison)	Regression adjusted difference estimate	Standardized mean difference[a]
Motivation for Entering Teaching	56	4.67	0.28	56	4.60	0.35	0.07	0.23
Confidence in General Teaching Skills	56	4.05	0.60	56	3.94	0.53	0.11	0.20
CD-RISC	56	4.09	0.65	56	4.00	0.67	0.09	0.14

Note. N = 112. CD-RISC = Connor-Davidson Resilience scale.

[a] The standardized mean difference is the regression-adjusted difference estimate based on a model with dummy variables to indicate cohort membership and a variable indicating the treatment condition, divided by the pooled standard deviation of the pretest distribution.

Survey 3 Algebraic Specification and Raw Result of Impact Models

The linear probability model estimates the probability of (a) graduating, (b) becoming a teacher-of- record in a Georgia public school the year after graduation, and (c) continuing to be a teacher-of-record in a Georgia public school in the second year after graduation from the program:

$$p_i = P(STATUS_i = 1|X_i) = \beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \beta_3 X_{3j} + \beta_4 X_{4j} + \beta_5 X_{5j} + \varepsilon_i$$

Where,

p_i = the probability of a student reaching the milestone ((a), (b), or (c), above) from the program for the i^{th} participant

$STATUS_i$= whether a student reaches (=1) or does not reach the milestone (=0)

X_{1j} = whether participant i is a member of the CREATE (=1) or comparison (=0) group

X_{2j} = motivation to enter teaching for participant i

X_{3j} = confidence in general teaching skills for participant i

X_{4j} = CD-RISC (resilience) for participant i

X_{5j} = a classification variable indicating the cohort for participant i

ε_i = a residual error term for participant i

β_0 = the intercept

β_2 = the effect of motivation to enter teaching

β_3 = the effect of confidence in general teaching

β_4 = the effect of CD-RISC (resilience)

β_5 = the effect(s) of cohort (With three cohorts assessed for the graduation outcome, Cohort 8 is the reference cohort, and β_5 is a vector of effect estimates for Cohorts 6 and 7; for retention into the first year of teaching, Cohort 7 is the reference cohort and β_5 is a scalar effect estimate for Cohort 6; we did not estimate retention into the second year of teaching β_5 because the entire sample belong to Cohort 6.)

The analysis counts may vary slightly from the sample numbers used for baseline equivalence tests. This is because we omitted cases with missing outcomes from the impact analysis. For example, the full baseline sample includes 112 teachers, whereas the analysis of the effect of CREATE on graduation includes 107 teachers. Five comparison group teachers are not in the retention analysis because their graduation status is unknown. See Table K2–Table K7.

Impact on Probability of Graduation

TABLE K2. FIXED EFFECTS IN REGRESSION-ADJUSTED IMPACT ANALYSIS OF CREATE ONGRADUATION

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	1.2463	0.3684	100	3.38	0.0010
Impact on CREATE[a]	0.007471	0.04823	100	0.15	0.8772
Confidence in teaching	-0.05312	0.04806	100	-1.11	0.2717
Motivation for teaching	-0.01213	0.07975	100	-0.15	0.8794
CD-RISC	-0.01355	0.04093	100	-0.33	0.7412
Cohort 6	0.05481	0.06155	100	0.89	0.3753
Cohort 7	-0.03374	0.05974	100	-0.56	0.5735

Note. We include 107 residents across Cohorts 6, 7, and 8 in this analysis. CD-RISC = Connor-Davidson Resilience scale.

[a] The impact is assessed at timepoint 1: graduation (cohorts 6, 7, and 8)

TABLE K3. RANDOM EFFECT IN REGRESSION-ADJUSTED IMPACT ANALYSIS OF CREATE ON GRADUATION

Cov Parm	Estimate	Standard error	Z Value	Pr > Z
Residual	0.05873	0.008030	7.31	<.0001

Impact on Probability of Being a Teacher of Record in the First Year After Graduation

TABLE K4. FIXED EFFECTS IN REGRESSION-ADJUSTED IMPACT ANALYSIS OF CREATE ON RETENTION IN TEACHING INTO THE FIRST YEAR

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	0.6132	0.6293	75	0.97	0.3330
Impact on CREATE[a]	-0.00824	0.08347	75	-0.10	0.9216
Confidence in teaching	-0.03130	0.08835	75	-0.35	0.7241
Motivation for teaching	0.05969	0.1270	75	0.47	0.6397
CD-RISC	0.008169	0.07107	75	0.11	0.9088
Cohort 6	0.09199	0.08402	75	1.09	0.2770

Note. We include 81 residents across Cohorts 6 and 7 in this analysis. CD-RISC = Connor-Davidson Resilience scale.

[a] The impact is assessed at timepoint 2: teaching in Year 1 (cohorts 6 and 7)

TABLE K5. RANDOM EFFECT IN REGRESSION-ADJUSTED IMPACT ANALYSIS OF CREATE ONRETENTION IN TEACHING INTO THE FIRST YEAR

Cov Parm	Estimate	Standard error	Z Value	Pr > Z
Residual	0.1324	0.02080	6.36	<.0001

Impact on Probability of Being a Teacher of Record in the Second Year After Graduation

TABLE K6. FIXED EFFECTS IN REGRESSION-ADJUSTED IMPACT ANALYSIS OF CREATE ON RETENTION IN TEACHING INTO THE SECOND YEAR

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	0.6980	1.1721	37	.60	0.5551
Impact on CREATE[a]	-0.05087	0.1335	37	-0.38	0.7053
Confidence in teaching	0.06084	0.1482	37	0.41	0.6839
Motivation for teaching	-0.2238	0.2123	37	-1.05	0.2987
CD-RISC	0.2014	0.1301	37	1.55	0.1300

Note. We include 42 residents across Cohorts 6 in this analysis. CD-RISC = Connor-Davidson Resilience scale.

[a] The impact is assessed at timepoint 3: teaching in Year 2 (cohorts 6)

TABLE K7. RANDOM EFFECT IN REGRESSION-ADJUSTED IMPACT ANALYSIS OF CREATE ON RETENTION IN TEACHING INTO THE SECOND YEAR

Cov Parm	Estimate	Standard error	Z Value	Pr > Z
Residual	0.1789	0.03903	4.58	<.0001

RESULTS INCLUDING MAIN EFFECT OF TEACHER RACE

The following three results include estimates of the main effect of minoritized status. The effect "Difference in outcomes for teachers who self-identify as Black" indicates the added-value probability of retention associated with being Black, controlling for the effect of other variables, and irrespective of whether a teacher is in CREATE or comparison. At each of the three stages—graduation, retention into the first of teaching, and retention into the second year of teaching—Black teachers are more likely to continue teaching than their non-Black colleagues. See Table K8—Table K13.

Graduated

TABLE K8. FIXED EFFECTS (INCLUDING RACE) IN REGRESSION-ADJUSTED IMPACT ANALYSIS OF CREATE ON GRADUATION

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	1.1788	0.3842	92	3.07	0.0028
Impact of CREATE[a]	-0.01930	0.05208	92	-0.37	0.7118
Confidence in teaching	-0.05252	0.04927	92	-1.07	0.2892
Motivation for teaching	-0.00054	0.08305	92	-0.01	0.9948
CD-RISC	-0.02364	0.04401	92	-0.54	0.5925
Cohort 6	0.05576	0.06313	92	0.88	0.3794
Cohort 7	-0.04106	0.06189	92	-0.66	0.5088
Difference in outcomes for teachers who self-identify as Black	0.1198	0.05066	92	2.36	0.0202

Note. We include 100 residents across Cohorts 6, 7, and 8 in this analysis. CD-RISC = Connor-Davidson Resilience scale.

[a] The impact is assessed at timepoint 1: graduation (cohorts 6, 7, and 8)

TABLE K9. RANDOM EFFECT IN REGRESSION-ADJUSTED IMPACT ANALYSIS OF CREATE ON GRADUATION

Cov Parm	Estimate	Standard error	Z Value	Pr > Z
Residual	0.05897	0.008340	7.07	<.0001

Teaching into First Year

TABLE K10. FIXED EFFECTS (INCLUDING RACE) IN REGRESSION-ADJUSTED IMPACT ANALYSIS OF CREATE ON TEACHING INTO THE FIRST YEAR

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	0.1997	0.6456	67	0.31	0.7581
Impact of CREATE[a]	-0.07268	0.08875	67	-0.82	0.4157
Confidence in teaching	-0.01958	0.08744	67	-0.22	0.8235
Motivation for teaching	0.1070	0.1262	67	0.85	0.3993
CD-RISC	0.02045	0.07527	67	0.27	0.7867
Cohort 6	0.1255	0.08446	67	1.49	0.1421
Difference in outcomes for teachers who self-identify as Black	0.2048	0.08589	67	2.38	0.0199

Note. We include 74 residents across Cohorts 6 and 7 in this analysis. CD-RISC = Connor-Davidson Resilience scale.

[a] The impact is assessed at timepoint 2: teaching in Year 1 (cohorts 6 and 7)

TABLE K11. RANDOM EFFECT IN REGRESSION-ADJUSTED IMPACT ANALYSIS OF CREATE ON TEACHING INTO THE FIRST YEAR

Cov Parm	Estimate	Standard error	Z Value	Pr > Z
Residual	0.1207	0.01984	6.08	<.0001

Teaching into Second Year

TABLE K12. FIXED EFFECTS (INCLUDING RACE) IN REGRESSION-ADJUSTED IMPACT ANALYSIS OF CREATE ON TEACHING INTO THE SECOND YEAR

Fixed effects	Estimate	Standard error	*df*	t-value	Pr > |t|
Intercept	0.1165	0.9935	33	0.12	0.9074
Impact of CREATE[a]	-0.1619	0.1148	33	-1.41	0.1679
Confidence in teaching	0.1388	0.1236	33	1.12	0.2694
Motivation for teaching	-0.1798	0.1762	33	-1.02	0.3151
CD-RISC	0.1693	0.1091	33	1.55	0.1302
Difference in outcomes for teachers who self-identify as Black	0.4667	0.1153	33	4.05	0.0003

Note. We include 39 residents across Cohorts 6 in this analysis. CD-RISC = Connor-Davidson Resilience scale.

[a] The impact is assessed at timepoint 3: teaching in Year 2 (cohorts 6)

TABLE K13. RANDOM EFFECT IN REGRESSION-ADJUSTED IMPACT ANALYSIS OF CREATE ON TEACHING INTO THE SECOND YEAR

Cov Parm	Estimate	Standard error	Z Value	Pr > Z
Residual	0.1203	0.02724	4.42	<.0001

Appendix L. Interview Protocol

CREATE Interview Protocol

2022–23 School Year

Participant Background Information

- Date / Time of Interview:
- Interviewer:
- Notetaker:
- Link to Jamboard:
- Participant Name:
- Participant Study:
- Participant Cohort:
- Professional Role:
- Years in Professional Role:
- Grades (If Teacher):
- School:
- CREATE Participation at School:
- Years at School:
- Previous Schools, If Applicable:
- CREATE Participation at Previous Schools, If Applicable:
- # of CREATE PL Opportunities Attended:
- Date (Month / Year) of Most Recent PL:
- Race / Ethnicity:
- Gender:

Questions	Instructions
GENERAL BACKGROUND QUESTIONS • **In what ways have you been associated with CREATE?** • **How long have you been associated with CREATE?** • **Why did you decide to join CREATE?**	**Note**, spend only a few minutes on these questions. This gathers some quick contextual information about who we are talking to, and eases the interviewee into the interview. **Notetaker**, copy and paste each question into the chat, for this section and for the remaining sections.

Questions	Instructions
WHAT DOES IT MEAN TO "BE" CREATE? • **Based on your experiences with CREATE, how would you describe CREATE's mission and vision?** • **For you in your specific role, and based on your experiences with CREATE, what does it look like to implement and embody CREATE's mission and vision?** • **Imagine an educator approached you and introduced themselves as a fellow member of the CREATE community.** • **What values or beliefs would you assume guide their work? Why?** • **What would you assume about their [teaching, mentorship, or leadership] approaches? Why?**	**Notetaker**, capture examples of CREATE mindset and behaviors the interviewee offers in a Jamboard that we can share when we ask questions in the next two sections. **Interviewer**, probe for specific examples of a specific mindset and actions that indicate you are a CREATE educator. Ideally, we also want examples of doing and implementing.
SCHOOL-LEVEL SUPPORTS FOR IMPLEMENTING CREATE • **How does your school support you in implementing CREATE's mission and vision, if at all?** • **Resources (training, funding, time)** • **Context (school climate, school culture, priority areas, quality of relationships)** • **Individuals (principal, coach/mentor, colleagues)** • **Teams (grade-level team, subject-area team)** • **District (engagement with CREATE, advocacy for CREATE)**	**Interviewer**, orient the interviewee to Jamboard so we can construct & visualize the interviewee's responses. Check our understanding along the way using what's written in the Jamboard. Invite the interviewee to help construct the Jamboard. **Notetaker**, capture supports the interviewee mentions in the Jamboard. Use arrows to link ideas (link behavior with support). **Interviewer**, probe for as many of the areas listed under this question as you can elicit (or that the interviewee has meaningful feedback for).
SCHOOL-LEVEL BARRIERS TO IMPLEMENTING CREATE • **What are barriers to implementing CREATE's mission and vision?** • **Resources (training, funding, time)** • **Context (school climate, school culture, priority areas, quality of relationships)** • **Individuals (principal, coach/mentor, colleagues)** • **Teams (grade-level team, subject-area team)** • **District (engagement with CREATE, advocacy for CREATE)** • **Are there any aspects of CREATE's mission and vision that you are able to implement in your individual classroom without receiving colleague or school-level support?** • **Are there any aspects of CREATE's mission and vision that you limit in your classroom or school, due to the broader context of teaching in Georgia (e.g., House Bill 1084 limiting so-called "divisive topics" such as white supremacy and racial oppression)?**	**Interviewer**, check our understanding along the way using what's written in the Jamboard. Invite interviewee to help construct the Jamboard. **Notetaker**, capture barriers the interviewee mentions in the Jamboard. Use arrows to link ideas (link behavior with barrier). **Interviewer**, probe for as many of the areas listed under this question as you can elicit (or that the interviewee has meaningful feedback for).

Appendix M. Interview Coding Scheme

Table M1 presents codes derived from thematic analysis, along with their description and illustrative participant quotes.

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Classroom barriers		
Educational resources on divisive topics	Participants described difficulty locating age-appropriate resources on potentially divisive subjects, such as curriculum topics about slavery or civil rights. These gaps made it harder to implement CREATE-aligned lessons.	It's finding resources that are age appropriate because something that I could handle and I can digest at 27, 28 is not something that they can digest at 12, 13 just because the language is not all there. But yeah, it's really, I don't know, it's a balancing act. (Participant G)
Individual skills	Some participants reflected on their own ability and comfort level in carrying out CREATE's mission and vision.	Barriers would be my ability, my comfort in doing it... I think that's the biggest thing and just being consistent with anything. But I think most of the barriers are me. I don't see that the school is ever put up any barriers for me or anyone else not to be able to do the things that need to be done. (Participant K)
Student buy-in	Participants noted that students were sometimes apprehensive at first when introduced to CREATE-aligned practices. Initial resistance required time, consistency, and relationship-building to overcome.	[Students are] like, "What is that?" I'm meditating. I'm literally closing my eyes and then when I'm trying to peek at them, they're looking dead at me, so it wasn't good. But this year, I mean, it was much better. (Participant B)
Tense co-teaching	Some participants experienced difficulty communicating and collaborating with their co-teachers. Challenges included disagreements about curriculum decisions and differing pedagogical approaches.	I realized even the co-teaching dynamic, that was really hard. That was extreme. That was the worst thing ever. That was the only thing I did not like about CREATE because we were very different in our teaching styles and he was very lax. Then our home lives, it was a lot of things I felt like they didn't consider, because I am a mom as well, and he was a regular student. His flexibility, his hours of time where we could plan, didn't align with mine. So instead of the co-teacher model, it was more of one teach, one assist, and then [CREATE] would push in and be like, "That's not what we want you to do." It's like, "That's all we really can do because he's not willing to do anything else." (Participant B)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
CREATE opportunities for support		
Consistent communication	Increase the frequency and reliability of communication and collaboration between CREATE, school leadership, and veteran teachers. Stronger communication channels help ensure shared expectations and better alignment.	I think communication is one. Obviously, I know everyone doesn't answer emails in one second, but I do think the communication and the access to the people is one. I honestly admit, for a little bit, didn't even know who was in charge anymore, to be honest. I had read an email when everything changed dramatically. I guess it was a year and a half ago and they went to Georgia State and the director left. I tried to piece through that and I talked to our instructional mentor, and then I feel like even she has made reference of it being different maybe like... I think when CREATE was more grassroots-ish, everybody was very, very close on the team, and the team was smaller, and now the team is huge. There's all these different positions added to it, multiple instructional mentors, all different levels of directors and coordinators. I think even for them, it's a very different setup, so yeah, I would say communication, access, just knowledge of what's happening. (Participant A)
Market CREATE to veteran & non-classroom educators	Increase efforts to promote CREATE to veteran teachers and non-classroom educators. Highlight how they can serve as mentors or take on supportive roles in implementing CREATE's mission.	It has helped me as a veteran teacher. I feel like CREATE is not just for recruiting new teachers and retaining those teachers, but it is helpful for the effectiveness of even veteran teachers as well. We could say something about that. Because you're open and honest and willing to do that reflection on your practices, I don't care how long you've been teaching, there's always something else you can learn in a different lens and something you can add or something you maybe take away. So I appreciate just how they continue to support the veteran teachers. It's needed. (Participant H)
Designate a CREATE lead within schools	Designate a CREATE contact person in each school (such as a Cooperating Teacher or experienced educator) to support communication, marketing, and networking within the school.	Maybe having a lead CREATE teacher at the school. So full transparency, the teacher that left was, I think she was our equity teacher. They called her because we have a team... It's like the counselor equity department... like a mini CREATE department at our school. And the teacher that left was really big on equity. She was really big on LGBTQ+ because that's just who she was... I think her daughters were identifying as LGB. So it was more personal for her. So she was more personally invested. So the buy-in was big. And since she's left, CREATE is still present, but it's not as boastful in a sense. (Participant L)
Characterize each school & inform residents	Understand and document each school's mission, values, and climate to help residents choose placements based on alignment with their own professional goals and beliefs.	They should do more research or provide an opportunity where student or their residents can dive into the community that they're about to go to... build a space where their residents can really understand the community that they're choosing to go to before they get there... Because I stayed with <redacted school name> for two years, despite the mental health challenges and other challenges based on this idea that I thought I was stuck... But I wasn't aware of what the community really gave and how the administrators and leaderships really responded. So with that, if there was an actual segment where the residents can understand the climate of the school first and the community, then they would make a better decision on how to move forward with their careers. That's something I believe that CREATE definitely did not understand... (Participant F)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Maintain individualized relationships with schools & leadership	Re-establish and maintain personal, relationship-oriented connections with early CREATE partner schools. These connections help trust, alignment, and sustainability over time.	When you go from being smaller to larger, when your success grows, when you get branded and you're starting to get popular and you get these different grants and you get a little more press and accolades, I think things can become cookie cutter when you try to make processes and structures for everything. I know it's obviously impossible to run any organization or school of flying by the seat of your pants... but I think when you do that, you do lose some of the personality or just the vibe of the work, which I do feel like CREATE is much less personalized, much more, I guess just like... I don't know... It's just like that personal warm fuzzy is gone. (Participant A)
District barrier		
Leadership lacks instructional expertise	Participants felt that some district-level personnel lacked the experience or patience to support new strategies long enough to see results. Promising approaches were sometimes abandoned too early due to unrealistic timelines or limited understanding of implementation needs.	I think a lot of times, the people who are directors or are on the board have never been a classroom teacher... they just kind of know what works for all kids or just students in general. But when we're in the classroom every day doing this all the time, some of us for many years, and we try to tell them like, "This isn't best for kids," or, "This isn't how kids learn," or, "This is what you should be doing." Just simple things. I think it's a lot of like, "No, that's not right because we don't have the time or we don't have the money or this," kind of thing... Even for example, with my board... We would try some, give a suggestion, the director would maybe try it, and maybe after a week, so not even enough time to see it play out. She would say, "Oh, this isn't working, so we're going to do this thing because this didn't work." (Participant I)
Compensation	Teachers described their pay in Atlanta as below a livable wage, creating barriers to recruitment and retention. While CREATE helped close the financial gap, participants believed that district-level compensation policies remained a challenge for sustaining long-term educator engagement.	I wanted to be a teacher when I was a kid... I feel my dream came true. And I've been loving it. So I want to do teaching as much as I can, as long as I can. But I really hope the state, they can pay the teacher a little bit more, really. I really hope that happens. (Participant C)
District plays unclear role	Participants expressed confusion about the relationship between CREATE and the district. Some were unsure if the district supported CREATE's mission, or were unsure because they didn't understand the relationship between their charter and the district.	Our charter school is set up so much differently... from my understanding, we have funding, we get certain resources through APS, but we also have a governing board where outside resources come, because our school initially was a traditional public school. But then it became a conversion charter school as well, so a lot of our decisions are placed on the head of school, which she works in the school, but she serves above the principal. I feel like she's more of the district person. I can't describe it... I don't even think our lunches come through APS anymore. Our students used to get bussed through APS. But yeah, other than that, I don't really know our relationships with APS. (Participant B)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Limited promotion	Some participants described CREATE as feeling like a "secret society," with limited presence. They suggested a need to actively promote CREATE, build awareness, and foster schoolwide engagement.	Sometimes it can seem like a secret society in a sense, because the word of mouth isn't there. Because my first year there, I was like, CREATE teacher that? Is that a student? I had no idea until I had more of a relationship with that teacher who left, who gave me more of a background on CREATE, and then I attended those meetings. Even my good friend was like, "How do I become a CREATE teacher?" And I had to forward him some emails and he's like, "Where was this information at?" (Participant L)
Leadership pushback	Charter directors or board members sometimes shut down ideas aligned with CREATE's mission.	So the directors of the charter, even maybe some of the board. I think that's where a lot of the pushback was coming from and the support, the lack of support I should say was coming from and just having conversations sometimes with my instructional coach or my AP, it made it seem like, "Yeah, we hear you, but because somebody higher than me is saying we can't do this or that," like it was out of their hands kind of thing. (Participant I)
District support		
Aligned language	Participants observed that language around racial equity used by the district often mirrored messaging from CREATE. This alignment in communication helped reinforce shared values.	[The district has] to approve of this connection and approve of the residents coming into the school building... so they're aware. And I think even more than just being aware, I'm assuming by proxy they support it because they wouldn't allow it if they didn't support it, I would think... but even more so than just that surface level, the superintendent just in communication that's sent to us electronically, emails, whatnot. Some of the wording is the same. So the racial justice, equity are used in her communication to us. So to me, even if she doesn't say CREATE, I feel like they support what CREATE supports. (Participant H)
Students as Change Agents	Students across the district actively pushed for change, including petitioning to rename a high school—a request that was ultimately approved by the district. Their actions demonstrated that student leadership supported CREATE's vision of equity and agency in the district.	I know students at the high school that we feed to, they changed the name of <redacted> because of him being a racist person and a leader in the Confederacy. I love the idea that students that I used to teach took on that, carried the torch, and pushed for change of the names of schools in Atlanta Public Schools of people who they felt were not right. (Participant D)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
GSU or state policies		
Evaluation requirements	Participants noted that Georgia's updated teacher evaluation system focuses on fewer standards for new teachers, which felt more realistic and aligned with CREATE's goals for new teachers. Some emphasized the importance of coordination between Georgia State University, Atlanta Public Schools, and CREATE around evaluation tools like lesson plans.	Starting next year, Georgia is actually only focusing on two standards with the first and second year teachers, which I think is great, which is planning and professional knowledge. To me, that really aligns with CREATE's work because really what they're focusing on is getting you up and running, right? Yes, in the past, Georgia has had to evaluate you on 10 standards, but in reality, on your first semester of teaching, you're not going to have an academically challenging environment. You're not going to be able to differentiate with ease. You're not probably going to be able to have this amazing well-run machine classroom where the kids are productive and on task and collaborative and independent and self-mentoring. I do think knowing Georgia's dialing back and recognizing that these newer teachers can only focus on a couple of things at a time and it builds up on this little pyramid... (Participant A)
HB1084	Participants noticed the removal of the CREATE mission from the website due to HB1084, reflecting concerns about political or parent backlash. The law created fear around teaching certain aspects of history, particularly topics related to race and equity.	I mean, some of Georgia's policies really makes it difficult to have some of the harder conversations, especially if kids are asking questions. (Participant G)
Resident-Cooperating Teacher matching & certification	Some participants raised concerns about the rigid requirement for student teacher certifications to exactly match that of their Cooperating Teacher. This limited flexibility, and participants saw it as a disservice to teacher preparation. Placing residents in niche classes, like accelerated courses, also posed challenges, as they weren't always prepared to teach every student group.	So for me to be like, oh, well, we can't give you a student teacher because your certification isn't in the same field as they're in, I thought that was a wall to run into, because if your certification is in education, I'm a teacher, if your certification is in mathematics, and though my original certification may not be in mathematics, it's just I think that did a disservice to whoever that teacher would've been to get that experience... Like the equity piece, you don't really know how to teach Black and Brown kids culturally, unless you're in a school with Black [students], you know what I mean? (Participant L)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Embodying CREATE's mission and vision		
Advocate, trailblaze, serve	Participants said they advocate for all teachers and students to feel supported, included, seen, and celebrated, no matter their background. Participants said if there aren't available resources or tools to support them, they work to find them. If there is an authority figure enforcing barriers, these educators say they are willing to address them. Sometimes, participants said these efforts led to school transformations, positioning them as leaders and trailblazers.	I think CREATE's vision, well, how I would describe it is just cultivating an anti-racist educator and then as well as giving educators the tools to basically make changes in the school and bring to light some of the inequities that you may see in the school and being able to speak up and provide resources to make it a better school environment for the kids and for your coworkers as well. (Participant G)
Dedicated, committed & persevering long-term	Participants said they are willing and committed to doing the work, whatever that takes, as demonstrated by their commitment to extra training and extra work through CREATE. They say they will stay in teaching despite challenges and setbacks because they are dedicated to their students' achievement and livelihoods, and have the tools needed to prevent burnout and remain resilient, ultimately increasing retention in the field.	I think I would think they are more dedicated. I would think they went through a lot of extra training and extra work. They're really serious about teaching. (Participant C)
Disruptors	Participants said that, even in the face of socio-political opposition, CREATE teachers express dedication to instruction that is culturally relevant, responsive, and sustaining. They report speaking up, confronting inequities, and teaching the truth while remaining committed to joy in the classroom and school -- making it a more positive and affirming environment for all.	...alluding to this disruption of normalcy in education right now -- how we think about curricula -- I'm living in Georgia and we all know of the different curriculum at laws that are being put out there as it relates to critical race theory and other divisive concepts in some eyes, what some people might think. Trailblazers in a sense of, well, how do we make the standards of teaching work for us but still be able to teach the truth and liberate students? That's that connection of thinking around society as it is set up now that I think about with trailblazing that is. (Participant F)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Critical frameworks & mindset	Participants said they view the world through a lens which critiques oppressive systems, and understands the overlap between history, society, and schooling, promoting equity whenever possible. Participants acknowledged personal biases, and aimed to elevate the genius of Black culture and influence. Participants report promoting culturally relevant, responsive, and sustaining pedagogy that acknowledges the true history of the United States, including Black patriotism, and aiming to serve students rather than save them.	From my understanding... they really wanted to cultivate educators who were culturally responsive and really conscious about the biases that we have. As humans, everybody has biases, and we kind of have that in our perspective when we go into teaching and all the different students that we're going to meet and all the different places and things that they're bringing in and coming from. So really framing our mindset around that and being mindful of what that might look like or feel like for not only ourselves but the students that we're teaching or we come in contact with. So I think their vision was just making you consciously aware of what's going on in the world, what's going on with your students or other people that you just encounter, and then how you can be conscious and thoughtful and mindful about those situations. (Participant I)
Identity work and mitigating biases	CREATE educators undergo extensive identity work and report understanding that all people have biases. Participants said they work to unpack their own biases and assumptions to avoid bringing them to the classroom, and refrain from pushing their own beliefs on students, opting to act as classroom facilitators rather than knowledge transferers. Participants said they do their best to remain neutral and unbiased in efforts to support and teach all students, and engage in extensive planning and research, relying on credible sources when discussing potentially divisive topics. Participants report approaching colleagues, students, and students' families without judgment or a deficit-lens, and remain mindful about how conversations are worded.	What would that look like in the classroom? Really dedicated to planning. They really push backwards design and identity. So, they focused a lot on knowing who I was as a teacher or knowing myself before I even pushed into the classroom. We spent a lot of summers around identity practices. So I feel like in the classroom, it looks like teachers being very secure within themselves and not really pushing on their beliefs, but creating an atmosphere where kids are able to... What am I trying to say? Search for who they are within their education, see themselves within their education. I feel like that's what it looks like. (Participant B)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Community building & collaboration	Participants said they were committed to fostering a welcoming community with ample professional support and collaboration, making an explicit effort to connect on a personal level with other residents, colleagues, and students—and to connect students with each other. Participants said they create opportunities to build personal connections with people of similar and different cultural backgrounds.	The connections. Oh my God. So when I first started this school, we would start off with connections, and I was like, what is this? I would be on my phone. I was irritated. I was like, I can be doing a million other things right now than this. And it wasn't until we had that CREATE training and just to live in the silence, no matter if it was just 10 minutes, just be present in the moment, whatever you feeling in yourself, it's very therapeutic. So I then start to transition that into the classroom. (Participant L)
Connecting with parents & families	Participants said they were committed to becoming better teachers and supporters of students by connecting with and getting to know and value their family, while working in partnership.	My children years ago also attended the school. So even though I didn't live in the community because my kids were there and participated in activities in the community, it helped me. And helped us to just have a different level of commitment to the community, which got us outside of the four walls of the school, so to speak and into working with the families. Because it's not just the kids, it's also the families. And that helps to teach children when you learn more about their backgrounds and their families and where they come from and what they believe. So that goes right hand in hand with teaching children and valuing their backgrounds and just participating with everyone in the classroom. (Participant H)
Focused on wellness & resiliency	CREATE focuses on self-care and wellness, helping participants balance all their responsibilities and set boundaries. Participants promote meditation, mindfulness, pauses, and reflection to understand how emotions influence and impact the classroom. Participants reported working with students to promote their wellness and social-emotional awareness, and to create a classroom environment focused on wellness through strategies like a calm-down corner, yoga, essential oil diffusers, and sensory tools.	I was mentioning wellness... teacher's emotions and life experiences can directly impact the classroom atmosphere and just students' and feelings. So I learned how to pause. For example, I know my limitation, I learned to know my limitation and knowing that if I'm struggling in my personal life, I should spoke up and tell people that "Hey, I need a break," or "Hey, I need some help." And also, I feel like because I'm paying attention to my own emotions that I will pay attention to my students' emotion too. In my classroom, I also share them the same thing that I learned from CREATE. Sometimes kids came to me exhausted. I told them, "How about let's go out for a walk?" So I actually take my students out for a walk during class a lot whenever they need it. (Participant C)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Geopolitical landscape and systems	Participants reported increased awareness of local and national history, politics, and current events, understanding the connections between these contexts. For example, understanding the origins of the school-to-prison pipeline, and how to disrupt it.	...some of the things that we talked about was the history of education and specifically in the US and how it wasn't really built for Black and brown students. So a lot of the things that affects Black and brown students in the schools, it's because of the system. And you're working in an educator in that system and you're trying to basically do all the you can to combat this system. (Participant G)
Highly trained, skilled educators with holistic teaching approach	CREATE teachers have undergone extensive training beyond their traditional credentialing program. Participants describe extra training on backward design and intentional lesson planning, and being well-versed in promising pedagogical practices, such as differentiation. CREATE participants receive extra training in their core academic subjects, but also mindfulness, social-emotional learning, and culturally and historically responsive pedagogy. Participants reported feeling experienced in navigating discussions about equity.	I really do see them creating these critical thinking, critically conscious educators who also are in tune with the holistic view of teaching. So that includes being, I would say mindfulness teaching as well and teacher training. (Participant F)
Classroom management skills	Participants described being well-versed in (a) classroom facilitation, (b) building rituals, routines, relationships, and a positive classroom atmosphere, (c) setting group and individual goals, and (d) implementing non-punitive classroom management and conflict resolution techniques. They aim to identify the root cause of behavioral issues and refrain from separating home life from school, seeing their students as whole people, and strive to understand students' individual backgrounds and tailor social-emotional and academic supports to help them succeed.	I think about the social emotional part of classroom management as well. So I've used a lot of my own techniques of how I control my emotions and be aware socially. I include yoga into my class, I include meditation. I also make sure that I'm intentional with setting up the actual classroom. I have a diffuser, with different essential oils and things of that nature. Because I went through CREATE, I had a chance to look at the inside of a school, I had a chance to look at images and hear from other teachers of how they set up their rooms and what's the science behind it, because of that experience, I'm able to be more intentional about how to set up my classroom, where it begins to manage itself. (Participant F)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Student voice & agency	Participants describe uplifting student agency, voice, and reflection in the classroom and encouraging students to draw their own conclusions based on credible sources and critical thinking. Some participants described making a concerted effort to ask questions so that students draw their own conclusions, rather than tell them what to believe.	When we create rules and expectations for our class, we get [students'] input. So it's never really like, "Oh, I'm the teacher, so these are the rules of the class." It's like, "What do you want this class to feel like and look like for you to be successful?" And then every morning we always go back to that. They say it every morning, they commit to which part of our agreement they're going to do. (Participant I)
Lifelong learners	Participants describe taking a stance as a lifelong learner and acknowledged that no matter your experience level, there is always room to learn and grow. Participants frequently relayed that there is always a new skill or technique to explore, and change is always happening.	I feel like CREATE also has a mission to guide teachers into just continuously learning, just continuously learning and just understanding that just change is always happening. I hope that makes sense. (Participant E)
Collaborative mentorship	Participants describe valuing their mentors and seeking them out for support, particularly around dilemmas and problems of practice. Mentors were described as respected peers, and residents are trained to be critical friends. Mentor teachers in CREATE have undergone identity work to understand how they work/show up with people, and are certified to train and guide pre-service and early-career teachers as well.	It feels like we have that support system from [our fellow cohort members], because even though we're at different schools, we experienced some of the same issues and some of the same problems that we've learned, but we've become really tight-knit... I used to call some of my cohort members and tell them like, "This happened today," or something like that. And sometimes you just need another person's perspective. And even though they're also new teachers, they might have also experienced it and been able to get through the situation. Yeah, we've just been a big support for each other. So I would say that they're one of my biggest supporters as well. (Participant G)
Openness and vulnerability with others	Participants describe being open to authentic and vulnerable human connection with mentors, peers, and students. They claim to be open-minded and receptive to new ideas and feedback, as well as people with different backgrounds and experiences. They say they encourage their students to be open-minded and open to authentic connections as well.	I would assume that they're someone who wants to continually be educated about, again, dismantling some of the systemic issues and education, someone who is open-minded and willing to learn, someone who provides support and then who is open to being supported in their journey as an educator as well. (Participant G)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Relationship oriented with students	Participants describe developing individual connections and relationships with students to teach and serve them better, and to motivate them to engage in learning. They said they give students grace when life outside of school interferes with their learning. They also describe encouraging students to develop relationships with each other to create an atmosphere of mutual respect, accountability, and support, and encourage restoration when a relationship breaks down.	I'm going to say that it's all about creating relationships. I think, and no pun intended, I think CREATE actually helps us learn how to create relationships regardless of background of any student and also praise, praise and having rituals and routines in the classroom. What I do is I play onto that. I build relationships with all my students in a sense where I can even sometimes for myself having a code language with each individual student, just knowing their background and what they like and their triggers. Triggers was a word that got to my mind because I think about the social emotional part of classroom management as well. (Participant F)
Praise, compassion, love	Participants describe leading with love and praise those around them, believing all children can learn and that they can uplift all children no matter their background. These participants said they value the differences of their peers and students, and they strive to approach all people with grace, compassion, and respect.	I'm giving my commitment and students are giving theirs in times maybe being motivated to do so. And it's going to always, always, always look as if I love the students. No matter how bad or how emotional they may be, but it's always going to look like [I love] the students. (Participant F)
Safe space, shared power	Participants describe valuing a safe space for educators and students, where everyone's experiences and opinions will be heard and respected.	CREATE's big on giving the safe space for students and teachers. So when I say a critical conversation or safe space, the students have their listening ears on. Everyone's opinion, if given, it's going to be heard and respected. (Participant L)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
School barriers		
Colleagues resistant to change	Some participants reported that their colleagues were not open to new ideas and were "stuck in their ways." Several colleagues felt that new pedagogical approaches were a waste of time, particularly when students were struggling with the basics, like reading. This resistance to changing mindsets sometimes created a "divisive" atmosphere among teachers.	I feel like they're just stuck in their ways for teaching. They complain about a lot of stuff. They don't like change. They really don't like change, to the point they don't want to change... Whatever works for them, works for them. I feel like now that they feel like they're in their groove of teaching and their culture to retirement, they're like, "If it's not broke, don't fix it." They really have that mentality. (Participant B)
Competing demands, stress, and limited capacity	Participants described some schools and educators being overwhelmed by competing demands such as staff shortages, heavy workloads, and standardized testing priorities, leaving little time or energy to engage with CREATE's mission. Sometimes teachers felt they were in "survival mode," managing classroom crises and basic needs rather than planning transformative approaches. Stress and lack of capacity prevented them from implementing CREATE's mission and vision, even when they supported it.	I just found out two days ago that I do have a student teacher, I didn't have one and that was because of one of my colleagues. She decided that she wasn't going to do it this year. It wasn't because she doesn't like CREATE. She's like me, she's a 30-year teacher and her time, she felt as if she didn't have the time that it took to spend with the student teacher. (Participant D)
Misalignment with CREATE mentors	Some CREATE mentors fit well with their school culture because they were open and flexible, while others were too rigid or set too many boundaries. Participants noted limited collaboration between CREATE mentors and school leaders, which sometimes created tension. Mentors grew frustrated when teachers were hesitant to deviate from standardized curriculum and pacing guides.	It was challenging because for that year, the typical instructional mentor that we've had all the years, they assigned her to some other school, so we had a new one... She did draw these very clear boundaries of what the school's responsibility is when it came to the residents and what her responsibility was, which I felt like was interesting because I felt like it was definitely more of a grassroots holistic effort when before CREATE was... taken over by the university.... the efforts were much more holistic and relationship-based prior to the structural change... When we were making the decision to move forward with [a resident] termination... I'd get on the Zoom with our mentor that year and one of the coordinators, and I would almost feel like they were questioning us, giving us lessons about how to be admin or whatever, which is very in stark contrast to how our relationship typically is. (Participant A)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Conflicting mindsets and approaches	Participants noted philosophical differences among staff regarding culture, discipline, and pedagogy. Some colleagues resisted culturally responsive practices, viewing new approaches as unnecessary or incompatible with standardized expectations. Disagreements over behavior management—including discipline challenges and differing views on cultural or socio-historical influences—made it difficult for schools to establish a unified approach to addressing student needs.	I think the barriers would be teachers who haven't gone through CREATE or even had some kind of CREATE training... They're not willing to change their mindset or think of things differently... because even in our school, our whole thing was like, "Oh, let's do restorative circles, whatever the case may be," which is great. But I also think if you don't train teachers properly or they don't really have the kinds of trainings or protocols that most do when they go through CREATE, then that can be a barrier because you're trying to make this change. But if other teachers aren't our board... then that can definitely be a barrier and that affects the school climate, which I've seen. (Participant I)
Need more CREATE members & influence	Several participants felt their colleagues might be more open to implementing CREATE's mission and vision if they had greater exposure to CREATE staff and professional learning. They suggested that having more CREATE educators employed at their school to model strategies could help demonstrate the positive impact of CREATE's approach.	"I don't know if it's the numbers. I mean, we are a fairly small school compared to some others. And I don't, to be honest, I don't think CREATE is ... What do I want to say? A big presence. It's like, okay, well Ms. Miles or Ms. Whomever else, they're two CREATE teachers. Oh, okay. It's almost like you guys are over there doing your thing with the CREATE people. Oh, okay. It's almost not inclusive of the whole school. Because it's, like I said, half the teachers not don't know, but they just don't have that deeper knowledge or the close-knit, the working thinking of how this can impact us all." (Participant H)
Culture and climate challenges	Several schools lacked stable leadership, cohesive policies, or adequate support and resources for teachers, which weakened school culture and climate. High turnover and reliance on substitutes made it difficult to build stability, while some schools had strong pre-existing cultures that overshadowed CREATE, leading teachers to abandon CREATE practices in favor of established norms. These cultural dynamics limited CREATE's ability to take root.	I am very sorry to say that this school year was a nightmare as far as, well the school building that I'm at... this was almost like our fourth year without a principal.... This year has been very topsy-turvy in the sense that we don't have a culture, a school culture... Also, a very diverse neighborhood, so without leadership in a school that's this diverse, not good. They hired two assistant principals to get us through the school year and then they hired a principal. Your question about the school: The school had a very hard time supporting CREATE because we were all over the place. We were not grounded. We were not rooted with tradition. We were not rooted with protocol. (Participant D)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Identity conflict	One teacher of color felt they were up against stereotypes, with colleagues and students expecting them to serve as a cultural expert and prove their expertise beyond cultural education. Another participant described language barriers that created interpersonal misunderstandings, and one recalled difficulty discussing Israel/Palestine out of concern for offending Jewish students.	When I was in <redacted school name>, it was really hard. I felt the need to prove that not all stereotypes are true. It was very lonely. And I think I was living in a situation where I felt very stressed. I went to the school and did my best for the day and came home and cried. (Participant C)
Parent and community pressures	Participants described how parent dynamics and broader socioeconomic contexts shaped school decision-making and affected CREATE's implementation. Some educators faced pushback from white, affluent parents on divisive topics or mistrust toward novice teachers, while others struggled to meet the needs of students from low-income backgrounds dealing with behavioral and mental health challenges. Several participants also criticized colleagues for failing to consider family financial limitations when planning programming and events, such as field trips, which they felt excluded some students and parents.	...if a parent calls [school leadership] and they're like, "Well, this teacher said this or brought this up, or allowed there to be a discussion about X, Y, and Z." And it might not even be that controversial, but some of the students at the school I was at, it was very, they cater to the parents a lot. And the parents could also be a barrier as well, just because the area that I was in. So yes, it is part of APS schools, but it was a more affluent area. So therefore a lot of, basically any issues that you would hear from a parent is not typically a parent of color, let's just say. (Participant G)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Staffing & responsiveness	Participants emphasized the need for leader buy-in to create momentum for implementing CREATE, but some schools lacked responsive or stable leadership. High teacher turnover and reliance on long-term substitutes left remaining teachers overworked and with little time or energy to innovate. Veteran teachers often dismissed newer teachers they expected to leave soon, and finding substitutes so educators could attend CREATE professional learning was also a challenge.	So at my school, we have a pretty high teacher turnover rate. So the first example I can think of... one of the CREATE teachers, she was in fourth grade. So the grade that was having some issues, basically all of the teachers on her team quit. So it was just her, and this is her second year teaching. She is still kind of new, and instead, the admin just put in a bunch of subs, but didn't really give any real direction to those subs, like, "Hey, this is what you do. This is how school works. These are the things." They kind of just left it to that teacher to do it. But at the same time, she had a class that was difficult. So because she had a class that was difficult and all the other fourth-grade classes, of course, since they lost their teacher, they're going to be chaotic, and then fifth grade is chaotic. So it's like she's trying to get support, but they can't really give it to her because they're trying to deal with fifth grade and all these other fourth-grade classes. So that's an example. A similar example happened in kindergarten. We lost three teachers, and they kind of just hired these kindergarten teachers and were just like, "Here's your classroom and you're just kind of on your own. Rely on the people on your team to guide you." But it's like, "We also have a lot of stuff we have to deal with. Yes, we want to be helpful, but at the same time, that's not our job. That's your job, you're the principal, so." Those are incidents that I think about when I say support is kind of there, but also not really. (Participant I)
Effective communication strategies	One experienced educator noted a lack of effective avenues for communication among CREATE participants in their school. Another teacher observed that some educators struggle to communicate with students or avoid sensitive topics, which limits the potential for meaningful reform.	I'm not sure exactly what the residents do... I don't see any communication between -- except for regular school-based stuff -- I don't see anything specific for CREATE between me and one of the residents, or someone else from one of the residents. (Participant K)
Lack of accountability structures	Some participants wanted school leadership to push educators more actively to engage with CREATE. Others felt there were no accountability systems or dedicated time for reflection to ensure CREATE's mission and vision were meaningfully implemented, even when teachers were generally supportive.	I think because there's not necessarily some accountability factor, and I get that too, that's hard because no one wants to be watched. No one wants to be check listed. We're already observed. But I also think this is super important. You know what I mean? This is beyond what we're being taught and what we're supposed to be doing is very important. I think because we're not being held accountable all the time, then the barrier is created, if that makes sense. (Participant E)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Leadership barriers	Participants described challenges both within school leadership and in their relationships with CREATE. Some felt school leaders were not open to transformative ideas, stifled innovation due to parent expectations, or failed to provide support and follow-through. Others struggled with communication and alignment between CREATE and school leadership, perceiving CREATE as positioning themselves as authority figures rather than collaborators. As CREATE scaled, some leaders felt the program became less personal and more disconnected from K–12 realities. Separately, some schools lacked stable leadership altogether, leaving staff without cohesive culture or shared policies.	I feel like if the leadership within a school is not really willing to accept new ideas, I feel like it really depends on the leaders in the school because CREATE does have a more... Well, I'm not going to say more of any type of approach. They just have a different approach to traditional teaching. (Participant B)
Mandated curriculum, pacing guides, and testing	Some participants felt unable to embrace the "risk-taking" that CREATE promotes due to highly standardized curricula and strict pacing guides. Non-CREATE instructional coaches sometimes reinforced this by urging teachers to maintain established content and procedures, and some educators preferred familiar methods they felt "worked." High-stakes testing also took priority, leaving little space for innovation or culturally and historically responsive teaching.	CREATE has these huge ideas of risk-taking in the classroom. I often tell them how difficult it is, especially when a curriculum has been introduced. I can't really take too many risks because literally, I have to stick to the script... In my school, you can't really get off course with, because if you do anything outside of that curriculum, it messes up the structure of it. (Participant B)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
School supports		
Need for CREATE (gentrification, affluent parents)	At least two schools faced neighborhood gentrification when they partnered with CREATE, which supported the challenges CREATE-trained educators were addressing and reinforced their work to improve equity in their schools.	<Redacted school name> started to have this really interesting dynamic where we were designed for the students living in these apartments over here... but then people are building $900,000,000 houses right behind us. When the cultures integrated, it became like there are some racial relations issues here. There are many equity issues here. There are access issues. There are issues with how the parent population is very rooted in the parents who have privilege, and have work from home jobs, and have time to come to meetings versus there's been multiple instances where the more privileged parents, black and white, have referred to the students who are economically disadvantaged living in the apartments as those kids or the other kids. When the work started with CREATE and iGroup, it all started at the same time, I think it was a perfect time just because we were experiencing a huge shift just in our school culture, our school community, our population that all that work has just really just helped to lend itself into the work that we do normally, too. (Participant A)
School markets CREATE	Leadership, CREATE participants, and alumni encourage teachers to attend CREATE professional learning. One school welcomed CREATE leadership to staff meetings to present their programs and offerings, though engagement decreased during and after the pandemic.	Sometimes the schools tend to have CREATE come out and just kind of explain to them again with the new staff what they're all about. (Participant H)
Staffing and personnel alignment	Schools with CREATE alumni benefited from their leadership in spreading CREATE-aligned practices and mentoring others. Alumni often recruited new educators into CREATE and brought back ideas that expanded the school's capacity for equity work. Some schools actively screened new hires for alignment with equity and wellness values, ensuring philosophical fit. Evaluation tools reflected CREATE principles, reinforcing consistent expectations. Goal-setting practices helped shift resistant teachers toward CREATE-aligned approaches. These staffing efforts helped sustain a shared commitment to CREATE's mission across the school.	In my school, I think I have a couple of them from CREATE and I feel like they tend to be more welcoming. I do have a couple of colleagues that were actually a graduate from CREATE too, and they will come to you and I just feel so nice and that bond already there. (Participant C)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Resource allocation and PL opportunities	Schools invested in resources and learning that supported CREATE-aligned practices. Funding covered professional learning fees, training time, classroom coverage, and materials like culturally responsive books or mindfulness tools. Some schools offered wellness funds, community gardens, or staff appreciation initiatives. Educators participated in professional learning communities, peer learning days, and affinity groups to deepen their engagement with CREATE's mission. Schools supported teachers in earning microcredentials or coaching certifications. These investments reinforced teacher development and helped sustain CREATE-inspired work across the school.	When I talk about the Lunar New Year celebration, we were able to get the funding from the parents. So the parents' association funded $1,000 for us. So that's what we use that money to get the lion dance and we're able to order some Chinese food. We actually have a food taste that a restaurant make like spring roll, dumpling, all the stuff, and the kids were able to get a taste. And also, the school cafeteria, our school made food from scratch by the chef team. And they will change their menu based on holidays or festival, things like that. So that day, I think we had noodle, that we all worked together. We would email the chef and tell him like, "Oh we're planning to do this." And he was super excited and he prepared a noodle for that day. (Participant C)
Classroom implementation support	Co-teaching and coaching structures provided teachers with the support needed to implement CREATE practices. Shared planning and instructional responsibilities allowed educators to experiment with new strategies, such as social emotional learning, mindfulness, and culturally responsive teaching. Instructional coaches played a key role in helping teachers design lessons and adapt materials to align with CREATE's mission. Some schools had systems—like vetting committees for potentially divisive topics—to ensure content was well-supported and aligned with policy. Teachers appreciated having trusted partners to reflect on their practice and troubleshoot challenges. These supports made implementation more sustainable and less isolating.	My instructional coach. I know my first year, my principal as well as my instructional coach, we'll meet once a month and just talk about the needs I might have in the classroom. Then they will ask what is CREATE doing? Whatever they're not doing, how can I help? (Participant B)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Leadership alignment and structural support for CREATE	School leadership actively collaborates with CREATE staff, aligns hiring practices with CREATE values, and adopts CREATE language and practices. Leaders acknowledge that CREATE fills gaps in school support—such as wellness, coaching, and equity-focused development—and they nominate teachers to participate. They provide planning time, invite CREATE-led sessions, and encourage teachers to attend professional learning. Some schools support coverage or data collection so CREATE educators can fully engage. In schools already aligned with equity ideals, CREATE was easier to implement. In others, a supportive culture allowed teachers to adapt curriculum and lead initiatives in line with CREATE's mission.	I think with our principal and vice principal, just a genuine inclusivity of people and lifestyles and also being open to hearing complaints... we had this field day activity. One of our activities was this hair station where kids could spray their hair different colors... I was like, "I think that sounds like a lot of fun." But I said, "But we really need to make sure that for our African-American students, that this is inclusive... Because I know me as a child, I could not get my hair sprayed." There were just certain cultural things around that... I could see people were just like, "Why is this a problem?" But then just having the space to be able to explain that... being able to express that, because I've been in other environments where I don't think I would've necessarily felt comfortable expressing that kind of concern. But that I think went over really well. I think it was helpful and it worked out. I think that's just one way. I think our principles are just supportive of ... People just speaking out, but in a way that's respectful and in a way that will actually help the situation. (Participant E)
Ideological and instructional alignment	School staff often share values aligned with CREATE, such as equity, wellness, culturally responsive teaching, and family engagement. This shared belief system creates a supportive environment for implementing CREATE's mission, such as eliminating punitive punishment. Teachers reported open communication with families and colleagues around these values, making practices easier to adopt. Academic freedom allowed teachers to choose or adapt curricula to meet student needs. Teachers felt trusted to explore sensitive or justice-oriented topics, sometimes with counselor support available. These school philosophies reinforced CREATE strategies.	My CO will tell us that, "Yes, we are your backup. You do it." So we were able to talk about a lot real, true story that happened in the history. And the kids were wondering and we all did research all together and we'll share about how that impacts our life today. We were able to talk about all those like race, equity issues in this country too. Yeah. I think we were given lots of freedom to teach. (Participant C)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Staff capacity and openness to CREATE	Some schools had highly committed staff or designated DEI personnel who played key roles in advancing CREATE's mission. These educators were described as available, engaged, and willing to take on leadership or extra responsibilities. In schools with younger teacher demographics, participants noted greater openness to new ideas and CREATE-aligned practices. In contrast, some veteran teachers were more hesitant or resistant to change, which slowed implementation efforts. The presence of motivated, equity-minded staff created momentum and reinforced CREATE's values. Overall, staff composition significantly influenced how fully CREATE could take root in a school.	Let's just say we're watching CNN 10, and they were talking about Black history... then students want to know more. They want to acquire more. And if they want to have to continue those conversations, I can either let my humanities teacher know like, "Oh, this was a topic that we had, they wanted to extend it." Or just let the counselor know, like, "Oh, this is a topic we were discussing. This seems like a 'critical conversation'." That's our keyword. This seems like a critical conversation. So then the counselor would then set some time aside to speak to students as a whole to address any racism and equity issues. (Participant L)
Culturally responsive and relational school culture	Schools fostered a sense of inclusion, connection, and openness that aligned with CREATE's values. The school encouraged teachers and students to share their identities, which strengthened relationships and supported equity goals. For example, families engaged in informal events, classroom visits, and cultural celebrations, often supported by PTA funding. Some schools held space for critical conversations and provided counselors to support students after heavy topics. Educators felt supported by colleagues who shared feedback, encouraged experimentation, and helped build a collaborative culture. These practices created a climate where CREATE-aligned work could thrive.	And there's just a different type of talk that we have with the students when their parents are there. And it's not just sitting in the classroom with the students. But also, we have activities like the Parkside Stride. There's a walk or run where we raise money for the school and for certain causes. But we are out into the community in the streets walking and talking. And you get to see the parents and the families again, just walking in the streets. Yeah, there's most recently a Belt Line, the Belt Line project in Atlanta. The Atlanta Belt Line in Old Fourth Ward where there were murals that the students painted. Fourth and fifth grade students painted. So families, we all got to go out there if you chose to and help participate and paint and work with that with the students in that area. Just something that builds community outside of the school. (Participant H)

TABLE M1. RESULTS FOR COHORT 5 YEAR 1 MATCHED CASES OF TEACHERS

Theme	Description	Example quote
Time for buy-in	Some schools were early CREATE adopters, allowing time for integration. Others struggled due to CREATE being relatively new or due to high teacher turnover. In one case, a school was previously aligned, but newer teachers lacked CREATE training or awareness.	I feel like our school, compared to just some people that I have interacted with, is a little farther ahead. But I wonder if it's partly because CREATE started with <redacted school name>. You know what I mean? We've just been in it since the beginning. I do think we're a little further along in that sense, because the program has really been a big part of our school. When we do have our [CREATE] monthly calls and we're talking about just some of the situations that other mentees are going through, it does make me think, "I wonder if they're having a hard time too, because the school culture is not ... it is just a little bit different." I think in a school you can have lots of professional developments and you can have lots of people coming in and saying and doing different things. But if there isn't buy-in from administration or if there isn't buy-in, then it's just another meeting. But yes, I have had conversations with other people that are just kind of like, "It's not necessarily a big part. It's not necessarily a part of the culture." (Participant E)

Note. We conducted 60-minute Zoom interviews in May and June 2023 with 12 participants across the larger CREATE evaluation: 4 residents, 4 Cooperating Teachers, and 4 experienced educators. PL = professional learning.

Appendix N. Economic Evaluation

Given empirical evidence of CREATE's positive impact on teacher well-being (see Chapter 4) and evidence of positive impact on teacher retention in a previous evaluation of CREATE under an Investing in Innovation (i3) grant (Jaciw et al., 2021), we understand that interest in implementing the CREATE program model may expand beyond GSU CEHD and APS. To help inform potential partners about the CREATE program model, we planned to conduct an economic evaluation. However, we could not obtain the necessary program data for a full economic evaluation. Thus, this appendix outlines our work toward an economic evaluation of CREATE by describing:

- the purpose of our economic evaluation,
- training received to conduct the evaluation,
- a definition of economic evaluation and its benefits, and
- an outline of the methods, assumptions, and sensitivity analyses we were prepared to conduct.

We conclude the appendix by discussing the complexities and challenges of conducting economic evaluations in education. Much of the existing literature about challenges in conducting economic evaluations is in the medical field and is not often applied to educational contexts (Tiesteel et al., 2024). As such, conducting economic evaluations for educational interventions is still a growing field (AIR, 2021), and there is not much guidance for education researchers looking to conduct an economic evaluation (Levin et al., 2018). We, therefore, hope our insights will be a useful contribution to the growing field of economic evaluations in education.

Purpose of CREATE's Economic Evaluation

To inform potential partners about the cost of the CREATE program model, CREATE leadership needs data that outlines the resources required to implement CREATE with fidelity and the cost of those resources. The primary purpose of the economic evaluation was to calculate the total incremental cost of CREATE compared to the business-as-usual (BAU) teacher preparation program (TPP) at GSU CEHD through a cost analysis. The following research questions guided the analysis.

✨ Intended Research Questions ✨

What is the total incremental cost of the CREATE program model for the final year of implementation during the 2020 SEED grant (the 2022–23 school year)?

What is the average cost per CREATE resident?

What is the marginal cost of serving one additional resident?

What is the average cost per student served?

We intended for the economic evaluation to include a cost-effectiveness (CE) analysis. A CE analysis allows stakeholders to compare similar programs based on cost and impact (further described under "Defining Economic Evaluation" below).

Staff Training

To prepare to conduct the economic evaluation, our Senior Research Manager attended two separate in-person trainings offered through the IES Methods Training in Economic Evaluation. These trainings, facilitated by experts from the Center for Benefit-Cost Studies in Education (CBCSE), occurred in July 2022 and May 2024. The July 2022 training focused on the fundamentals of conducting a cost analysis and implementing the ingredients method, while the May 2024 training included an overview of cost analysis but primarily focused on conducting cost-effectiveness and cost-benefit studies. Participants were provided a copy of *Economic Evaluation in Education: Cost-Effectiveness and Benefit-Cost Analysis* by Levin et al. (2018). Both training sessions also included an opportunity to work alongside experts to plan economic evaluations or troubleshoot current projects. Together, the trainings covered the content required to conduct the current economic evaluation study.

Background

Defining Economic Evaluation

The term "economic evaluation" refers to the group of evaluation methods "that explicitly address the cost of implementing a policy, reform, or intervention or that explicitly compare costs with benefits" (Levin et al., 2018, p.4). Lin et al. (2018) add that, to be considered an economic evaluation, the study "must examine *both* costs and consequences of at least two alternatives" (p. 152), such as a BAU comparator.

In economics, "cost" has a unique meaning, which includes market prices and the value of intangible resources and resources provided in kind. For example, volunteer time is a common resource used in educational programs. If one were to base the cost of volunteering on program expenditures, the cost would be $0. However, economists view the cost of volunteering as an **opportunity cost** of volunteers' time. Since a volunteer could have earned a wage working a job instead of volunteering, the cost of volunteering is equal to the wages foregone. According to Levin et al. (2018):

> Fundamentally, economics is the study of the allocation of scarce resources, with emphasis on the term 'scarce'. It is not sufficient to investigate how effective an intervention is–or whether one intervention causes a particular outcome. Although these are important, and in many cases, necessary investigations, they are not fully economic investigations: They do not incorporate the idea of scarcity and in particular do not weigh effects against the opportunity costs incurred to obtain the effects. (p. 3)

There are three primary methods of economic evaluation in education: (1) cost analysis, (2) cost-effectiveness analysis, and (3) benefit-cost (BC) analysis (Levin et al., 2018). We only define cost analysis and CE analysis in this report, as BC analysis is beyond the scope of this grant. For more information about BC analysis, see Levin et al. (2018).

A **cost analysis** is a summative calculation of *all resource costs* (anything with value, tangible and intangible) related to an intervention: a policy, program, or other initiative. We determine summative costs using the **ingredients method**, which "requires a descriptive itemization of all resources" (Levin et al., 2018, p.9). Once we calculate a total cost, potential partners can use the information to undergo a cost-feasibility analysis, which simply asks, "Can we afford the program given our available resources?" As stated above, this is an intended contribution of this economic evaluation.

A CE analysis is a type of economic evaluation that "compares policies or programs based on their ratios of costs to education results" (Levin et al., 2018, p.12). For example, if a math intervention costs $500 per student and results in an average increase of 10 standardized test points per student, then the CE ratio is $50 per point. Importantly, the programs being compared through CE analysis must have similar goals and outcomes (e.g., all must retain teachers, all must

increase reading test scores), which allows decision-makers to select *comparable* programs, even when they use dissimilar methods to achieve the same outcome (Yeh, 2010). For example, if a district intends to implement a reading program to improve standardized test scores, then they could compare the CE ratio of various reading programs with a unique theory of change to assess the programs' cost feasibility (can we afford it?) and choose between affordable alternatives based on which program produces the greatest effect (or a good enough effect for a lower cost). CE analysis "represents an attempt to choose approaches to educational improvement that allow the largest educational result for any resource restraint. By identifying those interventions that have the largest educational effectiveness per dollar, it is possible to maximize the impact of a given budget" (Levin, 2011, p. 412).

Benefits of Economic Evaluation

Beyond maximizing budgets, economic evaluations are useful for detailing the processes and resources needed to implement a program (Levin et al., 2018). To derive the total cost of a program, evaluators must name each resource required for the program and describe them in detail. For example, suppose a program requires a highly qualified content expert to deliver a professional development course. In that case, the cost analysis requires careful detail about that person's expertise and training to assign monetary value to their time. Additionally, to avoid counting costs outside of the scope of the program, it is necessary to have a well-defined theory of change or logic model to guide inclusion and exclusion criteria. These features of economic evaluations require that evaluators understand the program and communicate its details effectively.

Perhaps most importantly, economic evaluations are intended to inform decisions, as they present trade-offs between alternative programs (Lin et al., 2018). As stated by Levin et al. (2018), "economic evaluations must link with the requirements of decision-makers, stakeholders, and policy professionals" (p.4). In other words, if the economic evaluation will not aid in decision-making, then one must question the utility of performing an economic evaluation at all. Our work is grounded in its utility; the final products of an economic evaluation should be constructed in such a way that it will be useful for program staff and participants and for potential partners and policymakers.

The Cost Analysis Study

Methods

We used the ingredients method—as outlined in the IES (2020) *Cost Analysis: A Starter Kit* (to be referred to as "the Toolkit")—to plan the cost analysis and to structure the data collection instrument. There are five steps within the ingredients method: (1) identify the quantity and quality of all ingredients (program resources), (2) identify or estimate a price for each ingredient, (3) adjust costs for your context (such as geographical region), (4) calculate the total cost, and (5) conduct a sensitivity analysis (IES, 2020). Our data collection instrument was a Google Sheet with tabs for each domain of the CREATE organization (described under "Program Model" below), each with the following columns (adapted from the Toolkit).

- Ingredient type (category)
- Ingredient name
- Ingredient description
- Data source
- Units

- Cost per unit
- Quantity
- Funding source
- Total cost
- Assumptions

Notably, we were unable to gather data outside of the specific SEED grant expenditures, which limited our data collection efforts. For example, some program elements may have been in-kind or funded through other grants or philanthropic organizations. We could not identify those key ingredients and/or costs so they were missing from our dataset. Additionally, grant budgets and expenditures do not account for all costs (e.g., volunteer time). Thus, we could not complete the economic evaluation as expected.

Perspective

Economic evaluations typically adopt a societal perspective, necessitating the inclusion of all costs, regardless of funding source or in-kind contributions (Levin et al., 2018). Other perspectives include the individual (cost to the individual), school (cost to the school), district (cost to the district), and fiscal perspective (cost to the taxpayer or government agency) (IES, 2020; Levin et al., 2018). Our cost analysis study takes a societal perspective, as IES and other experts recommend. Thus, our cost analysis intended to include all resources required by CREATE that are above and beyond (i.e., incremental) to BAU.

Program Model

Importantly, a theory of change or logic model should guide the ingredients method, to establish parameters around which costs should be included (i.e., to avoid scope creep). To categorize costs effectively, we organized the key ingredients under five domains that together represent CREATE as an organization.

1. The Residency Program
2. Community Engagement and Continuing Education (CECE)
3. Research and Evaluation
4. Operations and Sustainability
5. Leadership

The Residency Program domain includes all aspects related to resident training and development, coaching, mentorship development, and financial support. The **CECE** domain includes broadly offered PL opportunities such as the Culturally Relevant Pedagogy Institute, CREATE 101, and Healing Centered Justice. The **Research and Evaluation** domain includes all internal and external research and evaluation activities (e.g., Empirical Education as the external evaluator). The **Operations and Sustainability** domain includes the human resources team, shared learning activities, finance team, marketing team, grant writing and mandated reporting activities, database development and upkeep, and program development and fundraising. The **Leadership** domain includes internal capacity and decision-making processes.

For the cost analysis and CE analysis, we focused on **The Residency** domain, as this part of programming directly impacts residents (re: our research questions) and would be replicated by potential partners. A simplified logic model is in Figure 1 in Chapter 1.

Assumptions

Guided by Levin et al. (2018) and the IES/CBCSE training, we adopted the following assumptions for our intended cost analysis.

- We prioritized documenting the largest costs. For relatively small costs with minimal impact on the overall program cost (e.g., printing supplies), we planned to estimate the expenses by collecting price data through standardized sources like Amazon rather than through CREATE expenditures.
- To the extent possible, we planned to use standardized costs (e.g., salaries recommended by the U.S. Bureau of Labor Statistics, supply costs from major suppliers) for comparability across studies.
- CREATE's start-up costs occurred long ago, nearly a decade before current federal funding. Thus, we planned to exclude some traditional start-up costs (e.g., purchasing laptops) from our analysis. However, we included a column in our Google Sheet data form to indicate whether each ingredient represented a start-up cost. This would allow us to conduct a sensitivity analysis to understand the extent to which excluding these ingredients might impact our overall estimates (discussed more below).
- For year 1 programming, we planned to include only incremental program costs above and beyond the cost of the traditional TPP at GSU CEHD in our list of ingredients. While we acknowledge that some aspects of CREATE programming may require additional time compared to the traditional TPP at GSU CEHD—such as matching CREATE residents to CREATE-trained CTs in a CREATE schools—we found it challenging to find a meaningful way to quantify these differences, given that staff in CREATE and school leadership positions are salaried, and we would risk double-counting costs. Therefore, we did not plan to include costs related to practicum courses, field placements, or other supports available to both CREATE residents and comparison GSU CEHD student teachers. However, we did intend to acknowledge these activities in the program logic model and when describing individual staff responsibilities in the key ingredients list.
- For year 2 and year 3 programming, we planned to include only incremental program costs above and beyond those of APS in our list of ingredients. In other words, we did not plan to include costs related to the district's regular recruitment and retention strategies and staff support. Again, we would account for these responsibilities in key ingredient descriptions and staff salaries, and we would risk double-counting.
- We planned to exclude overhead costs by GSU, such as Zoom membership costs and office space. However, we also planned to include these costs in a sensitivity analysis.
- We only planned to include incremental school costs. For example, we did not plan to cost out principals' time spent coordinating with CREATE to place pre-service teachers within their school, as we assume they would be spending time on this regardless.
- If a cost was not traceable for the 2022–23 school year, we planned to include the cost from the 2023–24 school year. We would adjust all costs for inflation to 2024 dollars.

- To annualize ingredient costs across multiple years, we planned to use an interest rate of 3%. Similarly, we planned to use a standard discount rate of 3% to calculate present value, as Levin et al. (2018) recommend for comparability across studies.

Key Program Ingredients

Working closely with CREATE leadership, we began collecting data by identifying and outlining the key ingredients for **The Residency** domain. We organized key ingredients under the following categories.

- Contractors
- Facilities
- Materials
- Participant support
- Mentor support
- Personnel
- Other

Table N1 illustrates how we would have reported the key program ingredients and cost data identified during the cost analysis.

TABLE N1. KEY INGREDIENTS AND COST DATA

Ingredient	Description and assumptions	Data source	Sensitivity analysis (Include/exclude)	Units	Cost per unit	Quantity	Total cost
Contractors							
Facilities							
Materials							
Participant support							

TABLE N1. KEY INGREDIENTS AND COST DATA

Ingredient	Description and assumptions	Data source	Sensitivity analysis (Include/exclude)	Units	Cost per unit	Quantity	Total cost
Mentor support							
Personnel							
Other							
TOTAL PROGRAM COST							$$$$
COST PER RESIDENT							$$$$
COST PER STUDENT IMPACTED							$$$$

Note. All costs adjust for inflation and are reported in 2024 dollars.

Sensitivity Analysis

While economic evaluations represent expected costs, it is still important to conduct sensitivity analyses to test assumptions made by the evaluation team. Thus, we planned to conduct three separate sensitivity analyses to adjust our estimates for some assumptions above. Specifically, we aimed to:

1. adjust the total amount of small costs (e.g., printing supplies, writing utensils) by +/- 10% to examine the potential impact on the total cost of the program,
2. include some traditional start-up costs (e.g., job posting fees, purchasing laptops) to examine the potential impact on the total cost of the program, and
3. include overhead costs (e.g., annual Zoom and Quickbooks membership) to examine the potential impact on the program's total cost.

The Cost-Effectiveness Analysis

In this section, we describe the outcomes and methods that would have been used to derive a CE ratio for **The Residency** domain of the CREATE Teacher Residency Program.

Research Question

Using effect sizes for teacher well-being (Chapter 4), what is the CREATE Teacher Residency Program's cost-effectiveness ratio?

Methods And Results

We calculate the CE ratio by dividing the total program cost (based on the cost analysis) by the measured effect to derive a dollar amount per unit of effect (Levin et al., 2018). Please see the "Cost Analysis" section above for the anticipated methods to derive the total program cost. Please see Chapter 4 of this report to understand the quasi-experimental design used to derive the effect sizes.

Limitations

At the forefront, it is important to highlight that cost analyses represent an estimate of *expected costs* and should not be reduced to an itemized list of actual costs. The end goal of economic evaluations is to produce comparable studies about the economic value of similar (CE analysis) or different (BC analysis) programs, policies, and other initiatives. However, expenditures are biased. For example, a person's level of education and years of professional experience, as well as factors like prestige bias toward their degree-awarding institution, the negotiating skills of the individual, and connections within the prospective organization, all influence their salary. Thus, pulling cost estimates from standardized sources like the U.S. Bureau of Labor Statistics allows us to eliminate *some* such bias.

Furthermore, even when derived from rigorous QEDs or RCTs, impact results are *estimates*. Study outcomes represent findings from a specific program, and each program contains a unique and highly contextual set of variables, including staff, program implementation, cohort, location (and respective geopolitical and sociohistorical context), timing, and confounding factors. Even the most seemingly "objective" measures are limited. For example, for this study, one source of employment data that we used to conduct the retention analysis (a suitable measure for the CE analysis) comes from Open Georgia—a government-maintained database. However, Open Georgia does not account for private school funding, so some teachers are not listed despite actively teaching in Georgia (and who may be otherwise considered "retained" in teaching under another definition). In addition, most teachers identify as female, and many change their names between their pre-service and in-service years due to marriage; yet, if we are unaware of these name changes, and those people are lost from our analysis.

The reality is that context and measurement introduce bias to even the most well-designed studies, and noise is unavoidable for both the economic evaluation and the impact study. That said, our CE analysis aimed to represent an *estimate we might expect of the effect per dollar spent*.

As suggested at the beginning of this chapter, the ultimate limitation of our work was the inability to collect cost data beyond SEED grant expenditures. Multiple challenges led to this issue, including retrospective data collection and leadership turnover. In the next section, we unpack these challenges and other complexities of this work.

Challenges and Complexities

By the conclusion of the 2020 SEED grant award, we could not complete the economic evaluation we initially set out to conduct. Two primary issues emerged during the process, which we believe warrant further discussion.

1. Retrospective Data Collection with Leadership Turnover
2. Methodological Complexities

Highlighting these challenges can provide valuable insights for other education researchers undertaking economic evaluations and contribute to the broader discussion about the complexities and obstacles when conducting economic evaluations in education.

Retrospective Data Collection and Leadership Turnover

A primary challenge in this work is that the CREATE Teacher Residency Program—established in 2011 at the Atlanta Neighborhood Charter School (ANCS) (GSU CEHD, n.d.)—is a long-standing program with cost and expenditure data spanning over a decade. A grassroots initiative led by educators within ANCS who sought to collaborate with GSU CEHD established CREATE. The program did not secure federal funding until 2015 when they developed a theory of change, budget, and other formal program elements for grant application purposes. As a result, critical cost data, such as startup costs like volunteer time, went undocumented during the program's first few years.

Additionally, the CREATE program has undergone significant transformation since its establishment. Originally designed to retain educators at a single school, it has evolved into a teacher residency fully integrated within the CEHD at GSU. The program now collaborates with district and school leaders across APS. This transformation highlights a common challenge in longitudinal programming: such initiatives are rarely static, and costs associated with program reform often go undocumented. This issue is exacerbated in education, where most of the cost lies in human capital. There are further complications when extended to cost-effectiveness studies, as measured program impacts also become challenging to link to a single theory of change. Thus, there is uncertainty in both the cost data and impact data.

For the 2020 SEED grant, expenditures started in the 2020–21 school year. However, this period was marked by additional challenges. The COVID-19 pandemic disrupted the 2020–21 academic year, causing CREATE to shift from an entirely in-person model to a predominantly virtual and hybrid format. This rapid shift required immediate changes to program elements, many of which went undocumented but likely impacted the program's cost. For example, CREATE would have largely met via Zoom rather than paying for meeting space within the university or an APS school site, but when and for how long is unclear. Another point of transition came in spring 2022 when the CREATE program made changes to key leadership personnel, resulting in the loss of institutional knowledge about COVID-19-related transitions, as well as program processes and cost data spanning back to 2011. For example, the program moved away from CBCT, an approach that prioritized individual well-being over collective well-being, and towards healing-centered justice. Previous cost data for CBCT-related programming was lost or incomplete. This also meant that impact data collected during the three years of implementation covered under the 2020 SEED award—some of which were influenced by old programming—would not accurately reflect the program's current model.

Methodological Complexities

The literature addresses additional complexities but the education or teacher preparation programs do not necessarily address them. These additional complexities include (a) defining the comparator, (b) the generalizability of place-based programming, and (c) the challenges and expenses of evaluations.

Defining a Comparator

Defining BAU for the comparator program can be as complex as defining the theory of change for the intervention program. Like longitudinal intervention programs, comparison programs are unlikely to remain unchanged over time (Petrou & Gray, 2005). The CREATE Teacher Residency Program is embedded within GSU CEHD and shares faculty and resources with the traditional TPP. As CREATE has expanded, it has become a normative part of CEHD programming, merging in some ways with the traditional TPP. For example, the text *Cultivating Genius: An Equity Framework for Culturally and Historically Responsive Literacy* by Gholdy Muhammad is a foundational text for CREATE, and undergraduates must now read it as part of their GSU CEHD coursework.

Despite CREATE having close ties with GSU CEHD, gathering specific program data about the traditional TPP has been challenging and aligns with issues highlighted in the relevant literature (Petrou & Gray, 2005). While our evaluation team has contacts within GSU CEHD who provide insights into the traditional TPP, we cannot require eligible comparison participants to join the study. Our limited ability to recruit non-CREATE TPP participants has resulted in only a partial characterization of their academic commitments and first-hand experiences. Given that CREATE is an intensive three-year commitment, a deficient understanding of the traditional TPP could lead to an overestimation or underestimation of the practical and monetary value of CREATE, a problem that Petrou and Gray (2005) identify. For example, we might overestimate the practical value of CREATE by assuming that comparison group participants have zero access to specific resources like mentorship when, in fact, they might have access to different but comparable resources. Alternatively, we might underestimate the societal value of CREATE by assuming that the intervention requires excessive costs beyond BAU, thereby reducing its cost-effectiveness ratio.

Place-Based Programming

As Petrou and Gray (2005) argue, assessing the transferability of a program in new contexts is a fundamental assumption before conducting an economic evaluation. It requires a deep understanding of the characteristics of participants for whom the program has been effective, as well as "the social, organisational and political settings in which the evaluation was conducted, and interactions between the programme and contextual factors (p.180)." In the case of the CREATE Teacher Residency Program, the social, organizational, and political contexts are multifaceted and continuously evolving, adding layers of complexity to any effort toward an economic evaluation.

Specifically, CREATE is deeply rooted in Atlanta, Georgia, drawing heavily on its unique historical and sociopolitical landscape to shape programming. In order to foster a deeper understanding of the historical education system in Atlanta, CREATE offers learning opportunities deeply embedded in the city's history of community organizing, social justice, and civil rights work. For example, the Baker-Hope-Clark (BHC) Summer Institute—offered to Year 1 residents—is named after three Black women who were advocates for social change in Atlanta. The BHC Summer Institute allowed participants to tour historic areas of metro Atlanta and learn about community organizing and social justice work in the area. Additionally, CREATE engages with leaders of local community organizations to learn how to integrate community-centered approaches and historical context into their educational practices. This focus on Atlanta's history is deeply connected to CREATE's goal of preparing educators to teach in the region.

Another example of CREATE's focus on place-based learning is CREATE 101, a multi-day PL opportunity where participants explore the history of education in Atlanta. While CREATE also acknowledges the relevance of histories from other locations in the South—as demonstrated through residents' visits to the Civil Rights Memorial Center Museum in Montgomery, Alabama, during their BHC Institute experience—their place-based orientation makes it challenging to assume that other contexts could scale or replicate CREATE. The program might not achieve the same effectiveness even in seemingly comparable locations, such as Birmingham, Alabama. Attempting to force-fit the model into different settings risks perpetuating a capitalist mindset of commodification and standardization, which stands in opposition to CREATE's liberatory goals.

Additionally, we often limit our understanding of the individuals within CREATE to surface-level demographic data, such as race, ethnicity, and age. While demographic information offers valuable insights, it cannot capture a group's diversity. For example, we cannot assume all Black female students who are 21 years of age share the same experiences, beliefs, values, and needs. Assuming shared experiences or perspectives based solely on demographic characteristics risks essentialist thinking (Delgado & Stefancic, 2023), further limiting the generalizability of CREATE's outcomes. Moreover, much of the impact data that we collect relies on Likert scales, which are susceptible to oversimplifying complex constructs like well-being.

From an economic evaluation perspective, these limitations pose significant challenges. The program's reliance on local context, its theoretical underpinnings that resist commodification, and the lack of detailed participant-level data complicate efforts to assess its cost-effectiveness and potential transferability.

The Challenges and Expenses of Evaluation

Many proposed solutions to challenges like defining a comparator or assessing transferability—such as diversifying, scaling up, or extending the evaluation period—introduce significant trade-offs. For example, Petrou and Gray (2005) suggest conducting additional subgroup analyses to address questions of generalizability. While they acknowledge that "this tends to increase the size and cost of studies," there is limited discussion of whether such efforts are feasible or desirable for specific programs.

From a practical perspective, forming and maintaining research partnerships with school districts has become increasingly challenging. Districts are understandably cautious about the purpose and intent of the research, making it difficult to recruit enough participants to meet the demands of a modest power analysis, let alone one that accounts for additional subgroups or longitudinal tracking of individuals. Furthermore, CREATE is specifically designed to support Black educators in Atlanta. Demanding additional evaluation resources to test the program's effectiveness for other populations—such as Black educators in New Orleans—may not align with the program's goals or priorities. Black and African American history in Atlanta vastly differs from other large metropolitan areas in the South.

Even if such efforts were deemed worthwhile, they would divert critical resources from the current program. This tension underscores a fundamental challenge in program evaluation: balancing the desire for comprehensive, transferable findings with the need to preserve the integrity and focus of the intended program. Also notable is the need to keep funds and resources in the hands of those doing the work (rather than in the hands of evaluation firms) and in service of populations that have been exploited for centuries in the United States.

Conclusion

This chapter outlines our attempted work to conduct an economic evaluation of the CREATE Teacher Residency Program, including a cost analysis and a cost-effective analysis. Given the challenges with retrospective data collection,

leadership turnover, and methodological complexities, we could not complete this project. However, we outline our intended project design and data collection and analysis activities and offer challenges and complexities for future evaluators to consider when conducting economic evaluations in education.

In the future, economic evaluations may be further complicated by the federal government's current ban on diversity, equity, and inclusion (DEI) initiatives. On April 3, 2025, the Trump administration sent a letter to all state leaders requiring them to certify that they will not engage in DEI in order to maintain federal funding for their schools (Mehta, 2025; U.S. Department of Education, 2025). In this light, economic evaluations, which require programs to articulate each component of their program clearly, risk turning an empirical process into a tool for de-funding equity-focused initiatives like CREATE. The current sociopolitical context will present evolving challenges and complexities for economic evaluations, and we look forward to collaborating with educators and economists to navigate these changes.

References

Aguilar, E. (2020). *Coaching for equity: Conversations that change practice*. Jossey-Bass.

American Institutes for Research (AIR). (2021). *Standards for Economic Evaluations Address a 'Glaring Need'*. https://www.air.org/resource/qa/standards-economic-evaluations-address-glaring-need

Butler, J., & Kern, M. L. (2016). The PERMA-Profiler: A brief multidimensional measure of flourishing. *International Journal of Wellbeing*, *6*(3), 1–48. https://doi.org/10.5502/ijw.v6i3.526

Davidson, J. R. T., & Connor, K. M. (2016*). Connor-Davidson Resilience Scale (CD-RISC) Manual*. Unpublished. Partly accessible at www.cd-risc.com

Delgado, R., & Stefancic, J. (2023). *Critical race theory: An introduction*. NYU press.

Doan, S., Greer, L., Schwartz, H. L., Steiner, E. D., & Woo, A. (2022). *State of the American teacher and state of the American principal surveys: 2022 technical documentation and survey results*. RAND Corporation. https://www.rand.org/pubs/research_reports/RRA1108-3.html

Georgia Department of Education (GaDOE). (2019). *An Assessment & Accountability Brief: 2018-2019 Georgia Milestones Validity and Reliability*. https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/Milestones/2018-19_Georgia_Milestones_Validity_and_Reliability_Brief.pdf

Georgia Department of Education (GaDOE). (2020). *Teacher Keys Effectiveness System*. https://www.gadoe.org/School-Improvement/Teacher-and-Leader-Effectiveness/Pages/Teacher-Keys-Effectiveness-System.aspx

Georgia Professional Standards Commission. (n.d.). *Certification Lookup*. Retrieved July 24, 2025, from https://www.gapsc.com/Certification/Lookup.aspx

Georgia State University College of Education & Human Development (GSU CEHD). (n.d.). *History – The CREATE Project*. https://thecreateproject.education.gsu.edu/history/

Heshmati, S., Kibrislioglu Uysal, N., Kim, S. H., Oravecz, Z., & Donaldson, S. I. (2023). Momentary PERMA: An adapted measurement tool for studying well-being in daily life. *Journal of Happiness Studies, 24*(8), 2441–2472. https://doi.org/10.1007/s10902-023-00684-w

Institute of Education Sciences (IES). (2020). *Cost analysis: A toolkit (IES 2020-001)*. U.S. Department of Education. https://ies.ed.gov/use-work/resource-library/report/technical-methods-report/cost-analysis-starter-kit

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/

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/

Lau, S., Lin, L., Zacamy, J., Jaciw, A. P., Nguyen, T., Nardi, C., & Heydel, Z. (2021). *Evaluation of CREATE's Professional Learning Activities for Experienced Educators: An Interim Report of a Formative Evaluation in Georgia*. [Unpublished report]. Empirical Education Inc.

Levin, H. M. (2011). The economics of education. *Government Law Review, 4*(2), 394–417. https://www.albanygovernmentlawreview.org/article/23840-the-economics-of-education

Levin, H. M., McEwan, P. J., Belfield, C., Bowden, B. A., & Shand, R. (2018). *Economic evaluation in education* (3rd ed.). Sage Publications

Lin, Y., Cheng, A., Hecker, K., Grant, V., & Currie, G. R. (2018). Implementing economic evaluation in simulation-based medical education: Challenges and opportunities. *Medical Education, 52*(2), 150-160.

Maslach, C., Jackson, S. E., & Leiter, M. P. (2018). *Maslach Burnout Inventory Manual: Fourth Edition.* Menlo Park, CA: Mind Garden, Inc.

Mehling, W. E., Acree, M., Stewart, A., Silas, J., & Jones, A. (2018). The Multidimensional Assessment of Interoceptive Awareness, Version 2 (MAIA-2). *PLOS ONE, 13*(12), e0208034. https://doi.org/10.1371/journal.pone.0208034

Mehta, J. (2025, April 3). Education Dept. warns schools: Eliminate DEI programs or lose funding. NPR. https://www.npr.org/2025/04/03/nx-s1-5350978/trump-administration-warns-schools-about-dei-programs

Minckler, C. H. (2011). *Teacher social capital: The development of a conceptual model and measurement framework with application to educational leadership and teacher efficacy* (Publication No. 3489646) [Doctoral dissertation, University of Louisiana at Lafayette]. ProQuest Dissertations & Theses Global.

Muhammad, G. (2020). *Cultivating Genius: An Equity Framework for Culturally and Historically Responsive Literacy*. Scholastic.

Muhammad, G. (2023). *Unearthing Joy: A Guide to Culturally and Historically Responsive Curriculum and Instruction*. Scholastic.

Neff, K. D. (2003). Development and validation of a scale to measure self-compassion. *Self and Identity, 2*(3), 223–250. https://doi.org/10.1080/15298860309027

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

Petrou, S., & Gray, R. (2005). Methodological challenges posed by economic evaluations of early childhood intervention programmes. *Applied Health Economics and Health Policy, 4*, 175-181.

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine, 16*(3), 199–202. https://doi.org/10.1093/abm/16.3.199

Tiesteel, E., Watkins, R. C., Stringer, C., Grigorie, A., Sultana, F., & Hughes, J. C. (2024). Where Are the Costs? Using an Economic Analysis of Educational Interventions Approach to Improve the Evaluation of a Regional School Improvement Programme. *Education Sciences, 14*(9), 957. https://doi.org/10.3390/educsci14090957

U.S. Department of Education. (2025, April 3). *Reminder of legal obligations undertaken in exchange for receiving federal financial assistance and request for certification under Title VI and SFFA v. Harvard*. https://www.ed.gov/media/document/reminder-of-legal-obligations-undertaken-exchange-receiving-federal-financial-assistance-and-request-certification-under-title-vi-and-sffa-v-harvard-april-3

Yeh, S. S. (2010). The cost-effectiveness of NBPTS teacher certification. *Evaluation Review, 34*(3), 220-241.

This appendix is a supplement to a report that was prepared for the Office of Elementary and Secondary Education (OESE) under Award Number S423A200035 funded by the SEED grant program.

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

Reference the report: Jaciw, A. P., Lau, S., Nardi, C., Zacamy, J., Lin, L, Maurer, L., & Waltower, M. (2026). Evaluation of Georgia State University's CREATE Teacher Residency Program—Cohorts 6 through 8: A Final Report of a Quasi-Experiment in Georgia. (Empirical Education Rep. No. Empirical_GSU-7036-FR-2026-O.3). Empirical Education Inc. https://www.empiricaleducation.com/create/