Mathematica
Progress Together

Findings from an Evaluation of KIPP School Leadership Programs

Final Report

September 30, 2022

Elias Walsh, Alicia Demers, Jeffrey Terziev, Erin Boyle, Elisa Steele, and Ijun Lai

This page has been left blank for double-sided copying.

Findings from an Evaluation of KIPP School Leadership Programs

Final Report

September 30, 2022

Elias Walsh, Alicia Demers, Jeffrey Terziev, Erin Boyle, Elisa Steele, and Ijun Lai

Submitted to:

KIPP Foundation
135 Main Street, Suite 1700
San Francisco, CA 94105

Submitted by:

Mathematica
600 Alexander Park
Princeton, NJ 08540

This page has been left blank for double-sided copying.

Contents

Tables

Figures

Exhibits

I. Introduction

A. Why the study?

In 2018, KIPP was awarded a Supporting Effective Educator Development grant by the Office of Innovation and Improvement at the U.S. Department of Education. Under this grant, KIPP supported, updated, and improved three school leadership programs it administers to prepare new principals: (1) the Successor Prep program, which prepares new principals to lead existing schools; (2) the Fisher Fellowship program, which prepares new principals to lead new KIPP schools; and (3) the KIPP Leadership Design Fellowship (KLDF) program, which disseminates effective school leadership development strategies outside of KIPP. The grant includes a requirement that these programs be independently evaluated.

This evaluation of the KIPP school leadership programs aims to support sustained implementation of effective programs to prepare and support school leaders. Effective school leaders, in turn, support teachers and help students achieve better outcomes (George W. Bush Institute 2016; RAND 2017). In addition to supporting KIPP to improve the design of the KIPP school leadership programs, the evaluation findings can guide the replication of effective tools and strategies that increase the number of highly effective principals among other charter and traditional public schools as well as across the principal preparation field.

B. The KIPP school leadership programs

1. The Successor Prep program

The Successor Prep program prepares participants to become principals at existing KIPP schools. Without support, some principal transitions are disruptive, causing dips in student outcomes, and the new principals may need three or more years to reach their full impact on their new schools (Clark et al. 2009; Coelli and Green 2012; Miller 2013; Walsh and Dotter 2018). The Successor Prep program aims to reduce possible disruptions and enable new principals to reach their full impact at their schools sooner.

The 15-month program begins in January of the school year before participants will become principals and continues to support the participants through their first year as a principal. KIPP also trains participants from outside of KIPP who will not become KIPP principals and may become principals of non-KIPP schools. The participants are nominated by KIPP regional staff, and those nominated complete an application package that includes written essays, video submissions, and information about their background and experience.

In January, most of the participants are placed as assistant principals, typically in the schools they will take over in the fall. This allows the participants to shadow the schools' current principals and gain experience as school leaders. To support the participants during the program, each is matched with a leadership coach who provides one-on-one support throughout the program. The leadership coaches, who are former school principals or have experience developing school leaders, also help each participant identify their key strengths and areas for improvement.

During the summer, the participants attend the three-week summer institute and complete required coursework. The Institute is led by education school faculty and other educational leaders. The Institute and the coursework help prepare the participants to take over their schools in the fall. Prior to taking over

their schools in the fall, the participants develop strategic leadership plans that detail how they will take over and lead their schools. In the plans, the participants must identify the challenges they anticipate facing, their desired outcomes, strategies for achieving their desired outcomes, and an approach to measuring their success. Program staff guide the participants in developing plans that will help them as they take over leadership of their schools; the participants are not evaluated on how well they achieve the goals they specify in their plans.

After the participants take over their schools in the fall, they receive additional training and support during the first school year. Across the full 15-month program period, the participants attend five two- to three-day in-person training sessions focused on the elements of effective and lasting organizations, including modules on change management, transition planning, strategic planning and execution, instructional coaching, and performance management. The participants also attend up to two brief residencies or focused school visits to high-performing KIPP schools across the country and continue working with their leadership coaches as needed.

Implementation of some aspects of the program, such as the content of the summer institute and the use of residencies, varied over time and across regions.

2. The Fisher Fellowship program

The Fisher Fellowship program trains new leaders to become founding principals of new KIPP schools. The program provides a year of training and support to individuals selected to become principals. The program begins in May and continues for one year. Participants then become a principal of a KIPP school in the fall after the end of the program.

The program includes a four-week summer institute, four multi-day training sessions held throughout the program year, eight residencies in KIPP schools or school visits, and support from a leadership coach. The residencies include time in schools from the KIPP region of the new school the participant will lead and schools in other KIPP regions.

As part of this programming, participants create a School Launch Plan that will guide their work to develop and lead a new school. Participants describe plans for engaging parents and the community to develop a vision for the school, plans for hiring and managing staff, a curricular and instructional philosophy and model, an approach to serving special populations, and a financial budget for the first two years. Because new KIPP schools typically start with one or two grades and grow by an additional grade each year, participants describe how their plans for the school will evolve as the school reaches full enrollment over multiple years.

The Fisher Fellowship program receives 200 to 300 applications each year from educators inside and outside of KIPP schools. Applicants undergo multiple rounds of interviews, and some are invited to a selection event. At each interview stage, applicants are evaluated against KIPP's Leadership Competency Model and School Leader Readiness Criteria. At the selection event, candidates participate in four interviews with experienced school leaders and KIPP regional staff. A selection committee of regional and national KIPP leaders evaluates candidates' performances in the interviews using rubrics with components that assess competency in instructional leadership, culture, self-awareness, decision making, communication, and others. Each candidate is scored by multiple committee members; the scores are then averaged across components and raters, and the highest-scoring applicants are selected into the program.

3. The KIPP Leadership Design Fellowship (KLDF) program

The KLDF program trains senior leaders of public school districts, charter school systems, and leadership training organizations on KIPP's leadership development model. Three summit events, each held over multiple days, explore a different theme such as KIPP's theory of leadership development, leadership pipeline development, adult learning, and KIPP's formal programming components. Throughout the three summits, participants observe a school, a portion of a principal selection event, and a portion of the Successor Prep program's summer institute. These opportunities allow participants a deeper understanding of how KIPP's approach is applied in schools, leadership selection, and trainings.

For each of these three programs, the length, number of participants that KIPP has typically trained each year, planned first year as a school principal, and program activities are summarized in Table I.1.

Table I.1. Overview of program characteristics and activities

	Fisher Fellowship	Successor Prep	KLDF
Program characteristics			
Length of time	12 months	15 months	6 months
Start month	May	January	September
First school year as principal in a school	After completing program	During 15-month program period	n.a.
Number of participants per year	10–15	About 25	15–35
Program activities			
Multi-day training sessions	✓	✓	✓
Coaching	✓	✓	
Summer institute	✓	✓	
Residencies	8	Up to 2	
Active after-program support	✓		
School visit	✓		✓

n.a. = not applicable.

C. Research questions

To understand how KIPP school leadership programs support the development of effective school leaders, this evaluation examined three research questions about KIPP school leadership programs:

1. What motivates participation in the leadership programs, which aspects of the programs do participants find the most useful, and are participants in the programs able to apply the lessons of the program in their own settings?
2. How do outcomes for teachers and students in schools that received Successor Prep principals compare to outcomes in similar KIPP schools that were not led by Successor Prep principals?
3. Is the Fisher Fellowship selection instrument a reliable measure of candidate potential, and do Fisher Fellowship participants with higher total or component scores on the selection instrument become more effective school leaders?

To answer the first question, we conducted a descriptive analysis of results from surveys that we designed and administered to Successor Prep and KLDF program participants. The surveys assessed participants' motivation, satisfaction, needs, experiences, and suggestions for improvement.

To answer the second question, we compared outcomes for teachers and students in schools that received Successor Prep principals to those for other KIPP schools in a comparison group. We constructed the comparison group to be similar to the group of Successor Prep schools based on the characteristics of students and teachers in the schools in the years before a newly trained Successor Prep principal began leading a school.

To answer the third question, we measured (1) the relationship between the three competencies measured in the instrument that KIPP uses to select candidates for the Fisher Fellowship program and how each competency contributes to overall scores, (2) the reliability of the instrument that KIPP uses to select candidates for the Fisher Fellowship program—including its internal consistency and interrater reliability—and (3) the predictive validity of the selection instrument, which is the extent to which scores on the instrument are associated with outcomes that reflect effective leadership.

These three research questions are broadly defined to capture the aims of the evaluation. In describing the analyses we conducted below, we also define the more specific research questions we addressed.

D. Prior research on the importance of school principals

School principals can produce meaningful impacts on student outcomes and may be responsible for as much as 25 percent of schools' contributions to student achievement (Chiang et al. 2016; Grissom et al. 2015; Branch et al. 2012; Dhuey and Smith 2014). However, novice principals might need a few years to develop their skills and settle into their school before they can reach their full potential. For example, novice New York City principals' contributions to student achievement improved between the principals' first and third years of experience (Clark et al. 2009). Even experienced principals go through an adjustment period when transferring to a new school; it may take three or more years for the principal to reach full impact in the new school (Coelli and Green 2012; Walsh and Dotter 2020).

Without support, some principal transitions are disruptive, causing dips in student outcomes (Miller 2013). Efforts to limit the negative impact of principal transitions on student outcomes by better preparing novice principals have had modest success. For example, a study of 10 districts that recruited principals from New Leaders—a program designed to recruit, train, and support highly effective principals—found that cumulative exposure to these principals over three years improved test scores by approximately 0.03 standard deviations of student achievement (Gates et al. 2014).

This evaluation of KIPP's school leadership programs aims to build on this knowledge by learning about the potential of these programs to support the development of effective school leaders.

E. Overview of findings

Based on the surveys of KIPP program participants, Successor Prep and KLDF participants expressed satisfaction with the programs and generally reported that they were able to apply the lessons of the program in their own settings. In particular, the participants valued cohort collaboration opportunities in the programs and appreciated the high caliber of program facilitators. The respondents also had some suggestions for improvements. Many expressed a desire for more follow-up opportunities with their cohort. Participants of both programs felt the diversity and equity training was inadequate and reported struggling to develop talent and manage competing priorities amid resource constraints. KLDF participants also experienced challenges cultivating a strong organizational culture.

Next, outcomes in Successor Prep schools generally appeared similar to outcomes in other KIPP schools. A majority of Successor Prep principals led their placement schools for at least three years, with many staying a fourth year, and the tenure of these principals was statistically indistinguishable from the tenure of other similarly experienced comparison principals in KIPP schools, although we cannot rule out substantial differences. Similarly, math and reading test scores, student retention rates, and teacher retention rates in Successor Prep schools were statistically indistinguishable from these outcomes in a group of similar KIPP schools, although we cannot rule out substantial differences.

Finally, the Fisher Fellowship selection instrument measures three distinct dimensions of leadership potential, as intended, and is also reliable. However, scores on each competency, when combined, do not contribute equally to a candidate's overall selection score, and we identified three items that had relatively low levels of one type of reliability. Due to data limitations, we were not able to assess whether participants with higher total or component scores on the selection instrument become more effective school leaders.

This page has been left blank for double-sided copying.

II. Data sources

The study involved the following three types of data:

A. Program participant lists

KIPP provided lists of the Successor Prep, Fisher Fellowship, and KLDF program participants between the 2014–2015 and 2018–2019 school years. The study used the participant lists to determine which KIPP schools received Successor Prep principals for the analysis of outcomes in Successor Prep schools. KIPP also provided current contact information for the participants, which the study used to administer surveys to former KLDF and Successor Prep participants.

B. Survey data

Mathematica, in collaboration with KIPP, designed and administered surveys to KIPP school leadership program participants to help KIPP understand how the programs can be improved to better support school leaders. Specifically, the surveys aimed to identify the professional roles of participants, participants' motivations for participating in the program, participants' satisfaction with the programs, how participants are bringing the program's tools and strategies to their own settings, and where they might need more support. Mathematica administered the surveys to Successor Prep principals and former KLDF program participants and during 2021 and 2022. We administered the surveys to participants within one to five years after they had completed a program. The survey data include responses from 46 Successor Prep leaders and 56 KLDF participants. The survey did not include Fisher Fellowship principals because KIPP planned to phase out the program and replace it with a program modeled after the Successor Prep program.

C. Administrative data

For the analysis of outcomes in Successor Prep schools and the analysis of the Fisher Fellowship selection instrument, KIPP provided the following:

- A list of KIPP schools in each of the school years from 2012–2013 to 2020–2021, including 265 unique schools across 33 KIPP regions, which provided us with information about when schools opened and for how long they were in operation.
- A list of the principals of all KIPP schools between the 2012–2013 and 2020–2021 school years, including 598 unique principals, which allowed us to identify new principals, measure principals' tenure, and follow them if they began leading a different school.
- Students' K–8 standardized Measures of Academic Progress (MAP) test scores from the 2010–2011 to 2018–2019 school years for reading and from the 2011–2012 to 2018–2019 school years for mathematics. We used spring test scores for the analysis of outcomes in Successor Prep schools, and fall and spring test scores for the analysis of the Fisher Fellowship selection instrument. KIPP provided test scores for 224 of the 265 KIPP schools, although some of these schools did not have test scores in every school year (for example, new KIPP schools that opened during the study time frame) and some were high schools in which the MAP was not administered to students. We standardized the test scores at the school level, by (1) averaging the test scores for students within KIPP schools separately by grade, subject, and year; (2) standardizing the school-level scores by

grade, subject, and year; and (3) averaging the standardized scores across grades within schools, separately by grade and subject.

- The percentage of each school's students in each of the school years from 2010–2011 to 2018–2019 who returned to the school the following year (excluding students who transitioned out of the school's grade span). KIPP provided this information for 172 of the 265 KIPP schools.
- The percentage of each school's teachers in each of the school years from 2010–2011 to 2018–2019 who (1) returned to teach in the school the following year and (2) returned to any KIPP school in any position (including non-teaching roles) the following year. For example, a teacher who transitioned to an assistant principal role in another school would be counted as retained for the second measure. KIPP provided this information for 172 of the 265 KIPP schools.
- Scores for 80 Fisher Fellowship candidates on KIPP's School Leader Readiness Criteria tool used to select participants during the 2017 and 2018 selection rounds. These scores included overall scores; scores for each of three competency areas, including Culture & Self-Awareness, Vision & Goals, and Instructional Leadership; and scores on 34 individual items within these areas. Two raters provided scores for each candidate.

III. Results from the surveys of Successor Prep and KLDF program participants

Summary

- Participants in both programs valued opportunities to collaborate with or learn from other program participants and others in the field. In fact, many desired more of these opportunities.
- Participants in both programs appreciated the high caliber of program facilitators and speakers and viewed them as keys to program success. KLDF program participants appreciated the substance and theory-driven curriculum of the program.
- Participants in both programs felt the diversity training—in terms of both depth and breadth—was inadequate.
- Participants in both programs generally reported that they were able to apply the lessons of the program in their own settings but struggled to develop talent and manage competing priorities amid resource constraints. KLDF participants also experienced challenges applying program content when it came to cultivating a strong organizational culture.

A. Detailed research questions and survey background

Understanding how participants perceive the KIPP school leadership programs can provide insight about how the programs support the development of effective school leaders. These lessons can inform improvements to KIPP's programs or to school leadership programs in other settings. We administered surveys to current principals who participated in KIPP's Successor Prep program and former participants in the KLDF program to understand the following:

1. What motivates participation in the leadership programs, and do the programs meet participants' expectations?
2. Which aspects of the two school leadership programs do participants find the most useful?
3. Are participants in the two leadership programs able to apply the lessons of the program in their own settings?

In 2019 and 2020, Mathematica, in collaboration with KIPP, designed and administered a first round of surveys to participants who completed KIPP school leadership programs to help KIPP understand how the programs can be improved to better support school leaders. Although response rates were low and survey administration was interrupted by the COVID-19 pandemic, this first round revealed several key findings:

- Respondents generally expressed satisfaction with both the content of the programs and the delivery of that content, and they believed that the programs were beneficial for their own professional development.
- Respondents reported that the programs proved effective at expanding their leadership skills, allocating an appropriate amount of time to each skill, and providing lessons that they regularly apply in their current jobs.
- Respondents in both programs tended to agree that the most useful program features included the high-quality facilitators and the unique networking opportunity. Successor Prep program participants

also highlighted the orientation and residencies, and the KLDF program participants highlighted the modeling of good professional development as useful.

- Respondents also reported a desire for follow-up learning opportunities, deeper investigations into real-world challenges, more time on talent management and school culture in the Successor Prep program, and more time and resources for how to effectively coach leaders in the KLDF program.
- When implementing program practices, Successor Prep participants attributed their success in creating and implementing strategic action plans to the program, whereas KLDF participants benefited in their hiring and coaching practices.
- Participants from both programs reported implementation challenges in balancing competing priorities and gaining buy-in from stakeholders.

To build on these findings, Mathematica administered a second round of surveys to KLDF and Successor Prep program completers in 2021 and 2022. The second-round surveys were administered to new completers from additional program cohorts of the Successor Prep and KLDF programs, and to completers from program cohorts already surveyed in the first round because response rates had been low. We administered the surveys to completers within one to five years after they had completed a program. Similar to the first survey, the second survey aimed to identify the professional roles of participants, how participants are bringing the program's tools and strategies to their own settings, and where they might need more support.

In this chapter, we summarize findings from the survey, emphasizing lessons for KIPP and other organizations to apply to their leadership programs. KIPP engages in continuous improvement of its leadership programs, thus, some of the action items we suggest throughout this chapter have already been planned or implemented since we surveyed program participants. Moreover, KIPP has recently phased out the Successor Prep program, replacing it with a new program administered by each KIPP region. While the action items we propose are influenced by the opinions and experiences of program participants who predate these organizational changes, most remain applicable to KIPP's plans going forward as well as to other organizations facing similar challenges.

B. Study sample

1. Sample frame and response rate

Mathematica administered the second round of surveys from December 2021 to June 2022 to five cohorts of participants from the Successor Prep program and four cohorts of participants from the KLDF program.

The second-round sample frame for the Successor Prep survey included school principals who had participated in one of five cohorts: those who began the program in January of 2016, 2017, 2018, 2019, or 2020. Participants who began the program in January of a calendar year generally took over as principal of a new school in the fall of that year, with the exception of some participants who began the program after they had already become a principal. Throughout this chapter, we refer to each cohort of participants in the programs based on the year during which the program began. When surveyed during the 2021–2022 school year, the earliest Successor Prep program completers were in their fifth year as school leaders (with the first year of the role concurrent with the Successor Prep program), whereas the latest cohort of completers were in their second year as school leaders. To be eligible for the survey, respondents must have confirmed they participated in the program in one of these years and needed to be

leading a KIPP school at the time they completed the survey. We had previously administered the first-round survey to participants in the 2017 cohort. These principals were included again in the second round because of a low response rate (23 percent) in the first round and because we wanted to administer them the updated surveys.

The second-round sample frame for the KLDF survey included participants in four cohorts: those who began the program in 2015, 2016, 2018, and 2019. The program was not offered in 2017. We had previously administered the first-round survey to participants in the earliest three cohorts, but the response rates had been low (14 percent).

Response rates were substantially higher in the second round than they were in the first round (Table III.1), particularly for the Successor Prep program. In part, the low response rates from the first round were due to closing the survey early because of the COVID-19 pandemic, but we also made changes to improve response rates in the second round. In the second round , we offered $50 gift card incentives to respondents, sent reminder emails and placed reminder phone calls, and, for Successor Prep principals, worked closely with KIPP to send reminders to the principals. Phone calls have been shown to be an effective strategy for improving response rates for principals (Neal et al. 2020), although we were not able to obtain phone numbers for all principals in the sample frame. Even with these efforts, it is possible that response rates were negatively affected due to ongoing time and resource demands associated with managing schools and organizations during the COVID-19 pandemic. Given the demands on these principals' time and lower response rates obtained from principals in other settings (such as those obtained in Neal et al. 2020), the Successor Prep response rate of 55 percent exceeded expectations.

Table III.1. Survey timing and response rates, by survey source

Program	Participant cohorts	Survey period	Participants administered the survey	Respondents who screened out	Remaining sample frame	Respondents with at least a partial response (response rate)
Successor Prep	2016, 2017, 2018, 2019, 2020	December 2021 to June 2022	95	11	84	46 (55%)
KLDF	2015, 2016, 2018, 2019	December 2021 to April 2022	180	3	177	56 (32%)

Notes: We surveyed Successor Prep participants in the 2016–2019 cohorts between December 2021 and April 2022, and participants in the 2020 cohort between February 2022 and June 2022. We waited to survey the Successor Prep 2020 cohort participants until they were well into their second school year as school leaders.

KIPP confirmed a list of 22 Successor Prep program participants who were no longer school leaders at a KIPP school at the time of survey administration. In addition, we did not have contact information for one Successor prep program participant. We did not administer the survey to these 23 participants, and they are not counted in any column in the table.

The sample frame refers to the number of participants who accessed the survey, consented to participate, and confirmed their eligibility to participate. Successor Prep program participants were eligible to participate if they confirmed they began the program between 2016 and 2020 and were a school leader at a KIPP school at the time they completed the survey. Similarly, KLDF program participants were eligible to complete the survey if they confirmed they completed the program in one of our study years and that they were currently working in an education-related field at the time they completed the survey.

Respondents with at least a partial response refers to the number of participants who provided sufficient responses to be included in our analyses presented in this report.

For KLDF leadership program participants, response rates were lower than for the Successor Prep principals (32 percent versus 55 percent). These lower response rates were likely due to our inability to reach participants at their professional email addresses, suggesting participants may have transitioned to

new organizations or were unable to receive our outreach emails due to email blocking filters that professional organizations often have. Moreover, as noted, KLDF program participants who responded to the survey but indicated that they did not complete the program in one of our study years or that they were not working in an education-related field at the time they completed the survey were not asked to complete the rest of the survey. We do not have similar information for participants we were unable to reach, and some of these nonrespondents may also have been ineligible for these same reasons. If that is the case, ineligible nonrespondents may be contributing to the low response rate.

Not all participants who responded completed the full survey. These partial responses are included in our analyses when possible; therefore, the samples going into each analysis do not always equal the number of respondents in the last column of Table III.1.

2. Limitations

Sample nonresponse. Although the response rates met or exceeded expectations for the study, they do give rise to two important limitations that should be considered when interpreting the results. First, we do not have the sample sizes necessary to conduct robust subgroup analyses (such as disaggregating by KIPP region or program cohort). It is possible that our analysis of all participants masks important differences across different groups. Second, our sample is highly unlikely to be representative of all leadership program participants. Specifically, we may have responses from many participants who were very happy with their program experiences and have been able to successfully implement the lessons of the programs in their current roles, or we could have responses from participants who were unhappy with their experiences and used this survey as an opportunity to let those views be known. We also do not observe the perspectives of participants for whom preparation was inadequate to the extent that they did not become principals or quickly left the role even though they completed the program. Although these limitations are significant, the responses we received can still help to identify program features and components that were useful for the respondents and highlight opportunities for improving the programs.

Principals who left KIPP. Another limitation to our analyses is unique to the Successor Prep sample. Specifically, the Successor Prep survey sample does not represent all participants who completed the program, but rather only those who became principals and remained principals up until the time of the survey. Any differences in responses we observe between those who have been principals only a short time (for instance, for one or two years) and those who have been principals longer could be due to the accumulated experience of the latter group, their greater recall error about experiences during the program, or differences between principals who remain in KIPP compared to those who left.

3. Background characteristics of the sample

Although our survey respondents may not fully represent all program participants, understanding the characteristics of respondents can provide general insight into program participation and design. Table III.2 provides sample characteristics for both sets of survey respondents. In general, the average age of a Successor Prep participant was 35 years, whereas the average KLDF participant was 41 years old. Most participants held a master's degree; others held various other degrees. KLDF respondents reported higher rates of advanced degrees. The sample was racially diverse, although relatively more Successor Prep respondents were Black or African American and relatively more KLDF respondents were White. Most respondents from both programs identified as non-Hispanic. Most participants were female and had caregiving responsibilities, with the latter suggesting some participants may appreciate remote learning options or limiting extended travel in leadership programs.

Suggested Action: Offer remote programming options

Consider continuing to offer remote programming options and limiting extended travel for leadership programs to accommodate participants with caregiving responsibilities.

Table III.2. Sample characteristics of survey respondents

Characteristic	Successor Prep		KLDF	
	Mean or % (standard deviation)	Sample size	Mean or % (standard deviation)	Sample size
Age	35.0 (5.4)	43	41.6 (7.9)	51
Gender (%)				
Female	81%	43	73%	51
Male	19%	43	27%	51
Race (%)				
White	38%	45	52%	52
Black or African American	56%	45	31%	52
Asian	2%	45	6%	52
American Indian or Alaska Native	0%	45	2%	52
Other	7%	45	0%	52
Unknown	6%	45	15%	52
Ethnicity (%)				
Hispanic or Latinx	16%	45	13%	48
Non-Hispanic	84%	45	88%	48
Marital status (%)				
Married or in a domestic partnership	71%	45	71%	52
Unmarried	29%	45	29%	52
Primary caregiver responsibilities (%)				
Children ages 0 to 6	42%	45	25%	52
Children ages 6 to 18	38%	45	38%	52
Children ages 18+	13%	45	17%	52
Older adults	2%	45	2%	52
Not a caregiver	33%	45	37%	52
Highest earned degree (%)				
Bachelor's degree	13%	45	0%	52
Master's degree	64%	45	54%	52
Education specialist degree	16%	45	17%	52
Principal certification	7%	45	n.a.	n.a.
Doctoral degree	0%	45	21%	52
Professional degree	0%	45	8%	52

n.a. = not applicable.

4. Previous, current, and aspirational work experiences

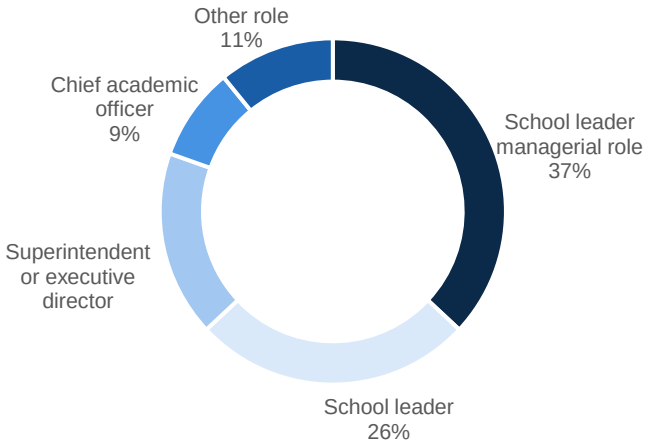
Understanding the type of environments, roles, and responsibilities that participants are familiar with before attending KIPP leadership programs, or that they aspire to take on after the program, can help KIPP appropriately tailor the leadership programs to draw on participants' prior experiences and leadership or organizational goals.

a. Previous, current, and aspirational work experiences of Successor Prep respondents

As expected, most of the Successor Prep program participants were affiliated with KIPP schools before they began the training program. Immediately before attending the program, all respondents held school leadership positions. Most respondents (61 percent) were assistant principals; a smaller share (30 percent) were school leaders or principals. A small number of respondents (9 percent) were deans of students or instruction. By design, at the time they completed the survey, all Successor Prep respondents were school leaders at KIPP schools.

All of the responding Successor Prep principals reported career goals that included holding leadership positions in an educational setting. As shown in Figure III.1, about one-quarter aspired to remain a school leader, while a larger share aspired to have a role managing school leaders in some capacity, such as a school leader manager, managing director of schools, or head of schools. Others had career goals such as becoming a superintendent or executive director, and a small share were interested in working their way toward becoming chief academic officers or holding other roles, such as talent and recruitment managers or teachers.

Figure III.1. Aspiring roles of Successor Prep participants



Suggested Action: Leverage participants' prior leadership experience and diverse career goals

Consider providing or expanding opportunities for participants to reflect on program content in the context of prior school leadership experience.

Consider acknowledging participants' varied career goals by demonstrating the applications of program content to roles managing school leaders, in addition to school principal roles.

b. Previous, current, and aspirational work experiences of KLDF respondents

Before attending KLDF, most participants (91 percent) were involved either directly or indirectly in training or preparing school leaders, and almost all had experience as a regional leader (or its equivalent) or were involved in the leadership development process. Specifically, 71 percent of respondents were regional leaders, superintendents, executive directors, or academic officers, and 27 percent were leadership development staff. Before attending KLDF, most participants had experience working for a school district or charter school organization. Specifically, 48 percent of the survey respondents worked for a school district or a state or local education agency, and 29 percent worked for a charter school or charter school management organization. Another 18 percent worked for an education nonprofit, and 5 percent worked in education consulting.

By design, at the time they replied to the survey, all KLDF respondents in our analyses were employed in an education-related field; 38 percent of them reported being employed in the same position they held when they started the program. Respondents held a variety of jobs, with some of the most common titles including executive director, superintendent or associate/deputy superintendent, and director of leadership development.

A vast majority of respondents (88 percent) were involved in training or preparing school leaders and candidates when they completed the survey. Approximately two-thirds of these respondents were working directly with school leaders and candidates, and the remaining third were working with staff members who work directly with school leaders or candidates. Common specific responsibilities in these roles included supporting and mentoring school leaders (68 percent), growing the pipeline of effective school leaders (66 percent), and identifying and recruiting school leaders (52 percent). Most respondents reported interest in continuing to work in the education sector, specifically in leadership development roles or as school administrators. A smaller number of respondents expressed having career goals in education consulting, as an executive director or chief people officer outside of a school system, or expressed a desire to remain in their current position.

Suggested Action: Leverage participants' prior experience training and mentoring school leaders

Consider providing or expanding opportunities for participants to discuss prior experiences training school leaders and reflect on program content in the context of prior experiences.

C. Research Question 1: What motivates participation in the leadership programs, and do the programs meet participants' expectations?

Understanding respondents' motivations for participating in KIPP leadership programs and how well their goals were met by the programs can help KIPP consider opportunities to expand the applicant pool and refine program offerings. Moreover, KIPP can also build on its own leadership pipeline knowing how Successor Prep program participants' expectations in terms of roles and placements fared against reality when they came out of the program.

Expected placements of Successor Prep program completers

When starting the Successor Prep program, 96 percent of respondents expected to lead a specific KIPP Successor school after completing the program (including some who were already leading that school), and all of those participants went on to lead the anticipated school after program completion. However, other participants may not have led their anticipated schools; participants who did not get placed in a KIPP school are not included in our sample.

To address the first question about motivation, we asked respondents of both surveys to identify up to three of the most important reasons for participating in their respective programs and how well the program met their expectations in fulfilling their goals. To address the second question regarding job placement expectations for Successor Prep participants, we asked respondents a series of questions to understand how they experienced and navigated the leadership pipeline. Responses to the questions about expectations are summarized in the box on the previous page.

1. Motivation for participating in the leadership programs

We presented survey respondents with a set of options for why they participated in a program. These options are listed in Appendix A.1.

An overwhelming majority of Successor Prep respondents and most KLDF respondents named the desire to improve their leadership skills as a top reason for participating in the program. Moreover, one-third of Successor Prep respondents and one-quarter of KLDF respondents participated primarily to have an impact on more children. Meanwhile, larger shares of KLDF respondents than Successor Prep respondents reported the desire to learn how KIPP trains and prepares school leaders (77 percent of KLDF respondents compared to 11 percent of Successor Prep respondents) and wanting to do so more effectively themselves (75 percent compared to 9 percent). This is consistent with the aims and purpose of each program—Successor Prep participants are training to becoming school leaders themselves and are already established in the KIPP model, whereas KLDF program participants generally train leaders and are motivated to learn about the KIPP leadership model.

Many respondents reported participating in the program for other reasons, including being encouraged to do so by a supervisor or a superior (43 percent of Successor Prep and 32 percent of KLDF respondents). Consistent with the aims of each program, some Successor Prep respondents participated in the program because it was required for their current or future position (33 percent cited this as a top reason for participating), but this was not the case among KLDF respondents.

Few KLDF respondents participated in the program primarily to learn about KIPP diversity, equity, and inclusion initiatives, although as discussed in the next section, many expressed that the program's diversity training could be improved. Similarly, although few respondents participated primarily to learn from other participating organizations, many expressed this type of collaborative setting to be a strength of the program.

Few respondents participated primarily because they saw the opportunity as a challenge or as a steppingstone to achieve a career goal. Similarly, no respondents from either program reported participating to increase their earnings. These findings do not mean the reasons mentioned here did not factor into their decision to participate in the program, but they were not the most important reasons for doing so.

Suggested Action: Expand application pool by tailoring promotional materials to participant motivations

For both programs, continue to emphasize leadership skill-building as a primary programmatic goal.

For the KLDF program, continue to emphasize program opportunities to learn about and become proficient in KIPP's leadership models.

2. Meeting expectations of participants

A well-defined program has objectives that align with participants' goals, and a well-executed program facilitates the realization of these goals. To better understand how the Successor Prep and KLDF programs lived up to their expectations, we asked respondents to reflect on their goals going into the program and consider whether the programs helped fulfill them. When looking across all goals identified by participants, we found that both programs overwhelmingly met or exceeded expectations: KLDF and Successor Prep respondents indicated that the program met or exceeded their expectations in fulfilling almost all goals (98 percent and 99 percent of goals, respectively). The KLDF program exceeded expectations in fulfilling about half of respondents' goals (51 percent) and met expectations for 46 percent of their goals, whereas Successor Prep respondents indicated their program less often exceeded expectations goals (38 percent had their goal expectations exceeded) and more often met expectations (61 percent of goals). Finally, the respondents indicated that the programs did not meet their expectations in fulfilling the small remaining share of goals (1 percent of KLDF and 2 percent of Successor Prep goals).

Among the three most frequently selected reasons for participating in the Successor Prep program, all respondents indicated the program met or exceeded expectations when helping to fulfill this goal (Figure III.2). Figure III.3 shows that for two of the top three reasons for participating in the KLDF program, most respondents reported that the program exceeded expectations. For respondents who sought to improve their own leadership skills, the program generally met expectations, although the program did not meet expectations for one respondent.

Figure III.2. The Successor Prep program met or exceeded expectations to fulfill participants' frequently cited goals

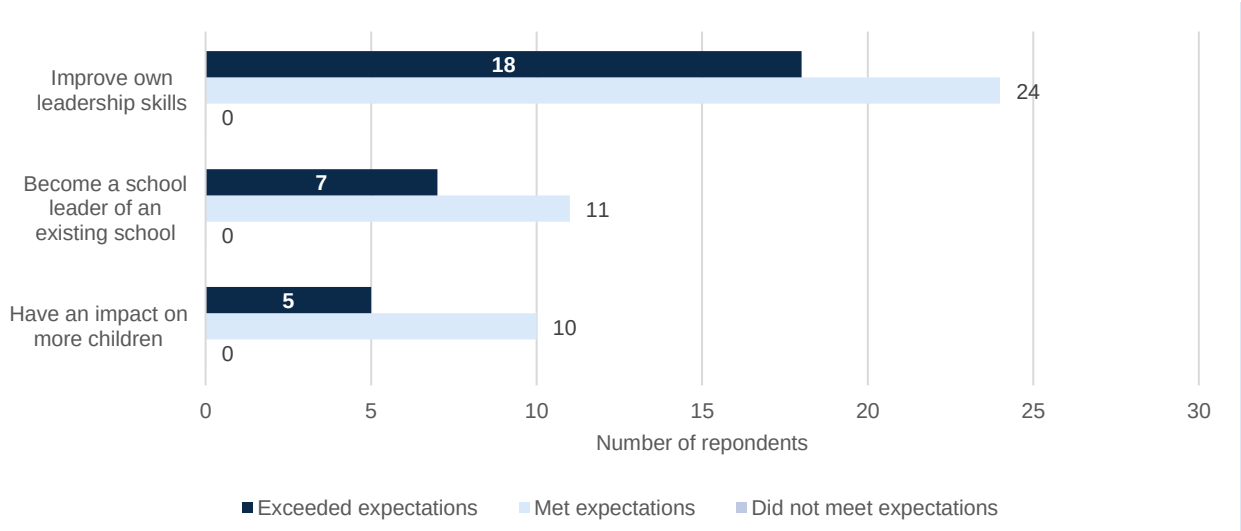
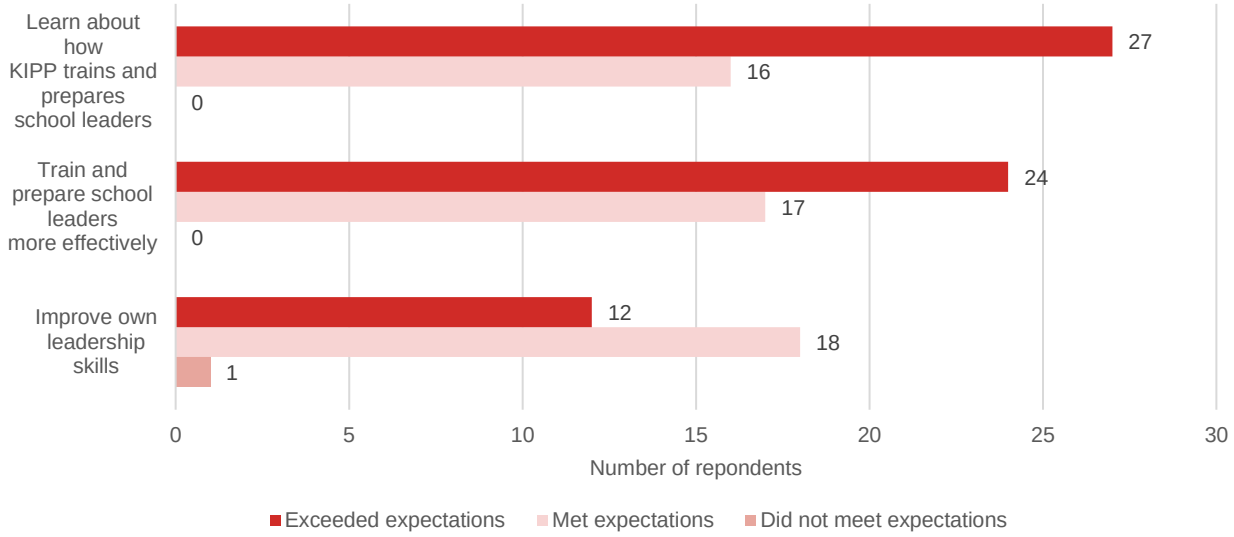


Figure III.3. The KLDF program generally met or exceeded expectations to fulfill participants' frequently cited goals



D. Research Question 2: Which aspects of the two leadership programs do participants find most useful?

Survey respondents were asked to assess the overall quality of their program, as well as how the program addressed various leadership themes and individual program features. To build on their overall perceptions about the program, we asked survey respondents to describe what they felt were the program's strengths and areas of opportunity. We then solicited feedback on how effective participants thought the program was in building each of the skills that were a focus of the programs, such as strategic planning for the Successor Prep program and talent development for the KLDF program. Last, we asked for feedback on specific program components or content to understand which program activities respondents found most and least useful. Much of what we learned came from open-ended responses, which generated a wide range of responses that we then categorized into common themes for ease of interpretation.

In general, respondents were satisfied with the program and believed that the program was effective in building their leadership skills, but they also pointed to some areas for improvement—particularly around enhancing both the depth and breadth of equity leadership training, building more collaboration and follow-up opportunities after program completion, and incorporating more personalized or real-world experiences into the training.

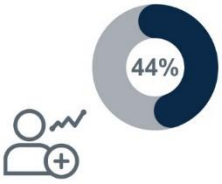
1. Overall program perceptions, strengths, and opportunities for change

Exhibits III.1 through III.4 show that a large majority of participants in both Successor Prep and KLDF programs reported that the leadership programs were relevant to their professional development, that they were satisfied with the program content, and that they would recommend the program to a colleague. Although most of the responses were positive, fewer respondents claimed that the program was essential to their professional development. Figures A.1 and A.2 in Appendix A.2 display the full distribution of responses for each of these measures of satisfaction.

Exhibit III.1

Relevancy to Professional Development

44% of Successor Prep respondents reported the program to be **essential** to their professional development



40% of Successor Prep respondents reported the program to be **very relevant** to their professional development

15% of Successor Prep respondents reported the program to be **somewhat relevant or not at all relevant** to their professional development

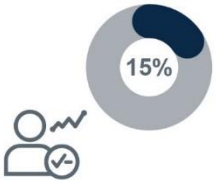
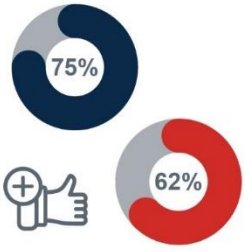


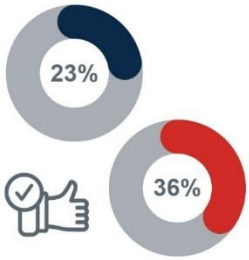
Exhibit III.2

Satisfaction with Program Content

75% of Successor Prep and 62% of KLDF respondents reported being **very satisfied** with program content



23% of Successor Prep and 36% of KLDF respondents reported being **somewhat satisfied** with program content



2% of Successor Prep and 2% of KLDF respondents reported being **somewhat or very dissatisfied** with program content

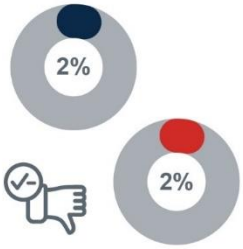
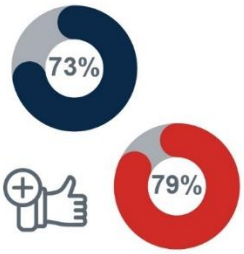


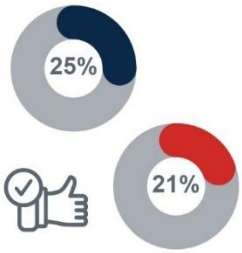
Exhibit III.3

Satisfaction with Delivery of Content

73% of Successor Prep and 79% of KLDF respondents reported being **very satisfied** with the delivery of content



25% of Successor Prep and 21% of KLDF respondents reported being **somewhat satisfied** with the delivery of content



2% of Successor Prep respondents reported being **somewhat dissatisfied** with the delivery of content

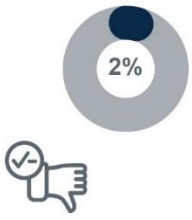
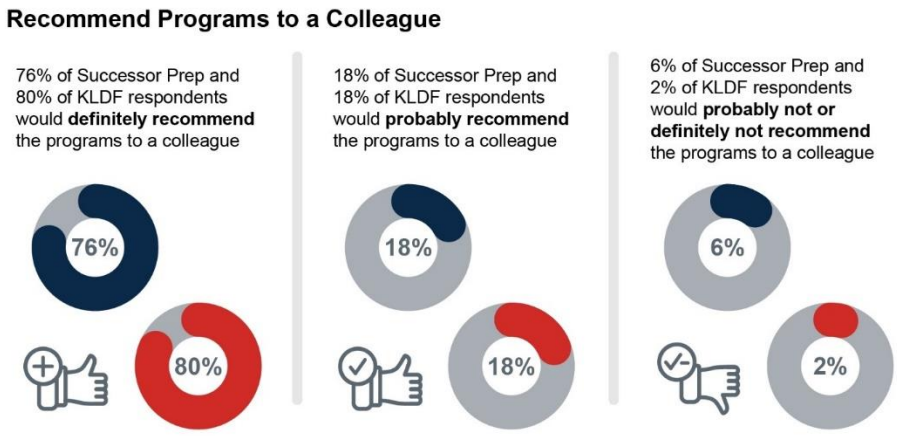


Exhibit III.4



2. Strengths and opportunities for change for the Successor Prep program

a. Successor Prep program strengths

Strengths of the Successor Prep program

- Most respondents highly valued their **cohort experience**, including collaborating and building partnerships with other school leaders.
- A majority also considered the high caliber of **speakers, presenters, and facilitators** to be a top strength of the program.
- About half thought a top strength was the **high-quality mentorship and individualized feedback and attention** they received from their coaches

With input from the KIPP leadership program coordinators, we identified eight primary features of the Successor Prep program to learn more about how participants perceived them. We presented the respondents with these eight features and asked them to select up to three that they felt were the greatest strengths of the program. The number and share of respondents who selected each option are provided in Table III.3 below. Because respondents could select only three options, an option selected by at least 40 percent of respondents represents a widely acknowledged top strength of the program. Over two-thirds of respondents considered the experience of collaborating and communicating with their cohort to be one of the top strengths of the program. Over half of respondents identified the program's strong facilitators and one-on-one coaching as top strengths of the program. Many indicated the program content as a top strength (43 percent), and some thought the program structure was a top strength (24 percent).

Table III.3. Top strengths of the Successor Prep program

Feature	Frequency	Percent of respondents
Cohort experience	31	67%
Strong facilitators	25	54%
One-on-one coaching	23	52%
Program content	17	43%
Program structure	10	24%
Equity training	6	15%
Role-playing sessions	6	13%
Residency experience	4	11%

b. Successor Prep opportunities for change

Our question about top strengths of the Successor Prep program also revealed that respondents generally did not perceive equity training, role-playing sessions, and residency experience to be top strengths of the program (Table III.3). Residencies were discontinued after the start of the COVID-19 pandemic, which may explain the low score we observed for this program feature. To build on these findings, we asked respondents to describe knowledge, skills, or abilities they had hoped to learn but that were not addressed in the Successor Prep program. Twenty-two respondents offered responses, which touched on four common themes of potential opportunities for change. These themes, along with example responses, are summarized in Table III.4.

Respondents most often expressed a desire to have learned more about talent management, including approaches to cultivating and maintaining positive staff culture and morale through investments in emotional supports. For responses that fit into the instructional leadership theme, we included responses that were related to instruction, coaching, and support. The technical skills theme included responses that were related to budget, process monitoring, and accountability systems. Last, the theme of regional leadership included responses related to navigating regional dynamics and stakeholders. Some responses fit into more than one theme and were coded into multiple themes where appropriate. Although respondents generally perceived that there was adequate attention given to equity training, no one described specific equity-related skills or knowledge they had hoped to learn but did not.

Table III.4. Common themes about knowledge, skills, and abilities not addressed in the Successor Prep program

	Frequency	Examples
Knowledge, skills, or abilities not addressed		
Talent management	10	"I was hoping to learn how to effectively engage my leadership team and entire staff in FYSAP throughout the year to maintain investment, make appropriate shifts, and maximize growth." "The post pandemic teaching force needed a different set of skills. I wish there was development on how to support humans in their needs or how to have conversations around emotional intelligence."
Instructional leadership	7	"I felt like the Successor Prep program gave me a strong foundation for stepping into the seat. In retrospect, I wish that I could have learned even more about developing APs on their instructional coaching…" "One skill I would have loved to have more focus [on], or better yet, I have noticed would have made a huge difference in my leadership—the focus on coaching teachers and leaders."
Technical skills	5	"I didn't feel as developed in managing the operational side."
Regional leadership	4	"I think that more focus could have been put on managing up, managing side-ways, up or preparing to navigate regional dynamics."

Suggested Action. Deepen talent management programming

Consider additional programming tailored to developing participants' talent management skills.

This could include providing participants resources with strategies for building positive staff culture.

Consider also structuring opportunities for participants to work collaboratively to design and practice implementing emotional supports for staff.

3. Strengths and opportunities for change for the KLDF program

a. KLDF strengths

In total, 47 KLDF respondents offered their opinions about the program's strengths. Table III.5 shows a summary of the frequency of themes describing program strengths mentioned by respondents, accompanied by an example response. Many responses mentioned multiple distinct strengths and are therefore counted multiple times.

Strengths of the KLDF program

Most respondents valued the strength and rigor of **the leadership framework model** ground in the KLDF program, as well as the opportunity to **network** with others in the field and share best practices and lessons learned.

Most KLDF survey respondents mentioned the models of leadership development covered by the program as a key program strength. Respondents noted that the program curriculum and reference materials were well structured and aligned to the leadership competency framework. Most respondents also highlighted the opportunity to network and collaborate with a diverse cohort of participating organizations as critical to their growth. Respondents valued the opportunity to share best practices with and learn from other participating organizations.

Many respondents highlighted the content knowledge and facilitation skills of the session facilitators as a program strength. Others highlighted the opportunities to think through real-world problems, either through case studies that were presented as part of the program or through opportunities in the course to share out about problems they are currently facing in their district or role. Some respondents also commented that learning from KIPP's experiences was a strength of the program, and others mentioned that the school visits and observing the leader selection process helped them see how the theories they discussed were used in practice. Few respondents considered training on leading for equity to be a strength of the program; many expressed this to be a potential area of improvement for the program, as discussed in detail below.

Table III.5. Common themes about KLDF program strengths

	Frequency	Example
Strengths		
Models of leadership development	27	"I feel the strength[s] are that the fellowship is grounded and aligned to the leader competency framework. The sessions/activities are well organized and facilitated by a good mix of former principals/building leaders and subject matter experts. The sessions/activities are grounded in research."
Networking	27	"Building a strong cohort of people doing this work around the country. Sharing best practices and lessons learned."
Facilitators	15	"The quality and experience level of the facilitators was outstanding."
Feedback on real-world problems	13	"Hands-on approach to learning; real world connections."
KIPP experiences	12	"One of KLDF's biggest strengths was that it provided the opportunity to unpack some of KIPP's strongest leadership development practices—seeing these exemplars continually pushed my thinking in terms of how to approach leadership development."
School visits and experiences	11	"Exposure to actual leaders who know and do this work; school visits and observation of actual training (seeing and participating in the actual KIPP fellowship programs)."
Equity training	4	"Strong equity foundation (new during this KLD Fellowship)—everything seemed to be rooted in equity and dismantling White Supremacy Culture—including a deep reflection on how we, as leaders, are perpetuating—consciously and unconsciously."

b. KLDF Opportunities for change

Respondents were asked to suggest opportunities for change or areas for improvement. Table III.6 summarizes the main themes about improvement, based on the 33 responses we received on this survey item.

Table III.6. Common themes about opportunities for change in KLDF program

	Frequency	Example
Opportunities for change		
More equity content	11	"Deeper dive into the areas of equity and leadership. Excellent but short overview. It could have been more in-depth."
Follow-up opportunities	7	"I know this could require too much capacity from the KLDF team, but I wonder if there's a way to maintain some contact in the year following KLDF to check in on implementation of various practices. For example, a quarterly or bi-annual call could be helpful to implementation back in our orgs."
Resources tailored to needs	7	"More focus on external recruitment of school leaders in the Recruitment & Selection summit." "I believe adding more equity scenarios and strategies to integrate Social Emotional Learning into engagement strategies."
More time and emphasis on collaboration	6	"More time to collaborate with other cohort members. The content where we practiced something or engaged in protocols with each other was more transferrable than the multiple sessions downloading frameworks or KIPP processes and approaches. That information could be shared in prework or asynchronous overviews so that the time together really maximizes on the cohort experience."
Practice in applications	4	"Additional time to work on application of the concepts."
Coaching and mentorship	2	"Intensive coaching with teams with KIPP leaders, facilitators, etc."

Many respondents suggested the program deepen its focus on equity, for example, by discussing equity content in more depth, using more equity-related scenarios, introducing new equity-focused sessions (for example, on power and politics or on interrupting white supremacy culture), or providing examples of organizations implementing equity practices well. Although equity training was brought to the forefront of the program in later years, some respondents from the most recent cohorts also cited a need for more in-depth equity content.

Respondents valued the collaborative nature of the program and suggested further opportunities to extend and deepen the cohort experience. Some respondents requested the program offer follow-up opportunities to strengthen long-term engagement between participating organizations and to monitor and support the implementation of program practices. Respondents suggested long-term mentoring with program staff, a quarterly or biannual call with all participants, or the formation of smaller follow-up learning communities that meet more regularly. Other respondents noted they would have liked more time to collaborate during the program, namely, through more frequent small-group discussions and opportunities for peer-to-peer feedback. Respondents suggested the program could consider decreasing the amount of structured content delivery time to allow for more opportunities to practice and engage with other participants.

In addition, some respondents suggested ways to personalize program content and materials to increase their learning in the program. For example, some respondents requested additional resources, specific to

their organizational contexts; these included social-emotional learning strategies, materials for digital learning, and more resources for academic leaders, external recruitment of school leaders, and technical systems, like dashboards. Others suggested the program build in more time for participants to practice applying program concepts to their organizational contexts. Several respondents mentioned that small-group coaching with KIPP staff would also support their leadership development.

Suggested Action: Explore ways to tailor program materials to participants' organizational contexts

Consider providing participants a curated list of additional reference materials.

Consider surveying participants about their specific organizational needs and using responses to compose small discussion groups with similar interests.

Finally, KLDF respondents were asked in an open-ended question about knowledge, skills, or abilities they were hoping to learn but that were not addressed in the program. Drawing on substantive answers from 26 respondents, we identified five common themes about desired content that was not addressed. These themes, along with sample responses, are listed in Table III.7.

Table III.7. Common themes about knowledge, skills, and abilities not addressed in the KLDF program

	Frequency	Example
Knowledge, skills, and abilities not addressed		
Equity concepts	10	"There was not much content about equity, or an investigation of how white supremacy culture might be at play in KIPP's practices and processes. I believe that's changed now."
Additional support applying the content	6	"I would have liked to have spent more time on the competencies, how they were developed and what they look like in action in different regions."
Gaining decision-maker buy-in	4	"I would like to have more opportunities to discuss and learn how to get the senior leaders invested as much as we are."
Personal leadership development	3	"I wanted to learn more about personal leadership development and goal setting."
Additional time on recruitment	3	"Retention and recruitment were addressed, but I would have appreciated more time on that!"

Respondents most often mentioned that the KLDF program did not address certain equity-related concepts that they were hoping to learn, echoing participants' suggestions about deepening the program's focus on equity in various programming components. In fact, 17 unique respondents mentioned equity programming when suggesting improvements for change or describing skills they had hoped to learn, meaning these feelings are widespread and not concentrated among a few respondents. Respondents suggested the program demonstrate how KIPP is investigating the impacts of white supremacy culture in its practices, how to support schools with their own equity initiatives, and specific practices for disrupting existing inequitable practices and driving systemic change in participants' organizations. Similar to the feedback received from Successor Prep respondents, this input was received from KLDF respondents in the earlier cohorts as well as those in the most recent cohorts who began after KIPP added more equity leadership content to the program curriculum. Therefore, despite the fact that the program has increased the emphasis on equity in recent years, participants believe that these changes have not sufficiently addressed this issue.

Other respondents had hoped for additional support applying program content, noting the program covered a significant amount of valuable content but may not have developed participants' skills in applying the content. Respondents also suggested that the program add programming to develop participants' ability to gain buy-in from key stakeholders and leverage personal leadership competencies such as goal setting.

Suggested Action: Deepen equity programming and prioritize opportunities for application, collaboration

Consider spotlighting additional examples of equity concepts in practice and providing opportunities for participants to practice applying an equity-based leadership lens to relevant scenarios.

Consider asking participants to review some programming content asynchronously and prioritizing time during sessions for reflection, small-group discussion, and application of concepts.

4. Success in building skills for participants of the Successor Prep program

Successor Prep participants were asked to assess their program's successes in three domains, where each domain focuses on the kinds of skills KIPP aims to help participants develop (see definitions in the box at right): managing to achieve positive outcomes in the school, strategic planning, and leading for equity. Survey respondents were asked to rate the program's effectiveness at developing participants' skills in each domain, the amount of time the program spent on each domain, and how important the skills they have learned relate to their success (which is explored further under the third research question, below).

Successor Prep program domain definitions

1. **Managing to achieve positive outcomes in the school:** collecting data for monitoring school performance; monitoring school performance metrics; providing data-driven leadership; providing continuous instructional performance support; engaging in continuous learning
2. **Strategic planning:** providing visionary leadership; providing mission-driven leadership; supporting staff learning and growth; setting and pursuing instructional goals; setting and pursuing operational goals; establishing decision-making processes; planning, executing, and committing
3. **Leading for equity:** providing equity leadership; striving to dismantle systemic inequities; identifying guiding principles for culture systems; setting direction and modeling expectations; encouraging team leadership; encouraging constructive dialogue; demonstrating cultural competence

As shown in Table III.8, most respondents found the program effective at helping them manage to achieve positive outcomes in the school and helping with strategic planning. Respondents thought that the Successor Prep program was somewhat effective at leading for equity (see Tables A.1–A.3 and Figures A.3–A.5 in Appendix A.3 for distribution of responses across all survey items related to each domain). Mean effectiveness scores for managing to achieve positive outcomes in the school and strategic planning suggest that most survey respondents believed the program was either very effective or effective. Although still considered effective on average, the program was rated least effective in helping participants lead for equity.

Box 1. Method for calculating program domain summary measures

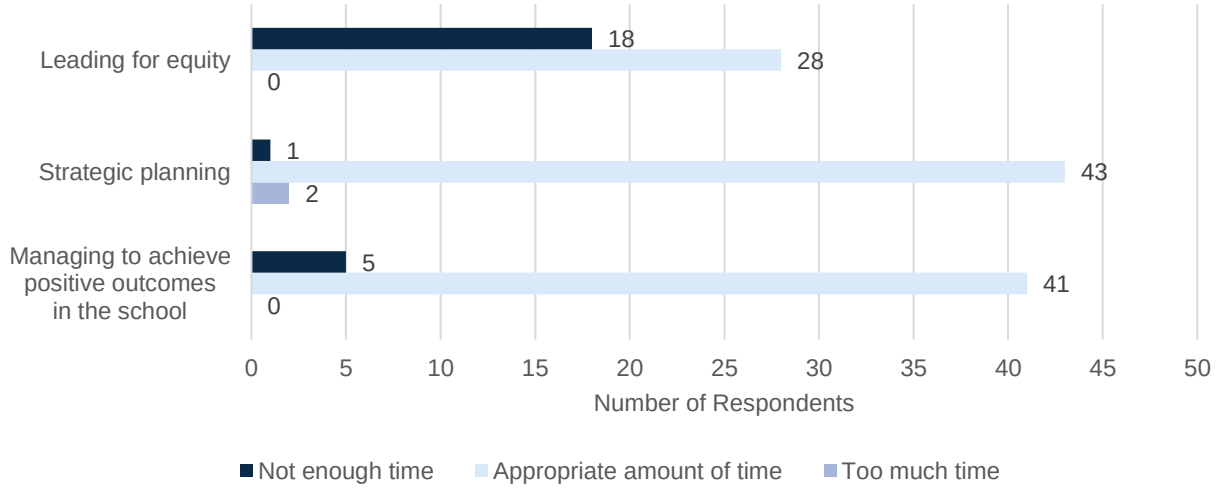
Survey respondents were asked to rate the effectiveness of various program aspects within each program-specific domain (as defined in the callout box on page 25 for Successor Prep and callout box on page 27 for KLDF). Program effectiveness was measured on a scale of 1 to 4, with 1 meaning very effective, 2 meaning effective, 3 meaning somewhat effective, and 4 meaning not at all effective. Because of the large number of components across the domains, we created and analyzed summary measures to simplify the presentation of the results. To do this, we examined correlations in the responses across these components and found that within-domain responses on all domains were highly correlated and consistent. Therefore, we computed a composite for each domain equal to average ratings of individual items. Our discussion focuses on composite scores; however, Tables A.1–A.5 in Appendix A include descriptive statistics of each individual component.

Table III.8. Successor Prep program effectiveness by program domain

	Number of responses	Mean
Managing to achieve positive outcomes in the school	46	1.89
Strategic planning	46	1.92
Leading for equity	46	2.02

Respondents were also asked to rate whether the program spent an appropriate amount of time developing skills in each domain. Figure III.4 shows that most respondents believed the program spent an appropriate amount of time on each program objective. No respondents thought the program allocated too much time developing skills around managing to achieve positive outcomes, although a small share believed the program did not spend enough time on this skill set. Few respondents thought the program spent too much time on developing strategic planning skills, but a small share also thought the program did not allocate enough time to this skill set. No respondent thought the program spent too much time on developing skills around leading for equity, but many respondents noted a desire for more time devoted to leading for equity. When asked about specific skills related to this objective that they would have liked to spend more time on, this latter group indicated a desire for more time building skills around providing equity leadership, dismantling systemic inequities, identifying guiding principles for culture systems, and encouraging constructive dialogue.

Figure III.4. Successor Prep participants' ratings of program time use by domain



Suggested Action: Spend more time building equity leadership skills

Consider ways to shift time spent on the strategic planning domain to the leading for equity domain, including concepts related to dismantling systemic inequities, identifying guiding principles for culture systems, and encouraging constructive dialogue.

Consider ways to incorporate equity leadership concepts throughout programming in the other two domains—for example, by pausing to reflect on strategic planning content from an equity lens or prompting equity-related reflections during group discussion of monitoring school performance data.

5. Success in building talent development skills of others for participants of the KLDF program

KLDF participants were asked to assess their program's successes in two domains (definitions are in the box at left): talent development and leading for equity. Survey respondents were asked to rate the program's effectiveness at building participants' capacity to develop the skills of potential or current leaders in each domain, the amount of time the program spent on each domain, and how often they personally focus on building the skills of potential or current leaders in their everyday practice (which is explored further under the third research question, below).

KLDF program domain definitions

1. Talent development: recruiting and selecting highly effective staff; providing on-the-job development; retaining talented and valued employees; managing staff who do not meet expectations or fit the organizational culture; providing coaching and one-on-one support to staff; establishing standards for effective leadership; developing leaders and preparing successors
2. Leading for equity: providing equity leadership; striving to dismantle systemic inequities; identifying equitable organizational policies; demonstrating cultural competence

For both of the main skill domains, survey respondents were asked to rate the program's effectiveness on each of four to seven specific components (listed in the box). Because of the large number of components across the domains, we created and analyzed summary measures to simplify the presentation of the results, similar to the approach used for the Successor Prep survey described previously in Box 1 on page 25.

Table III.9 presents the mean and number of observations for program effectiveness across both skill domains. On average, participants found the program moderately effective at building their capacity to develop talent development and leading for equity skills in potential or current leaders.

Table III.9. KLDF program effectiveness across program domains

	Number of responses	Mean
Talent development	52	2.20
Leading for equity	47	2.34

Overall, more respondents believed the program spent an appropriate amount of time on skills in the talent development domain than on skills in the leading for equity domain. Figure III.5 shows that most respondents believed the program spent an appropriate amount of time on most skills in the talent development domain. However, respondents believed not enough time was spent on managing staff who do not meet expectations. More respondents indicated that more time should have been spent on building this skill than on all other skills addressed across both domains. Relatedly, some respondents also believed not enough time was spent on retaining employees and providing coaching and one-on-one support. Figure III.6 shows that most respondents believed the program spent an appropriate amount of time on two skills in the leading for equity domain: providing equity leadership and demonstrating cultural competence. However, respondents believed not enough time was spent on identifying equitable organizational practices and striving to dismantle systemic inequities. These findings align with respondents' suggestions to deepen the program's equity focus, including suggestions to provide more examples of equitable systems, opportunities to see equity practices in action, and content on changing existing inequitable systems. Together, these findings reveal that although participants want more time devoted to several key topics, they do not want the program to be cut back elsewhere suggesting the available time for the program may be insufficient.

In all, these findings suggest that KLDF respondents, on average, desired more time invested in key topics than Successor Prep program respondents. However, such comparisons of findings across programs should be interpreted with caution because participants' needs and expectations were different. For instance, KDLF participants are typically already in a role that is the focus of the training, so it may be the case that it is more difficult to help them develop further skills than it is for the Successor Prep participants, most of whom are being trained for a role they have never held.

Figure III.5. KLDF participants' ratings of program time use: Talent development domain skills

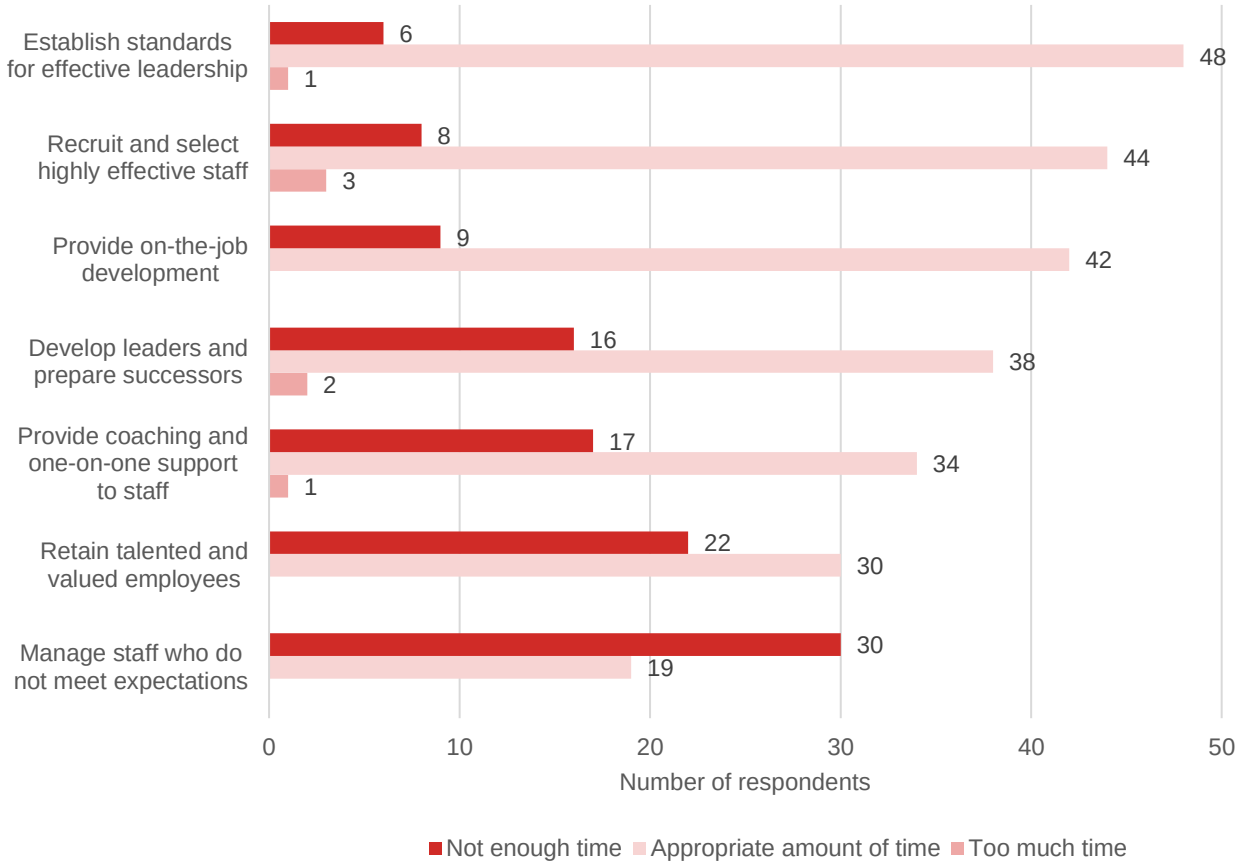
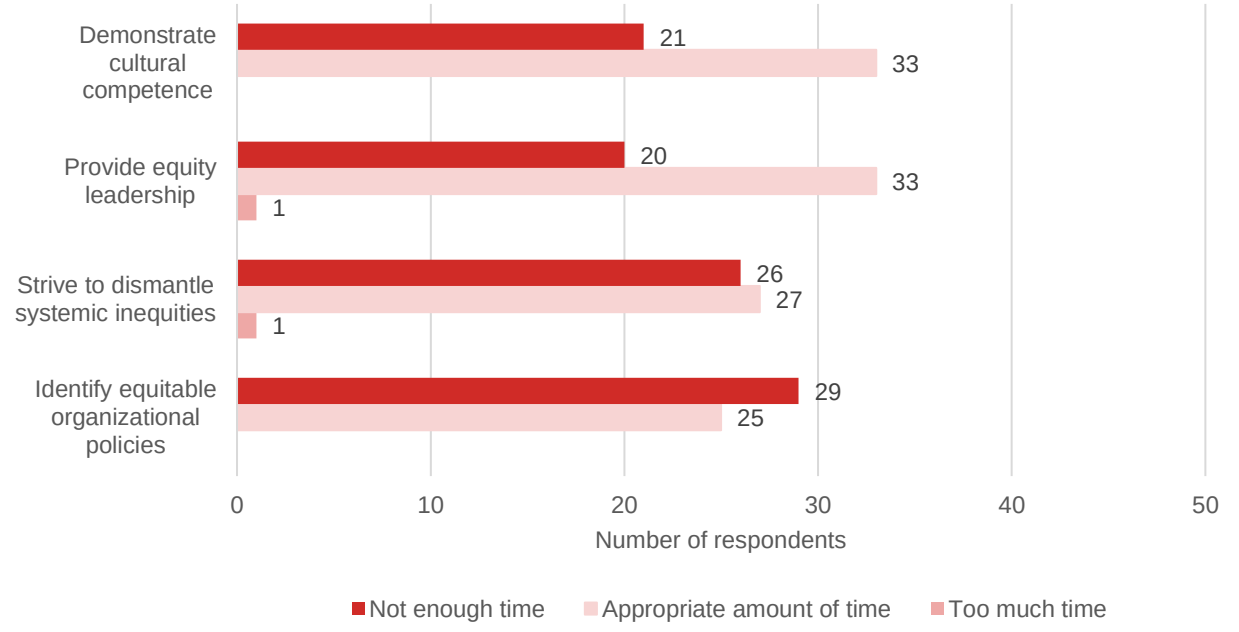


Figure III.6. KLDF participants' ratings of program time use: Leading for equity domain skills



Suggested Action: Restructure opportunities to practice managing staff and weave equity concepts throughout domains

Consider incorporating more explicit opportunities for staff to practice managing staff who do not meet expectations. Practice of this skill could be included during time already spent on providing coaching and one-on-one support.

Consider ways to incorporate equity leadership concepts—namely, identifying equitable organizational practices and striving to dismantle systemic inequities—throughout other programming.

This could include pausing to reflect on development of leadership standards from an equity lens or prompting equity-related reflection during discussions of recruiting and supporting staff.

6. Perceptions of program components and content

Understanding what program activities respondents found useful can help KIPP determine where to invest limited program time and resources. Survey respondents for both leadership programs were asked to assess the value of specific program components or activities (Successor Prep) or content (KLDF) on a scale of 1 (very beneficial) to 4 (not at all beneficial).

Primary components of the Successor Prep program

1. **Orientation** (a multi-day event to build key relationships with teammates and begin critical coursework)
2. **Professional development/intersessions** (multi-day intensive coursework or sessions to develop skills in various leadership areas)
3. **Coaching** (working one on one with a leadership coach)
4. **Summer institute** (multi-week intensive with rigorous coursework)
5. **Residencies** (multiple days working in a school inside or outside the participant's region)

a. Perceptions of Successor Prep program components

As shown in Table III.10, when applicable, Successor Prep survey respondents found most program components beneficial to their professional growth. The mean score for all components was below 2. Successor Prep respondents rated the professional development and coaching components as most beneficial; participants who completed residencies also rated these experiences as beneficial.

Table III.10. Participant views of usefulness of Successor Prep program components

	Number of responses	Mean
How beneficial was each component to your professional growth? (1 = very beneficial, 2 = beneficial, 3 = somewhat beneficial, 4 = not at all beneficial)		
Orientation	42	1.8
Professional development/intersessions	45	1.6
Coaching	44	1.7
Summer institute	39	1.9
Residencies	16	1.7

Respondents who perceived a content component as somewhat beneficial or not at all beneficial were asked to describe how the component could be improved to be more beneficial for their professional growth. As noted above, each component of the Successor Prep program was generally well received, so there were a limited number of suggested improvements. Although we received limited responses, the feedback discussed could indicate there is some room for improvement or change. We received responses

for the program components of orientation, professional development or intersessions, coaching, and summer institute, but we did not receive any for residencies.

Orientation program component feedback (four responses). We identified two common themes: improvement of the format of orientation delivery and relationships or team building. Although some may appreciate the flexibility that comes with virtual learning, some respondents left the experience feeling disconnected. For instance, two respondents noted that they completed the orientation virtually, with one of them writing that the virtual format "took away from the overall experience." Another respondent wrote that they would have benefited more if there were "more social and authentic relationship building … less sit and get" during orientation.

Similar to the orientation component, the area of improving relationships and team building was one of the themes, along with content, specifically related to the desire for more programming on monitoring goals and responding to leading indicators. For the relationship and team building improvements, one respondent wrote, "I would have loved to see more team building to bring life lines"; another respondent noted they had virtual sessions, which made it difficult to build relationships.

Professional development/intersessions component feedback (three responses). One respondent wanted more time devoted to learning to monitor goals and respond to leading indicators. The other two respondents suggested the sessions would have been more beneficial if they placed more focus on relationship building. One respondent attributed their concern to the virtual format of the sessions: "Our sessions were virtual and many people did not attend or were in and out of sessions. It made it hard to feel the sessions were important and build relationships."

Coaching program component feedback (four responses). These respondents all felt they did not get a good coaching match. One suggestion was to "ensure that the leadership coaches are matched with the SL [school leader] based on background and experiences."

Summer institute component feedback (five responses). The themes found in these responses addressed the format of the summer session delivery and the content. The response related to the format of summer sessions was, "It's just too hard to be away from families and life for that long. I think virtual programming or smaller week institutes are more effective." Some responses related to content included lack of coverage on how to monitor goals and respond to leading indicators; lack of preparation for the upcoming school year; and too much time spent on task completion and instruction and not enough focus on culture.

Suggested Action: Explore ways to deepen relationship building during virtual sessions

Consider incorporating more extended, smaller group activities during orientation and professional development sessions to support participant relationship building.

Consider gathering information on participants' background and experience to structure small-group activities and leadership coaching pairs that build meaningful connections.

b. Perceptions of KLDF program content

We asked KLDF respondents to rate how beneficial content areas that were delivered through the program—rather than specific components—were to their professional growth. The content of the KLDF program can be grouped into seven categories, listed in the box to the right. Table III.11 shows that respondents indicated the opportunities to see practice in action (including school visits, principal selection, and observing cohort learning) as most beneficial. Programming around establishing leadership standards (by exploring KIPP's practices and developing leadership competencies) was also rated highly, as well as content focusing on developing participants' skills in building leadership pipelines. These ratings align with responses indicating a key strength of the program was the models of leadership development featured during sessions and observed during interactive experiences. KLDF respondents rated the leading for racial equity, selecting for equity and excellence, and learning team components of the program as least beneficial.

Content of the KLDF program

1. Seeing practice in action (school visits, principal selection, observing cohort learning)
2. Leading for racial equity
3. Establishing leadership standards (exploring KIPP's theory of leadership development, principal leadership competencies, and school leader progression)
4. Building equitable and excellent leadership pipelines (exploring effective practices in creating an equitable and excellent leadership pipeline)
5. Selecting for equity and excellence (exploring equitable hiring practices and KIPP's selection event design)
6. Developing for equity and excellence (exploring and practicing effective leadership coaching, adult learning practices, and formal programming components)
7. Learning teams (convening and collaborating with members of your fellowship cohort between summits)

Table III.11. Participant views of usefulness of KLDF program content

	Number of responses	Mean
How beneficial was the following content to your professional growth? (1 = very beneficial, 2 = beneficial, 3 = somewhat beneficial, and 4 = not at all beneficial)		
Learning teams	52	2.2
Selecting for equity and excellence	50	2.1
Leading for racial equity	44	2.0
Developing for equity and excellence	49	1.9
Building equitable and excellent leadership pipelines	53	1.8
Establishing leadership standards	54	1.7
Seeing practice in action	54	1.5

In line with respondents' suggestions to provide more equity-based content, some KLDF respondents thought the existing equity-related programming components could be improved. Thirteen respondents suggested improvements to KLDF programming concerning selecting for equity and excellence (defined as exploring equitable hiring practices and KIPP's selection event design). Most of these respondents suggested the program demonstrate how to embed equity at all stages of the selection process, including the design of selection systems and how equity is cultivated before and after the candidate interview. Eight respondents suggested expanding programming related to leading for racial equity, and another three respondents suggested deepening the equity focus of programming tailored to developing leaders for

equity and excellence. In addition to requesting more time on these components, several respondents suggested more opportunities to see these frameworks in practice, for example by observing equity-focused discussions or practices in action at a school site, seeing examples of equity embedded into day-to-day school systems, or experiencing equity coaching directly from KIPP program staff. To cultivate productive dialogue on these topics, a few respondents also noted they would have appreciated facilitation techniques that were more open-minded and responsive to participants' varied viewpoints.

In addition, 11 respondents suggested improvements to the structure of the learning teams: several respondents emphasized the lack of engagement from learning team members and noted the group members' varying organizational contexts was a barrier to meaningful collaboration. Respondents suggested the program provide more structure to the learning teams, for example by providing discussion guides, experimenting with role-alike and smaller groupings, and/or incorporating more time for learning teams to build rapport during the summit.

Suggested Action: Broaden examples of equity in practice and provide more structure to learning teams

Consider broadening selections for equity content to include examples of equity during the design of selection systems and after the candidates are interviewed.

Consider modeling equity coaching practices and spotlighting equity frameworks during site visits.

Consider restructuring learning teams to increase engagement and collaboration. This could include providing learning team discussion guides, experimenting with role-alike and smaller groupings, and incorporating more time for learning teams to build rapport during summits.

c. Desired KLDF program features

To identify program features that KLDF participants would have found most valuable had they been offered, we asked respondents to rank seven potential program offerings from 1 (most valuable) to 7 (least valuable). Table III.12 shows that, on average, respondents ranked the following potential program offerings as most valuable: additional one-on-one or small-group coaching with KIPP Leadership Design Fellowship facilitators (average ranking of 3.10) and additional equity training (average ranking of 3.54). Forty-two percent of respondents ranked these components as the most or second-most valuable potential program offerings. In open-ended response items, respondents named a similar desire for greater personalization of programming and a deeper equity focus. Respondents ranked additional sessions for practicing providing feedback or having difficult conversations with potential or current school leaders (average ranking of 4.69) and a KLDF alumni convening (average ranking of 4.38) as least valuable. That respondents ranked an alumni convening as less valuable is slightly surprising, given that open-ended responses suggested additional follow-up opportunities be rolled out for program participants.

Table III.12. Average value of potential program offerings through KLDF

Potential offerings	Average value rating (from 1 = most valuable to 7 = least valuable)
Additional one-on-one or small-group coaching with KLDF facilitators	3.10
Additional equity training	3.54
A session exploring rapid onboarding of school leaders in trying or unusual circumstances	3.96
Additional direct feedback on your prototype design from KLDF facilitators	4.04
Additional time to collaborate with learning team during summits	4.17
KLDF alumni convening	4.38
Additional sessions for practicing providing feedback or having difficult conversations with potential or current school leaders	4.69

Suggested Action: Pilot follow-up opportunities and additional opt-in coaching

Consider piloting less time-intensive or opt-in follow-up opportunities to better understand participants' desire for follow-up supports, and what kind of additional opportunities would be most useful.

Consider piloting an opt-in opportunity for additional small-group coaching with facilitators during or after the program.

E. Research Question 3: Are participants in the two leadership programs able to apply the lessons of the program in their own settings?

School leaders often juggle several job responsibilities at once, and these responsibilities may vary by school and by the day. Given these complex realities, we surveyed program participants to assess how much time participants spend on job responsibilities that were emphasized in the program or how important these responsibilities, or skills, were to their success. We then asked respondents to assess how influential the programs have been in their leadership approach and to describe their successes and challenges in implementing program practices. The first assessment provides a sense of whether the leadership programs are targeting topics and areas that are common across participants. The second assessment addresses how well the programs prepare participants to put these practices into action.

1. Relevance of program focus

a. Relevance of the Successor Prep program

To assess the relevance of KIPP's training topics for Successor Prep, we asked respondents to rate how important each of the program domain skills (as defined in the callout box on page 27) are to their success in their current job. Respondents rated each skill on a 4-point scale (1 = very important, 2 = important, 3 = somewhat important, and 4 = not at all important). Table III.13 shows that respondents considered each domain of skills important to their success. Moreover, individual skills under each domain were each rated very important or important, on average (see Appendix Figures A.6–A.8 for skill-specific ratings under each domain).

Table III.13. Importance of skills to Successor Prep participants, by program domain

	Number of responses	Mean
In your current job, how important are each of the following skills to your success? [1 = very important, 2 = important, 3 = somewhat important, and 4 = not at all important]		
Managing to achieve positive outcomes in the school	46	1.3
Strategic planning	46	1.3
Leading for equity	46	1.3

b. Relevance of the KLDF program

We measured the relevance of KIPP's training topics for KLDF participants by asking KLDF respondents to rate how often they focus on building skills from each program domain in potential or current leaders. Table III.14 shows that KLDF respondents often focus on both building talent development and leading for equity skills in potential or current leaders in their current roles (Tables A.4 and A.5 in Appendix A.4 presents the average rating for individual skills within each domain).

Table III.14. Frequency with which KLDF participants focus on building skills in leaders, by program domain

	Number of responses	Mean
Currently, how often do you focus on building the following skills of potential or current leaders? [1 = very often, 2 = often, 3 = sometimes, and 4 = never]		
Talent development	55	2.1
Leading for equity	55	2.1

2. Influence of the programs

Although KIPP programs focus on topics that are consistently top of mind for many school leaders, the previous survey questions do little to inform whether participants have incorporated what they learned from the programs into their jobs. To that end, we asked participants to rate the extent to which participating in the program helped increase their ability to perform essential responsibilities.

As a result of the Successor Prep program, participants …

are more **school mission focused**, more **frequently use data** to inform decisions and reach milestones, and **anticipate instructional roadblocks** or setbacks and plan accordingly.

The program least influenced Successor Prep participants' …

ability to provide clearer guidance to teachers, as well as their understanding of systemic racism and ways to establish more equitable teacher practices and learning environments.

a. Influence of Successor Prep program

Successor Prep participants were asked to rate the degree to which participating in the program helped increase their overall ability to support teacher practice and support improvements in student outcomes. Table III.15 shows that, on average, survey respondents believed that the program had moderately increased their ability to support teaching practice and improvements in student outcomes. Responses to these questions were mixed, as shown in Appendix Figure A.9, indicating that while most respondents answered favorably, many did not feel that the program greatly influenced their ability to carry out these key responsibilities.

Table III.15. Influence of Successor Prep program on participants' approach

	Number of responses	Mean
To what degree has participating in the Successor Prep program increased your ability to … [1= extensively, 2 = a good deal, 3 = somewhat, and 4 = not at all]		
Support teacher practice	45	2.3
Support improvements in student outcomes	45	2.2

Next, we aim to understand the specific ways in which participating in Successor Prep influenced participants' leadership abilities. Respondents who indicated the program at least somewhat increased their ability to support teacher practice or improvements in student outcomes were then asked to indicate up to three specific changes they made in their leadership approaches as a result of the program. As shown in Table III.16, the Successor Prep program most commonly influenced successors in becoming more school mission-focused (51 percent). Some (36 percent) also reported that they more frequently use data on student and teacher performance. The same share also feel better equipped to anticipate and develop plans to address instructional roadblocks. Thirty-one percent indicated they are more strategic in professional development planning or are generally more aware and competent in their roles. About one-quarter changed their hiring approach, and one-fifth collaborate with other school leaders as a result of the program. To a lesser extent, participants provide more effective constructive feedback to teachers and clearer guidance on monitoring student outcomes and identifying supports. A small share also gained a stronger understanding of systemic racism and established more equitable teaching practices and learning environments as a result of the program.

Table III.16. Behaviors influenced by the Successor Prep program to support teacher practice or improve student outcomes

	Number of responses	Percentage of respondents
As a result of the program … [participants were asked to select up to 3]		
I am more school mission focused, meaning I am aware of the importance of a school vision and how to create and maintain one.	45	51%
I more frequently use data on student and teacher performance to inform my decisions and reach school milestones.	45	36%
I anticipate instructional roadblocks or setbacks and plan accordingly.	45	36%
I plan instructional professional development sessions more strategically.	45	31%
I am generally more aware and competent in my role in supporting teacher practice and improving student outcomes.	45	31%
I have changed my hiring approach.	45	24%
I collaborate more with other school leaders on instructional leadership practices.	45	20%
I provide more effective constructive feedback to teachers	45	18%
I provide clearer guidance to teachers on how to monitor student outcomes and identify when and where students need additional support.	45	13%
I have a stronger understanding of systemic racism and have established more equitable teacher practices and learning environments.	45	13%

As demonstrated earlier, respondents value their cohort experience, including talking and collaborating with other participants about responsibilities related to the successor role. To understand how successors maintain connections with their cohorts after the program, we asked them how often they talk or collaborate with their cohort and for additional context around topics they discuss. Most respondents (48 percent) indicated they talk or collaborate with other cohort participants less than once a month and 20 percent do so one to two times per month. smaller share (16 percent) of respondents talk to other school leaders in their cohorts once per week, while the remaining 16 percent indicated they never talk or collaborate with their cohort.

For respondents who indicated they talk or collaborate with other cohort participants at all, we asked for up to three examples of when these conversations were helpful in solving a problem in their role as a school leader. From the 25 responses to this question, we identified several themes: Their cohort was helpful when it came to sharing resources, strategic planning, approaches to handling staff issues and staff development, hiring and retention strategies, staff discipline approaches, instructional support strategies, improving and cultivating school culture, and vision setting.

In addition to ways collaboration helped solved problems, many responses cited general benefits of collaborating and having "lifelines" available to them that seemed noteworthy to summarize here. For instance, some respondents said having lifelines was especially important during the pandemic. Others wrote that having the friendships from their cohort gave them moral support and they often shared uplifting messages.

b. Influence of KLDF program

KLDF participants are typically responsible for identifying leadership competencies for their organization, developing an internal pipeline of leaders, establishing equitable hiring and selection practices, and, more broadly, developing leaders. When asked how much participating in the KLDF program helped them carry out these responsibilities, respondents, on average, felt that the program influenced their overall approach to developing leaders the most (see Table III.17). The program also had a good deal of influence on their ability to identify leadership competencies and more of a moderate influence on developing pipelines and equitable hiring and selection practices.

Table III.17. Influence of KLDF program on participants' approach

	Number of responses	Mean
To what degree has participating in the KIPP Leadership Design Fellowship … [1 = extensively, 2 = a good deal, 3 = somewhat, and 4 = not at all]		
Helped identify leadership competencies for their organization	48	2.1
Helped develop an internal pipeline of leaders	44	2.4
Helped establish equitable hiring and selection practices	44	2.3
Influenced approach to developing school leaders	51	1.9

Over two-thirds of KLDF respondents also reported that as a result of participating in the program, they have changed their approach to developing leaders a good amount or extensively. Many went on to note their organization now uses clearly defined leadership competencies to drive all phases of their leadership pipeline, from selection to coaching to evaluation. Other respondents described how the program helped them think through all aspects of the leadership pipeline and develop systems that had never before existed. Echoing responses about key program strengths, they cited the impact (on their leadership skills and organizations) of being exposed to strong models of leadership development throughout the program.

"It pushed us to establish a clear set of competencies that were transparent, and everyone held sacred as a lever for leadership at our organization. Before KLDF, our competencies were not codified, and our performance evaluations were subjective and vague. KLDF helped us define leadership competencies and how to thread them into our day to day being as an organization."

The remaining share of respondents (31 percent) indicated they made little to no change to their approaches as a result of the program. Some respondents explained that the leadership development models discussed in the program already aligned with existing systems at their organizations, making the programming less impactful. Several others said they ran into difficulty applying program learnings in their organizational contexts or that they were not currently in a role where they had the opportunity to apply the content from the program.

One approach the KLDF program takes to increase influence on participants' leadership practices is to help participants identify a problem of practice in their organization and support them in developing a plan to address this problem. The most common problems of practice identified by our respondents were related to designing effective leadership pipelines, developing better-aligned leadership competencies, and implementing equitable leadership and discipline models. Many respondents (62 percent) left the program with a prototype that addressed their problem of practice, and many of those respondents continue to use the prototype, or some modified version, in their current role. Still, several respondents

reported not leaving with a prototype or leaving with one but not using it beyond the program, suggesting that this exercise might be more influential for some participants than others.

Suggested Action: Improve the usefulness and relevance of the problem-of-practice prototypes

Consider ways to ensure all participants leave the program with an actionable prototype on a relevant problem of practice, such as providing additional coaching support or time for developing prototypes.

3. Success in implementation

Participants from both leadership programs were asked to describe their successes in implementing program practices. Successor Prep participants were asked to focus on practices that support improvements in teacher practice and student outcomes, and KLDF participants were asked to focus on practices related to developing school leader pipelines. Responses were grouped into common themes and summarized below.

a. Successes among Successor Prep participants

When asked to describe up to three examples of successes in implementing program practices, Successor Prep respondents overwhelmingly cited successes with assembling and executing strategic action plans at their schools (25 out of 31 respondents, Table III.18). Specifically, they more effectively implemented systems to collect and monitor data to drive schoolwide and instructional decisions as a result of the program. Respondents noted that strategic planning has improved their school each year, has set them up for success, and has helped monitor and respond to progress toward goals.

Table III.18. Common successes in implementing Successor Prep program practices

Theme	Frequency	Examples
Strategic action planning	25	"I actively engaged our leadership team in analyzing SY20-21 data in order to co-create a strategic action plan with clear big rocks, goals, and strategies for SY21-22, which has allowed our team to remain vision aligned and mission driven in times of great flux!" "I have learned how to establish and progress monitor a strategic plan based on data. Staying focused on the one thing has allowed us to get better in that area to drive results for kids."
Talent development	15	"My roles and responsibilities documents that I created during SP helped me to onboard my leadership team in a way that set clear expectations for them to support me in our vision." "I have been able to coach my leadership team in a stronger way, which has increased my leadership bench depth and the stability of the school."
Creating and supporting school vision	14	"Creating a vision for my school which ultimately became our identity for who we are." "Starting with the why session has really helped me anchor all the work we do to our vision. In addition, my coaches have created a vision for their departments that they ground the work that the do…"
Addressing cultural competency, equity, and DEI	12	"DEI work with Successor Prep translated so well to how we started to tackle some inequities in our school." "As a result of the successor programming, my teacher, leader and staff became more diverse in thought, skill, age, race, gender, and identity."
Instructional support	7	"In Successor Prep, we learned about facilitating AP O3s that were grounded in student outcomes and created space to discuss teachers' development. I still follow that O3 process with my APs." "I have a balanced coaching approach which incorporates observation feedback along with checking in emotionally with teachers."

About half of the responses indicated becoming more effective and intentional in developing talent, including selecting, onboarding, and training staff. Many successors also attributed to the program their

successes in creating and supporting their school vision, further citing that having a vision and creating areas of accountability for their school or team, known within KIPP as "Big Rocks", were critical for measuring and tracking their school's priorities and progress. Others were successful in promoting cultural competency and developing various equitable practices, noting the program helped them tackle inequalities in their school and sharpened their vision of equity. Few cited successes in providing instructional support, including more effective classroom observations and follow-up opportunities.

b. Successes among KLDF participants

KLDF respondents were also asked to reflect on their implementation successes. Among 35 respondents who provided a response, many described successes developing their leadership pipeline systems. Table III.19 summarizes the main implementation success themes, pulled from survey responses.

Table III.19. Common themes of successes among KLDF participants

Theme	Frequency	Example
Leadership pipeline systems	18	"As an organization we have been able to support organizations in developing systems and practices (PD/JD/etc.) to support both internal and external pipelines. Additionally, within our organizations we have been able to support our own teacher fellows with establishing a pathway to leadership."
Leadership competencies	12	"Codifying our competencies was a game changer for us. It helped ensure we had the right skills for the roles versus a hope and see mentality."
Hiring and selection practices	11	"We implemented a new recruitment website, interview questions, performance tasks, leadership readiness criteria, and leadership competencies.
Professional development delivery	9	"I still use the PD models from the second session. I frequently use the various participant engagement strategies we were introduced to in the multiple sessions (playing card), Switch book."
Coaching practices	5	"Implemented a leadership competency framework that guides all of our coaching of current and potential leaders."
Equity practices	5	"We have completely changed the equity lens through which we select, develop, and prepare school leaders."

Many respondents identified their ability to improve their organization's leadership pipelines as an implementation success. They described formalizing and establishing better alignment between pipeline components, using models shared during the program as a guide. Many also mentioned their ability to strengthen their organization's leadership competency framework and use the framework to drive each segment of the leadership pipeline. Some respondents highlighted specific successes improving their hiring and selection and coaching practices to better serve internal pipelines. These included improvements to candidate interview and evaluation practices and the intentionality of leader coaching. These successes align with respondents' reports that the program had the most influence on their ability to develop competency-aligned leadership pipelines.

In addition, some respondents highlighted their success applying professional development and other engagement techniques modeled during the program. A few respondents highlighted successes applying an equity lens to their leadership development practices. Despite many respondents highlighting the KLDF program's equity components as an area for growth, these responses suggest that the existing

programming is supporting some successful advances toward more equitable systems in participating organizations.

4. Challenges in implementation

Leadership program participants were also asked to detail their challenges in implementing program practices. Similar to the previous section of implementation successes, Successor Prep participants were asked to focus on practices that support improvements in teacher practice and student outcomes, and KLDF participants were asked to focus on practices related to developing school leader pipelines.

a. Challenges in implementing Successor Prep program practices

We received 28 open-ended responses to the survey question about implementation challenges and two responses that stated that they have not faced challenges implementing practices to support improvements in teacher practice and student outcomes. The remaining respondents all expressed challenges in various areas within six different themes included in Table III.20, which shows these common themes, the frequency of each theme, and example responses that were coded for each theme.

Table III.20. Common themes of challenges among Successor Prep participants

Theme	Frequency	Examples
Talent development	11	"The biggest challenge was developing PD sessions for teachers at all levels of learning. This continues to be a push for our school." "It has been challenging to develop other leaders when those leaders were not people you chose or people on your bench."
Competing priorities and resource constraints	8	"Maintaining focus on my Big Rock while being reactive to the day to day challenges (the whirlwind)." "I've struggled successfully implementing 70-20-10s during the pandemic. I feel like I have not gotten to develop my APs due to the constant 'fire fighting.'"
Regional-related challenges	6	"KIPP foundation was way ahead of our region in regards to equity. Some plans were not plausible because the region wasn't 'ready' for it." "Attending performance management sessions was not always super productive given that our district has very structured systems."
Implementing systems and monitoring data	5	"I found it challenging to develop strong systems to monitor academic progress."
Continued training and support	3	"We were given a significant amount of resources with little support on implementing them and no follow-up."
Equity	3	"I would have wanted to get more from the Equity project. I think in theory the idea of the project was great. But there was not much guidance around it."

Suggested Action: Provide reflection opportunities, reference materials to support implementation

Consider embedding discussion questions throughout group activities that prompt participants to reflect on how implementing key learnings may be a challenge and to collaborate on solutions.

Consider curating a library of additional resources for participants to leverage as they implement program concepts. These could include professional development templates and data monitoring procedures.

b. Challenges in implementing KLDF program practices

When KLDF survey respondents were asked to describe their challenges in implementing program practices, the general theme was disjointed organizational priorities, values, and staffing. Table III.21 summarizes the main implementation challenge themes, the frequency of each theme, and a sample quote that was coded for each theme.

Table III.21. Common themes of challenges among KLDF participants

Theme	Frequency	Example
Competing and shifting priorities	11	"Being in such a large and ever-changing district, it has been challenging to create full coherency and consistency with our supports to pipeline development throughout the city."
Organizational culture	8	"We do not have the infrastructure, really we do not have the mindsets necessary to make the shifts in people's thinking and behavior—that takes so much time to cultivate."
Talent development challenges	8	"We don't have the capacity, time, or number of people with the full knowledge of everything we learned to implement what we learned to the degree that we would like."
Need more resources and support from the Fellowship	6	"I would have appreciated a space to connect with cohort members again and/or receive additional learning after the fact." "I would have loved more examples and PD resources to be used that we could turnkey."
Not enough access or involvement	5	"Our organization is one that influences schools and their leaders but does not have any evaluative authority over them, therefore, some tools are not within our 'lane' to implement."
Not enough buy-in	4	"Our challenges are that our most influential school and network leaders prefer a more subjective basis for selecting and promoting talent."

Respondents cited competing organizational priorities, often resulting in limited time and resources for changes in leadership systems, as barriers to implementing KLDF program practices. They also described the challenges of implementing new systems when their organizational culture lacks an aligned vision and may be inhospitable to change. Similarly, respondents described how frequent turnover in staffing, along with the associated time needed to get new staff on board with KLDF practices, makes it challenging to find the capacity and momentum to make meaningful improvements. One respondent wished their organization could have sent more than one pair of participants to help address this challenge.

Respondents also cited the challenges they faced in gaining enough buy-in from other important stakeholders, including government and network leaders. Others noted they were simply not in positions that allowed them access to day-to-day decision making, or that their organizations had limited ability to influence school leaders. Some respondents suggested they would have needed more resources and

support from the program to avoid implementation challenges. Echoing their general improvement suggestions, respondents mentioned that more time spent on equity practices, additional follow-up opportunities, and more content on KIPP practices for recruiting external leaders may have eased their barriers to implementation.

Suggested Action: Provide strategies to support participants in implementing program content

Consider explicitly acknowledging various challenges participants may face in implementing content, to increase participant buy-in and encourage solutions-oriented problem solving.

Consider asking past program participants to provide written or video insights about their successes and challenges in implementing program content.

Consider embedding discussion questions throughout group activities that prompt participants to reflect on how implementing key learnings may be a challenge and to collaborate on solutions.

This page has been left blank for double-sided copying.

IV. Student and teacher outcomes in Successor Prep schools

Summary

- A majority of Successor Prep principals led their placement schools for at least three years, with many staying a fourth year.
- Successor Prep principals were similarly likely to remain in their schools for a second and third school year compared to other similarly experienced comparison principals in KIPP schools, although we cannot rule out substantial differences.
- Schools led by Successor Prep principals had similar student test scores, student retention rates, and teacher retention rates at the end of the principals' first and second years compared to outcomes in comparison KIPP schools, although we cannot rule out substantial differences.

A. Detailed research questions

The evaluation of the Successor Prep program sought to determine whether the program resulted in successful principal transitions that avoided the disruptions and dips in student outcomes that are sometimes observed (Clark et al. 2009; Coelli and Green 2012; Miller 2013; Walsh and Dotter 2018). To do this, we examined the following research questions:

1. How long do Successor Prep principals stay in their placement schools?
2. Are Successor Prep principals more likely to stay longer in their placement schools than other new principals who did not participate in Successor Prep?
3. Do schools that receive a Successor Prep principal have better outcomes—including student test scores and student and teacher retention rates—in each year following the leadership transition compared to outcomes in similar schools that did not receive a Successor Prep leader?

B. Study methods

The study used the following methods to answer each detailed research question. We label Successor Prep cohorts based on when the cohort began the program. For example, the January 2014 cohort became assistant principals in their placement schools in January 2014 and continued to participate in program activities as principals during the 2014–2015 school year. We measure the key outcome of student test scores using spring MAP scores in math and reading.

1. Research question 1

To answer research question 1, we examined how long Successor Prep principals remained in their placement schools. We calculated the proportion of the Successor Prep principals who were still leading their placement schools for a second, third, fourth, fifth, sixth, or seventh year. We calculated the proportions separately for each Successor Prep cohort, measuring retention for as many years as possible. For example, for the January 2014 cohort, we can measure principal retention for seven years (2014–2015 to 2020–2021). Because we can measure principal retention for at least three years for all five cohorts, we also calculated the proportions of Successor Prep principals who were still leading their placement schools for a second and third year across all five cohorts combined.

The sample for research question 1 includes 82 Successor Prep principals across five cohorts (Table IV.1). The number of Successor Prep principals in the sample is smaller than the total 150 who participated across all five cohorts because some participants came from outside KIPP and some participants never led KIPP schools. The 82 Successor Prep principals led 79 distinct schools because three schools were each led by two Successor Prep principals simultaneously.

2. Research question 2

To examine whether Successor Prep principals stay in their placement schools longer than non-Successor Prep principals, we used regressions to estimate the difference in the likelihood that a Successor Prep principal was still leading their placement school for a second and third year compared to non-Successor Prep principals. We restricted this analysis to both Successor Prep and comparison principals who were new to leading their schools in the same years to ensure that differences in tenure across the groups did not impact the results. Because differences in school performance might influence principal retention, we also accounted for differences in test scores using propensity score weights to achieve a comparison group of principals in schools that had test scores similar to the Successor Prep schools in each of the three years prior to the leadership transition. This requirement means that we excluded some principals in schools that did not have these test scores. However, prior research on principal transitions (in particular, Miller 2013) has shown that it is potentially important to account for trends in student achievement over multiple years prior to a principal transition.

We calculated the propensity score weights based on the prior test scores, a set of cohort indicators, and a middle school indicator. Even with the weights, small differences remained in the baseline characteristics of the weighted intervention and comparison groups, so our analyses controlled for the measures in the propensity score model, as recommended in the literature about propensity score weighting (Ho et al. 2007).

We also tried to restrict the comparison principals to those who are in the same KIPP regions as the Successor Prep schools. We could not make this restriction when focusing on new principals, because most new principals who are not Successor Prep principals were in different regions. However, we obtained similar results from a sensitivity analysis that restricted to schools in the same region by including more experienced principals in the analysis, as discussed in Appendix B. For this research question, we prefer the analysis that included only new principals because principal tenure is likely to be highly related to the amount of prior experience the principals have leading KIPP schools.

The analysis sample for this research question included 44 Successor Prep principals and 74 non-Successor Prep principals (Table IV.1). The number of Successor Prep principals in the sample is much smaller than for research question 1 because 18 of the Successor Prep principals had led their placement schools for at least a year before participating in Successor Prep and thus were excluded from the analysis. Another 20 led schools that did not have the prior test scores we needed so we could control for the possible impact of differences in school performance on principal retention. The 44 Successor Prep principals led 44 distinct schools. The 74 comparison principals led only 49 distinct schools because some of the schools were led by more than one principal in a given year or had multiple leadership transitions during this time period.

Table IV.1. Number of principals used in the analysis of research questions 1 and 2

	Successor Prep program cohort					
Condition	**January 2014**	**January 2015**	**January 2016**	**January 2017**	**January 2018**	**Total (unique)**
Research question 1: How long do Successor Prep principals stay in their placement schools?						
Successor Prep	21	10	15	15	21	82
Research question 2: Are Successor Prep principals more likely to stay longer in their placement schools than other principals who did not participate in Successor Prep?						
Successor Prep	11	4	11	9	9	44
Comparison	8	9	13	22	23	74

Source: Administrative data from KIPP.

Note: The sum of counts across cohorts in the comparison group row for research question 2 is 75 principals, which is larger than the count of 74 unique comparison principals because one principal served in the comparison group for two Successor Prep cohorts.

3. Research question 3

We assessed whether the Successor Prep program resulted in better test scores and student and teacher retention rates after the first two years of the transitions than if there hadn't been a principal transition. To do this, we used regressions to estimate differences in the outcomes between schools that received Successor Prep principals from one of five program cohorts and similar schools that never received Successor Prep principals from these cohorts. For each cohort of Successor Prep principals, we identified a comparison group of schools within the same KIPP regions that never received a Successor Prep principal and that had outcome data in multiple years before and after the Successor Prep principals first took over their schools. We then calculated propensity score weights that we used to achieve comparison groups that had outcomes similar to the Successor Prep schools in each of the three or four years before the Successor Prep principals took over leadership of their schools.

We calculated the weights based on the prior outcomes, a set of region indicators, a set of cohort indicators, and a middle school indicator. As in research question 2, our analyses controlled for the measures in the propensity score model.

We measured outcomes for the Successor Prep and comparison schools in each of the two years following a Successor Prep leadership transition. Unlike research question 2, which compared outcomes between Successor Prep and comparison group *principals*, for this research question we compared outcomes between Successor Prep and comparison group *schools*. In particular, we retained Successor Prep schools in the sample contributing outcomes even if the Successor Prep leader left the school. In effect, this approach considers any leadership transitions that follow the Successor Prep-supported transition to be a consequence of the Successor Prep program. We also retained comparison schools that experienced leadership transitions.

The comparison group pool for this research question differs from the one used in research question 2 for several reasons. First, we relaxed one of the eligibility criteria for comparison schools to retain as many schools in the analysis as possible. For this analysis, we did not restrict the sample only to Successor Prep and comparison principals who were new to leading their schools in the same years. We think this is justified because student and teacher outcome measures may be less likely to be influenced by leadership tenure compared to the principal retention outcome that is the focus of research question 2, lessening the

need for the restriction. However, three additional differences result in a smaller comparison pool for this analysis:

- This research question focuses on schools instead of principals, and there are fewer schools than there are principals in KIPP during this time period—research question 2 included 74 unique comparison principals, but only 49 unique comparison schools.
- To help ensure that Successor Prep and comparison schools are similar, we require that the comparison schools come from the same KIPP regions as the Successor Prep schools, which we did not require for research question 2 because of the focus on new principals.
- Although leadership tenure is known for nearly every KIPP principal in the analysis time period, for research question 3 we drop several schools that are missing student and teacher outcomes. Some schools that were eligible to be included in the comparison group for research question 2 are not eligible to be included in this analysis because they do not have outcomes in enough consecutive years—the analysis for research question 3 requires reading test scores in one additional prior year and requires a second follow-up year for the outcome measure. We examined the sensitivity of the results to excluding the additional prior year of test scores so that a few additional schools could be included in the analysis, and we found no substantive differences to the main results.

See Appendix B for more details on the study methods.

For the analysis of student test scores and student and teacher retention rates, the sample includes between 33 and 40 Successor Prep schools and 50 to 68 comparison observations, depending on the outcome measure. The number of observations included in the analysis for the comparison group is larger than the number of unique comparison schools, because the same school can be included in the comparison group for multiple cohorts. The comparison group includes between 36 and 42 unique schools, depending on the outcome measure. These counts overall and by Successor Prep cohort are shown in Table IV.2 for the math test outcome.

Table IV.2. Number of schools used in the analysis of math scores for research question 3

	Successor Prep program cohort				
Condition	**January 2014**	**January 2015**	**January 2016**	**January 2017**	**Total (unique)**
Successor Prep	12	6	12	10	40
Comparison	12	11	32	13	42

Source: Administrative data from KIPP.

Note: The sum of counts across cohorts in the comparison group row is 68 schools, which is larger than the count of 42 unique comparison schools because the same school can be included in the comparison group for multiple cohorts. A comparison school would contribute outcome data from different years for each cohort.

C. Characteristics of Successor Prep and comparison schools before the principal transitions

We use comparison groups to understand how the Successor Prep principal transitions may have influenced principal, teacher, and student outcomes in research questions 2 and 3. The differences in outcomes between the Successor Prep and comparison groups can be more plausibly attributed to the leadership transitions if the two groups were more similar before the transitions occurred. Therefore, we examine whether the Successor Prep and comparison schools appeared similar on pre-transition student

and teacher outcomes. Table IV.3 reports baseline outcomes for the year before participants in each Successor Prep cohort became principals, which is the same school year that they began the program. For example, the baseline data for the January 2014 cohort come from the 2013–2014 school year. Baseline outcomes in earlier years are reported in Appendix B.

Compared to all KIPP schools, the weighted samples of Successor Prep and comparison principals included in the analysis of research question 2 about principal tenure tend to be leading schools that are lower performing. These schools have lower math and reading scores compared to the scores for all KIPP schools at baseline. This is evident because the test scores reported in Table IV.3 are standardized relative to the full sample of KIPP schools, so a negative mean reflects a below-average score. For comparison, the unweighted means for all KIPP schools are reported in Appendix Table B.3 for all outcome measures.

In contrast, the weighted samples of Successor Prep and comparison schools in the analysis of research question 3 about student and teacher outcomes tend to be higher performing. In the weighted analytic sample for the student retention outcome, the average school has about 89 percent of students returning the following year. In the weighted analytic samples for the teacher retention outcome, the average school has about 64 percent of teachers remaining as a teacher at the school the following year, and about 73 percent remaining in some position (including non-teaching positions) at any KIPP school the following year. Each of those figures is lower than the overall KIPP averages. Just under two-thirds of the schools were middle schools, and the rest were elementary schools.

Differences between the weighted samples of Successor Prep and comparison schools at baseline were generally small (last column of Table IV.3). The What Works Clearinghouse (WWC) requires that baseline differences do not exceed 0.25 standard deviations in absolute value to Meet WWC Standards With Reservations. None of the baseline differences we measured in the weighted samples exceed that threshold. The WWC also requires that when baseline differences exceed 0.05 standard deviations in absolute value, the analysis include an appropriate statistical adjustment for the baseline measure. Our analyses controlled for the baseline outcomes and school characteristics. Because the Successor Prep program primarily aims to improve school-level outcomes, the effect sizes we report are measured in terms of school-level standard deviations. So an effect size of 0.04 for math test scores reflects a difference of 4% of the standard deviation of achievement across schools. If we instead measured this effect size in terms of student-level standard deviations, it would be smaller—roughly half as large—because there are larger differences in achievement scores across students than there are across schools.

Although the weighted samples of Successor Prep and comparison schools are similar, the Successor Prep schools overall tend to be lower performing than the schools in the comparison pool before we apply the weights, as reported in Appendix Table B.3.

Table IV.3. Baseline differences between Successor Prep and comparison schools are substantively small

Sample and baseline measure	Successor Prep			Comparison group			Difference (effect size)
	Number of observations	Mean	Standard deviation	Number of observations[a]	Mean	Standard deviation	
Principal retention analytic sample (research question 2)							
Math MAP scores	44	-0.12	0.96	75	-0.10	0.80	-0.02
Reading MAP scores	44	-0.12	0.95	75	-0.11	0.84	-0.01
Middle school	44	63.5	--	75	63.9	--	-0.01
Math test score analytic sample							
Math MAP scores	40	0.22	0.96	68	0.19	0.79	0.04
Middle school	40	59.6	--	68	59.8	--	0.00
Reading test score analytic sample							
Reading MAP scores	36	0.09	0.95	53	0.11	0.81	-0.02
Middle school	36	66.9	--	53	66.2	--	0.02
Student retention analytic sample							
Percentage of students remaining at the school the following year	33	88.5	6.8	50	89.3	5.6	-0.12
Math MAP scores	33	0.14	0.91	50	0.18	0.86	-0.04
Reading MAP scores	33	0.08	0.96	50	0.11	0.82	-0.03
Middle school	33	64.0	--	50	66.3	--	-0.05
Teacher retention in the same school analytic sample							
Percentage of teachers still teaching at the school the following year	33	63.8	11.8	50	65.5	12.9	-0.14
Math MAP scores	33	0.09	0.91	50	0.14	0.86	-0.06
Reading MAP scores	33	0.03	0.96	50	0.06	0.82	-0.04
Middle school	33	66.8	--	50	61.3	--	0.11
Teacher retention in any KIPP school analytic sample							
Percentage of teachers in any position in any KIPP school the following year	33	72.1	13.1	50	74.1	12.5	-0.16
Math MAP scores	33	0.17	0.91	50	0.16	0.86	0.01
Reading MAP scores	33	0.12	0.96	50	0.08	0.82	0.04
Middle school	33	61.5	--	50	60.8	--	0.01

Source: Administrative data from KIPP.

Note: The baseline measures reported in this table are those from the year before Successor Prep participants became principals. The means shown in the table have been adjusted using the propensity score weights calculated for each analysis. Standard deviations are unweighted. The number of observations reflects the number of principals for research question 2, and schools for the other outcomes. The means and standard deviations for the test scores are measured in units of school-level standard deviations. Whether a school is a middle school is a dichotomous measure, so the standard deviations are not reported. Effect sizes are Hedges' *g* effect sizes representing school-level standard deviations with an adjustment for small sample sizes (WWC 2022), including for the dichotomous middle school measure.

[a] For the comparison condition, the 75 observations used in research question 2 reflect 74 distinct principals who led 49 distinct schools, including schools that were led by more than one principal in a given year or that had multiple leadership transitions during this time period. For the other outcomes, the same comparison school can be counted multiple times because it can appear in the comparison group for multiple Successor Prep cohorts. For example, the number of unique comparison schools for the analysis of math scores is just 42.

D. Limitations

Our analyses of the Successor Prep program have five limitations:

- Like any non-experimental comparison group design, the findings may reflect preexisting differences between the Successor Prep and comparison schools, rather than the actual causal effect of the Successor Prep-supported principal transitions. We use propensity score weights to construct similar groups of Successor Prep and comparison principals or schools, which successfully create groups that are similar on important baseline characteristics of schools. However, they do not guarantee that the two groups of schools are similar in every way that might be related to outcomes. If the comparison group schools tend to be those that have access to additional resources, for example, then our results will understate the extent to which the Successor Prep-supported transitions might lead to more favorable outcomes.
- Because nearly all KIPP schools from within Successor Prep regions that experienced principal transitions were Successor Prep schools, we could ensure either that the Successor Prep and comparison schools were in the same regions or that the principals in these schools had the same amount of leadership experience, but not both. For research question 2, we prioritized obtaining novice Successor Prep and comparison principals because experience likely has a strong relationship with the principal tenure outcome. For research question 3, we prioritized obtaining Successor Prep and comparison schools in the same region because the student and teacher outcomes are less likely to be influenced by the prior experience of leaders. However, we did examine the sensitivity of both of these decisions to different choices about the comparison group and obtained similar results, as discussed in Appendix B. In particular, we conducted (1) an analysis of the principal tenure outcome using a sample that included more experienced principals but excluded principals who led schools in regions that did not have Successor Prep principals and (2) an analysis of the student and teacher outcomes that controlled for tenure. Few leadership transitions outside of the Successor Prep regions occurred in schools that met the other eligibility requirements for research question 3, so it was not feasible to conduct a sensitivity analysis for the student and teacher outcomes that restricted to new principals.
- Because the comparison schools in our analyses of student and teacher outcomes include experienced KIPP principals, we compare KIPP Successor Prep schools to comparison group schools in KIPP that did not generally experience leadership transitions. Because of this, we are unable to assess how well the Successor Prep program supports principal transitions, since we do not observe outcomes in a comparison group of schools that experienced leadership transitions without being supported by Successor Prep. It would likely be necessary for KIPP to test different versions of the Successor Prep programs or selectively support only some leadership transitions to measure how well the Successor Prep program supports principal transitions.
- The findings may not be generalizable to other settings. The study schools are KIPP schools, which may be dissimilar from other charter or traditional public schools. Consequently, the impacts of the Successor Prep program may be different in these other settings. Also, the leadership programs were implemented by KIPP in KIPP schools, and different implementation challenges may arise in other settings. In particular, the impacts of the Successor Prep program in KIPP may be different from the impact that might be obtained when providing the Successor Prep training to other principals because Successor Prep principals are intentionally selected by KIPP; our findings reflect both the effects of participating in the program and the effects of KIPP's selection process.

- The study's sample is small, which limits our ability to detect statistically significant differences in principal retention rates and in the student and teacher outcomes between the Successor Prep and non-Successor Prep schools. Although we sometimes detect substantively large differences between Successor Prep and comparison schools on outcomes, when these differences are not statistically significant, we cannot rule out that those differences were due to chance.

E. Results

1. Most Successor Prep principals led their placement schools for at least three years.

Across all five cohorts, 82 percent of the Successor Prep principals stayed in their placement schools a second year (that is, the year after they first took over leadership of their schools), and 67 percent stayed in their placement schools a third year (last row of Table IV.4). Few of the Successor Prep principals stayed in their placement schools more than four years.

The results were generally similar across each cohort, with some exceptions. The number of principals in each cohort is small, so we recommend caution in interpreting these exceptions. The proportion of principals returning for a second year is lower for the later cohorts—it was 95 percent for the January 2014 cohort, but just 71 percent for the January 2018 cohort. Except for the 2018 cohort, this pattern is not as pronounced for the proportion of principals returning for a third year, although retention rates for the January 2018 cohort were still lower. Just 57 percent of principals in this cohort returned for a third year, compared to 67 to 73 percent of principals in the earlier four cohorts. This cohort experienced the onset of the COVID-19 pandemic in the spring of their second year as a Successor Prep principal, which might have influenced whether they returned for a third year (the 2020–2021 school year). However, retention rates for the 2020–2021 school year were not noticeably lower for other cohorts compared to those for cohorts who experienced the COVID-19 pandemic later in their careers. For, example, 53 percent of principals in the January 2017 cohort returned for a fourth year in 2020–2021, a rate that is at or above the fourth-year retention rates for the earlier cohorts.

Table IV.4. Most Successor Prep principals remained in their placement schools for a second and third year

Cohort	Number of leaders	Percentage of Successor Prep leaders still leading their placement school for a: Second year	Third year	Fourth year	Fifth year	Sixth year	Seventh year
Jan. 2014	21	95%	67%	38%	19%	10%	5%
Jan. 2015	10	90%	70%	40%	10%	10%	
Jan. 2016	15	80%	73%	53%	27%		
Jan. 2017	15	73%	73%	53%			
Jan. 2018	21	71%	57%				
Across all cohorts	**82**	**82%**	**67%**				

Source: Administrative data from KIPP.

Note: The table reflects school leader placements from the 2014–2015 to 2020–2021 school years. The table shows leader tenure for each cohort for as many subsequent years as are available.

2. Retention rates for Successor Prep principals in their second and third years were statistically indistinguishable from those for non-Successor Prep principals, but we cannot rule out substantial differences.

We did not find statistically significant evidence that the 44 Successor Prep principals who were new to leading their placement schools were more or less likely to continue leading those schools for a second year than comparison principals who were also new to leading their schools (Table IV.5). Although the results were not statistically significant, the Successor Prep principals were 7 percentage points more likely to remain in their school a third year compared to the comparison leaders. We conducted a similar analysis that compared the principal retention rates of a larger group of 55 Successor Prep principals and 91 non-Successor Prep principals, included regardless of whether they were new to leading their schools, which also did not find differences in the retention rates (see Appendix B).

Table IV.5. Successor Prep and non-Successor Prep principals were similarly likely to stay in their schools for a second or third year

Principals still leading their school:	Percentage: Successor Prep	Percentage: Comparison	Percentage point difference (standard error)	Effect size
For a second year	83.0	79.3	3.8 (7.8)	0.09
For a third year	66.9	59.9	7.1 (9.7)	0.14

Source: Administrative data from KIPP.

Note: The regression sample included 44 Successor Prep principals from 44 schools and 75 comparison group records, representing 74 distinct leaders from 49 distinct schools. The regressions are weighted using propensity score weights, control for prior math and reading test scores, and include a set of cohort indicators and a middle school indicator. The standard errors are calculated accounting for clusters by school. Effect sizes are Hedges' *g* effect sizes.

* Significantly different from zero at the .10 level, two-tailed test.

** Significantly different from zero at the .05 level, two-tailed test.

3. Student and teacher outcomes in Successor Prep schools at the end of the first two years were statistically indistinguishable from outcomes in non-Successor Prep schools, but we cannot rule out substantial differences.

We did not find statistically significant evidence that the Successor Prep program changed student and teacher outcomes within the first two years beyond what might have happened if the previous leader had remained in the school. Test scores and student and teacher retention rates of the Successor Prep schools at the end of the first and second years after the transition were not statistically different from those for similar non-Successor Prep schools in the same school years (Table IV.6).

Although no findings were statistically significant, we cannot rule out large differences in outcomes between Successor Prep and comparison schools. We find a substantively large difference between the reading test scores in the Successor Prep schools at the end of the first year after the transition, compared to test scores in comparison schools. The estimated difference in reading scores at the end of the first year of -0.11 standard deviations of school-level student achievement (based on the KIPP-wide distribution of scores) is equivalent to a reduction in the average school's performance by 4 percentiles, and is larger than the negative impacts of principal turnover found in other studies (for example, Miller 2013). An impact of this magnitude indicates there may be potentially meaningful costs to students from the leadership transition, at least in the first year. By the second year, there was no longer an estimated negative impact on reading test scores.

The effect sizes we measure for test scores are reported in units of school-level standard deviations to emphasize how outcomes for schools change after a Successor Prep leadership transition. However, other studies report effect sizes in units of student-level standard deviations. To allow this comparison, we also used the student-level test scores to report effect sizes translated into student-level effect sizes in Appendix Table B.5. For example, the school-level effect size of -0.11 for reading scores at the end of the first year is -0.05 when measured in student-level standard deviations.

The findings in Table IV.6 are supported by secondary analyses that the study team conducted of these outcomes, which generally found similar results (see Appendix B). In particular, even for the early cohorts that had outcomes after three, four, and five years, the Successor Prep schools still had similar outcomes after those years compared to non-Successor Prep schools.

Table IV.6. Successor Prep schools did not have better student and teacher outcomes than comparison schools

Outcome	First year the Successor Prep principals led their schools				Second year the Successor Prep principals led their schools			
	Successor Prep	Comparison	Difference (standard error)	Effect size	Successor Prep	Comparison	Difference (standard error)	Effect size
Math MAP scores	0.16	0.20	-0.04 (0.08)	-0.04	0.27	0.25	0.02 (0.09)	0.02
Reading MAP scores	0.04	0.15	-0.11 (0.09)	-0.11	0.21	0.19	0.02 (0.09)	0.02
Percentage of students remaining at the school the following year	89.2	89.1	0.1 (0.7)	0.01	89.3	88.3	1.0 (0.8)	0.15
Percentage of teachers still teaching at the school the following year	65.3	64.8	0.5 (3.2)	0.01	64.4	67.0	-2.6 (4.3)	-0.17
Percentage of teachers in any position in any KIPP school the following year	73.9	73.4	0.5 (2.6)	0.01	74.0	73.7	0.4 (4.1)	0.02

Source: KIPP administrative data.

Note: The math test scores regression samples include 40 Successor Prep schools and 68 comparison school-cohort observations. The reading test scores regression samples included 36 Successor Prep schools and 53 comparison school-cohort observations. The student and teacher retention regression samples include 33 Successor Prep schools and 50 comparison school-cohort observations.

The regressions are weighted using the propensity score weights, control for prior outcomes, and include a set of cohort indicators, a set of region indicators, and a middle school indicator. The test scores are measured in units of school-level standard deviations. The standard errors are calculated accounting for clusters by school. Effect sizes are Hedges' *g* effect sizes representing school-level standard deviation units with an adjustment for small sample sizes (WWC 2022).

* Significantly different from zero at the .10 level, two-tailed test.

** Significantly different from zero at the .05 level, two-tailed test.

This page has been left blank for double-sided copying.

V. The reliability and validity of the Fisher Fellowship selection instrument

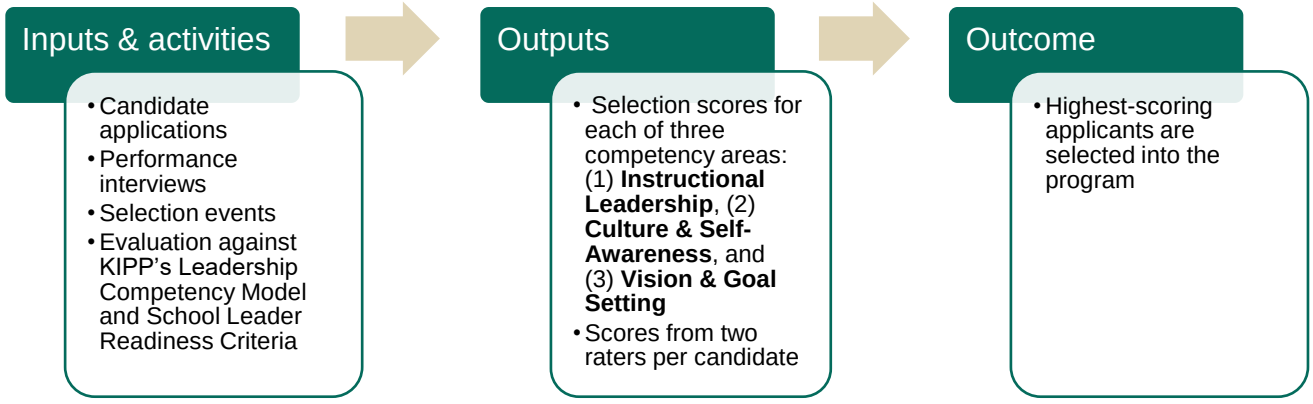
Summary

- Scores on each competency, when combined, do not contribute equally to a candidate's overall selection score. Alternative weighting techniques can be used to address this.
- Each competency within the selection instrument consistently measures a distinct dimension of leadership potential.
- Multiple raters tend to agree on scores for the same candidate, indicating distinctions between scores across all candidates are likely due to differences in leadership potential rather than rater judgments. For some items, KIPP may consider improving guidance and training for raters.
- We were unable to learn whether the instrument predicts who will be an effective school leader. A larger pool of principals, additional information about how Fisher Fellowship principals are placed into schools, and outcomes measured before a leader might influence school outcomes would support measuring the predictive validity of a selection instrument in the future.

A. Detailed research questions

One component of an effective school leadership program is an approach to selecting participants with potential to become effective school leaders. As shown in the selection model below (Figure V.1), several inputs go into selecting participants for the Fisher Fellowship program to prepare principals to lead new KIPP schools. Applicants who pass through an initial interview are invited to a selection event. During these events, KIPP uses its School Leader Readiness Criteria selection tool to identify promising candidates who demonstrate key leadership competencies. Candidates with the highest scores are typically selected into the program.

Figure V.1. Fisher Fellowship program selection model



Understanding how well KIPP's leadership program selection process identifies promising candidates with high potential for effective leadership can provide insight about how to improve the selection process. A refined and enhanced selection tool can better support KIPP to identify these promising candidates. We analyzed data from KIPP's School Leader Readiness Criteria selection tool to understand the following:

1. **How do the three competencies measured in the selection tool contribute to overall scores?** Understanding these descriptive properties of the competency scores can inform changes to the design of the selection instrument that might better guide selection decisions. Calculating an overall score supports KIPP to make quick assessments about candidates' potential. KIPP can combine this overall assessment with information from the scores on each competency about candidates' relative strengths and weaknesses when selecting program participants and considering opportunities to tailor the training to better develop their talent. To understand what KIPP can learn about candidates from the overall scores, we assessed (1) how much each competency is weighted in the overall score and (2) whether candidates who score highly on one competency also tend to score highly on the others.
2. **Do multiple items and raters provide consistent measures of leadership potential?** When an instrument has high reliability, it has the potential to make meaningful distinctions between participants. We measured the reliability of the instrument that KIPP used to select candidates for the Fisher Fellowship program using data from the 2017 and 2018 selection rounds. We examined two measures of reliability: Interrater reliability measures whether multiple raters assign the same scores to the same participant, and internal consistency measures whether scores on different items intended to measure the same competency are similar.
3. **How strongly are scores on the selection instrument associated with leadership outcomes?** When an instrument has high predictive validity, participants with higher scores on the instrument are more effective once they become school leaders—that is, the instrument can predict who will be an effective principal. In this context, predictive validity is the extent to which scores on the instrument are associated with outcomes that reflect effective school leadership. To assess the predictive validity of the instrument, we had intended to correlate scores on the instrument with leadership outcomes including those based on student test scores. However, we were not able to conduct this analysis as planned due to significant limitations of the data, as discussed below.

B. Study methods

1. Measuring how the three competencies in the selection tool contribute to overall scores

KIPP calculates competency scores as the simple sum across all items measuring that competency. Overall scores for candidates are computed as the simple sum across all items. KIPP may wish to give equal consideration to each competency when making selection decisions or may wish to prioritize one competency above the others. In either case, KIPP can be intentional about the relative weighting of the competencies in the overall score to achieve their goals. However, the simple sum is unlikely to be the optimal choice for achieving those goals.

If the score distributions across the three competencies are dissimilar, an overall score calculated as either a simple sum or average across competencies will not give equal weight to each competency. This could result in KIPP making decisions about leadership potential without giving due consideration to each candidate's strengths. On the other hand, if the same candidates tend to score high on all three competencies while other candidates tend to score low on all three, then the relative weighting of items is not important, but KIPP might need to revisit whether revisions to the tool could better distinguish candidates on each skill.

The extent to which a competency contributes to the overall score is referred to as the effective weight for that competency. A competency will have a larger effective weight when it contributes more points to an overall total or has a greater share in the spread of scores (that is, a greater standard deviation). To

understand how much each competency contributes to the overall score, we calculated the average competency score, standard deviation of the scores across candidates (a measure of dispersion in the scores), and the range of scores among the candidates for each competency. Following KIPP's approach to determining candidates' scores on each competency, we calculated competency scores as the sum of all items within the competency, where raters assigned a score from 1 to 4 (including half increments) to candidates on each item.

We also calculated correlations to understand whether candidates who score high on one competency also score high on the other two competencies. Correlations can range from –1 to 1, where larger, positive correlations indicate that scores on one competency are more similar to scores on another competency. Results from this analysis can indicate whether the competencies are measuring similar or distinct skills—for instance, whether candidates who demonstrate high competency in Vision & Goal Setting also demonstrate high Instructional Leadership skills. High correlations between each pair of competencies would indicate they are not complementing each other and instead are measuring related skills.

2. Measuring the reliability of the Fisher Fellowship selection instrument

To measure reliability, we calculated the internal consistency and interrater reliability of the measures:

- **Internal consistency** measures the degree to which a set of item scores—such as those for the 13 items within the Instructional Leadership competency—yield consistent assessments of leadership potential as measured by the overall score on the competency. Internal consistency is high when applicants' scores tend to be similar across items. High internal consistency could reflect similarity in the content and skills measured by each item, suggesting the competency consistently measures a single dimension of leadership potential. Low internal consistency signifies that scores differ on items within a competency, which could be due to items that do not consistently measure the same dimension of leadership potential or measurement error that results from poor question wording or an item not measuring what it is intended to measure.
- **Interrater reliability** describes the degree of agreement between scores calculated by different raters for the same candidate. Interrater reliability is high when both raters of a candidate come to the same conclusions and assign the candidate the same score on each item. Interrater reliability is highest when raters have clear instructions, are well trained, and agree on the leadership traits they seek to identify and how to measure them. Low interrater reliability could indicate a need for better guidance or training for raters or a need to reword some items so that raters do not rely on their own interpretations.

When reliability on either type of measure is high, KIPP can be more confident that distinctions between candidate scores are due to differences in leadership potential. Also, with high internal consistency, KIPP can rely more on group summaries when the items in the group are more consistent.

We measured internal consistency of each competency using Cronbach's alpha, which indicates the degree to which it is permissible to replace a group of items with a summary measure. Cronbach's alpha ranges between zero and one, where larger values are associated with higher levels of internal consistency (Cronbach 1951). The formula to calculate Cronbach's alpha for a weighted average of items is:

$$\alpha = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1}^{K} \omega_{c_i}^2 \sigma_{c_i}^2}{\sigma_S^2}\right)$$

where K is the number of items, σ_S^2 is the variance of the overall score, $\omega_{c_i}^2$ is the weight on item i in the calculation of the overall score, and $\sigma_{c_i}^2$ is the variance of item i. We measured internal consistency using scores on the instrument for 80 candidates from the 2017 and 2018 selection rounds. The instrument has 10 items in the Culture & Self-Awareness competency, 11 items in the Vision & Goals competency, and 13 items in the Instructional Leadership competency.

Researchers have reached different conclusions about what value of alpha should be considered "good." The answer can depend on a scale's use. For instance, Wasserman and Bracken (2003) recommend that alpha values should exceed 0.8 for psychological assessment scales if the scales have high-stakes consequences for individuals. Many other researchers have recommended 0.7 as an acceptable alpha value (Cortina 1993). As we are not aware of any existing guidelines for the internal consistency of measures used to assess leadership potential, we apply David de Vaus's recommendation, in his widely cited textbook on surveys in social research, that alpha values above 0.8 are considered good and alpha values above 0.7 are considered acceptable (de Vaus 2002).

We also calculated a measure of each item's contribution to the internal consistency of the overall score on a competency, thereby providing information that could enable KIPP to determine whether particular items ought to be included in the instrument. To do this, we calculated internal consistency when excluding one item at a time, which is sometimes called a leave-out alpha. If internal consistency is notably larger when excluding a particular item, this would tell us that the left-out item may be less important to measuring the construct of interest or measures a different competency or skill.

To measure interrater reliability, we calculated the intraclass coefficient, or the proportion of variance in the summary scores due to differences in leadership potential rather than to differences in rater judgments about leadership potential. We calculated this using a series of regressions, one for each item, with a random effect:

$$s_{ir} = \mu + \theta_i + \varepsilon_{ir}$$

where s_{ij} is the score on the item for candidate i given by rater r, μ is an intercept representing the mean score across the sample of candidates, θ_i is the leadership potential of candidate i, and ε_{ir} is an error term that reflects differences in how raters evaluate the same candidate. We estimated the variance of θ_i, given by σ_θ^2, representing the dispersion in scores we would expect to see if all raters always agreed, using a random effect. The intraclass coefficient is then given by $\hat{\sigma}_\theta^2 / \hat{\sigma}_S^2$, where $\hat{\sigma}_S^2$ is the observed variance in the scores across participants and raters. We measured internal consistency on each item using selection scores from 79 candidates from the 2017 and 2018 selection rounds that had two scores per item from two different raters.

As with internal consistency, researchers have not agreed to a single acceptable level of interrater reliability. However, one point of comparison is the reliability of classroom observations used by schools to measure teachers' use of certain practices. Ho and Kane (2013) report that one of these measures, the Danielson Framework for Teaching, has an intraclass coefficient of about 0.6 when scored by two trained observers.

Internal consistency and interrater reliability provide different information about the appropriate level of confidence in the scores, and these measures cannot be compared to each other. We calculated both sets of measures using combined scores from participants from the 2017 and 2018 cohorts.

3. Measuring the validity of the Fisher Fellowship selection instrument

If the instrument is faithfully measuring leadership potential, then participants who are higher scoring during selection will also be more likely to have success as school leaders. The association between scores on the instrument and performance as a school principal measures the instrument's predictive validity. When predictive validity is high, KIPP can be more confident that high-scoring candidates have potential to be successful.

To measure predictive validity in this study, we had intended to calculate the correlation between scores on the selection instrument and schoolwide measures of student achievement after Fisher Fellows completed the program and led the school after one year. Specifically, we planned to correlate selection scores from candidates who went through selection in 2017 with outcomes from spring 2019, as well as those who went through selection in 2018 with outcomes from spring 2020, at the end of their first year leading a school. We had also planned to adjust these correlations by controlling for fall 2018 and fall 2019 test scores, respectively, to account for how leaders might be assigned to schools. This adjustment might be necessary if leadership outcomes, such as spring 2019 test scores for the first cohort, are at least partially outside of the control of the Fisher Fellowship principals. For example, some principals might be placed in more challenging settings. In particular, if leaders with higher scores on the instrument are more likely to be placed in more challenging schools, then student outcomes could appear worse in these high-scoring leaders' schools compared to those in schools of lower-scoring leaders. The adjustment for fall 2018 test scores is intended to account for this possibility.

However, we could not measure the predictive validity of the Fisher Fellowship selection instrument because we could not include enough Fisher Fellowship principals in the analysis. First, although we had selection scores for both the 2017 and 2018 Fisher Fellowship cohorts, we could not include the 2018 cohort because of the COVID-19 pandemic. The primary outcome of interest is student test scores in spring 2020, but these were not collected. Second, there were too few Fisher Fellowship principals from the 2017 cohort who could be included in the analysis. Based on program data, 38 Fisher Fellowship candidates went through selection in 2017, and of those, 12 participated in the program in 2018. Among these participants, eight went on to lead a school at KIPP the following year and had mathematics test scores from spring of that year; seven of those eight had reading scores. A small sample limits the strength of our conclusions about the relationship between a leader's selection scores and student achievement. This is because the leaders who we can include may not be representative of the full population of Fisher Fellowship principals and because the results based on so few leaders are highly imprecise. Because of the small sample, we do not report the results of this analysis.

The analysis would have been feasible had there been available a larger number of Fisher Fellowship principals who had selection scores and outcome data one year after completing the program. Two additional pieces of information could also strengthen a future analysis: (1) an understanding of the process by which Fisher Fellowship principals are placed into schools and (2) additional measures of the characteristics of students in the schools and of the school environment shortly after the schools open, such as the school climate or teachers' instructional practices.

C. Findings

1. Research question 1: Understanding the overall scores

a. Contributions of each competency to the overall score

The scores on each competency varied widely across candidates, as shown in the minimum and maximum scores in Table V.1. This means that the tool has the potential to make distinctions between candidates on all three competencies.

Next, the properties of the scores on each competency differed. In particular, the average competency score and standard deviation of the scores differs for each competency in Table V.1. For example, scores on the Instructional Leadership competency tended to be higher on average and had a higher standard deviation. Because its mean and standard deviation are higher, the Instructional Leadership competency has a larger effective weight in the summed score calculated by KIPP compared to the other two competencies. This means that this competency is given more importance when considering only the overall score. In other words, a candidate scoring at or near the top on the Instructional Leadership competency will, on average, tend to have a higher overall score than a candidate scoring at or near the top of either of the other two competencies.

Table V.1. Fisher Fellowship 2017 and 2018 selection score summary statistics, by competency

Competency	Minimum score	Maximum score	Average score (standard deviation)	Highest possible score
Culture & Self-Awareness	35	67	56 (7)	80
Vision & Goals	36	81	61 (8)	88
Instructional Leadership	40	94	64 (12)	96

Source: Administrative data from KIPP.

b. Whether candidates who score high on one competency also tend to score high on the others

As shown in Table V.2, each competency is positively correlated with the other two. Although the correlations are positive, they are not strong (below 0.4), meaning candidates who score high on one competency tend not to score as high on the others. This suggests that each competency is measuring a distinct set of skills, or competency.

Table V.2. Correlations between Fisher Fellowship selection instrument competency scores

Competency	Correlation with Culture & Self-Awareness	Correlation with Vision & Goals
Culture & Self-Awareness	0.39	
Vision & Goals	0.36	0.22

Source: Administrative data from KIPP.

c. Implications and recommendations for calculating overall scores

If each competency is considered equally important to a candidate's leadership potential, then KIPP should give each one equal weight in the overall score. To do this, KIPP may want to develop scales that result in more similar means and standard deviations for each competency. However, it is not necessary to revise the tool to do this. Instead, KIPP could standardize the scores on each competency across candidates by subtracting its mean and dividing by its standard deviation. This approach would ensure equal weighting of items. If KIPP prefers to retain the simple sums for each competency, an equivalent solution would be to calculate the nominal weights that would be needed to achieve equal effective weights. Because candidates do tend to differ in their performance across the three competencies, the weighting of the scale is potentially consequential. Equal weighting will prevent candidates who excel on the instructional leadership competency to be overrepresented among candidates with the highest overall scores.

If KIPP does not wish to give each scale equal weight, one approach to choosing the weights would be to give the competencies weights that reflect their relationships to outcomes that measure effective leadership. An analysis of the instrument's validity such as the one we were unable to complete for research question 3 could inform these weights.

2. Research question 2: Internal consistency of the Fisher Fellowship selection instrument

a. Overall internal consistency

Summary measures for each competency have high internal consistency. As shown in Table V.3, Cronbach's alpha for the Culture & Self-Awareness competency is 0.85, Vision & Goals is 0.87, and Instructional Leadership is 0.84. These findings indicate that the simple sum of scores on the items within a competency used by KIPP (or, alternatively, a simple mean) provides a meaningful summary of candidates' potential for that competency. Instead of having to consider each item when making offers to candidates or dividing the competencies into smaller groups of related items, KIPP can confidently rely on the three summary scores.

Table V.3. Competencies in the selection instrument have high internal consistency

Competency	Internal consistency (Cronbach's α)
Culture & Self-Awareness	0.85
Vision & Goals	0.87
Instructional Leadership	0.84

Source: Administrative data from KIPP.

b. Items that contribute to lower internal consistency

No item contributes to a meaningfully higher or lower internal consistency within any competency. As expected for a tool that includes relevant items within each competency, the internal consistency of each competency is generally slightly higher than the one reported in Table V.3 when excluding individual items one at a time. These leave-out alphas are reported in Appendix Tables C.1–C.3. For the Culture & Self-Awareness and Instructional Leadership competencies, there are a small number of items that, if removed, would slightly improve the reliability of the summary score, but not in a meaningful way.

c. Interrater reliability of items

For most instrument items, there was moderate to strong agreement between scores given to the same candidate by different raters. Just three items had low interrater reliability, meaning their intraclass coefficients were below 0.60. These included one item from the Culture & Self-Awareness competency and two from the Vision & Goals competency (see Table V.4). These low reliabilities may indicate that raters tend to disagree on a leader's competency level. This could happen for several reasons, including that the raters might misunderstand the item or bring their own opinions and interpretations to scoring the item, or the item itself is unclear or ambiguously worded. These issues could lead to mismeasuring a candidate's true leadership potential, if not addressed. However, these intraclass coefficients are just below our threshold for acceptable values and affect only three of the 31 items. The remaining 28 items all have interrater reliabilities that exceed 0.60 (reported in Appendix Tables C.4–C.6), so the three items in Table V.4 do not pose a major concern for the tool.

Table V.4. Fisher Fellowship selection instrument items with low interrater reliability

Competency	Item	Interrater reliability (intraclass coefficient)
Culture & Self-Awareness	1b: Self-awareness in response to "Why do you want to lead a KIPP school in [this community]?"	0.59
Vision & Goals	1b: Student focus in response to "What is your vision for the school you want to lead? Please describe both your instructional vision and your cultural vision."	0.52
Vision & Goals	1c: Achievement orientation in response to "What is your vision for the school you want to lead? Please describe both your instructional vision and your cultural vision."	0.58

Source: Administrative data from KIPP.

d. Implications and recommendations for the instrument

Because the instrument has high reliability overall, there is no pressing need to revise the instrument, improve training, or provide new guidance to raters. However, we recommend considering the following:

- **Focus improvements to the tool on the items with lower interrater reliability.** Although we do not have detailed information about the guidance or training provided to raters, there may be opportunities for improvement in these areas. Consider refining guidance or instructions on how to score these items, implementing tailored rubrics for scoring them with distinct definitions or examples of what type of response should score a 1, 2, 3, or 4. Consider also providing additional training to raters on how to objectively complete the tool, rather than ranking candidates relative to other candidates the rater has evaluated.
- **Consider whether some items are redundant**. High internal consistency means that the item scores are all highly related to one another. Knowing this, KIPP could potentially reduce the number of items without losing valuable information. This could allow KIPP to identify promising candidates more quickly and efficiently, requiring less time from raters and in interviews. However, if raters can already complete the tool efficiently, this may not be necessary. Also, there is not enough evidence to conclude that these items are actually redundant and that they are measuring very similar and specific

skills. It would be reasonable to keep the length of the instrument as is, as the results indicate that each competency is well measured.

3. Research question 3: Predictive validity of the Fisher Fellowship selection instrument

Although we could not conduct this analysis as planned, we do offer some recommendations that could support measuring the predictive validity of a selection instrument in the future. There may be actions KIPP could take to prepare to learn whether a future instrument can identify leadership potential.

First, a larger pool of principals is necessary to conduct the analysis. It will be necessary to plan to include multiple cohorts, or larger cohorts. This evaluation had intended to include multiple cohorts, but even had the COVID-19 pandemic not occurred, the sample would have been small. Second, KIPP should document how Fisher Fellowship principals are placed into schools. Our research design included an effort to account for the possibility that leadership outcomes, such as student test scores, are not fully in the control of principals, and that principals with higher scores on the selection tool may be more likely to be placed in more (or less) challenging settings. However, efforts to account for this possibility as we had planned depend on measuring the aspects of school settings that might affect outcomes, but these aspects are often not measured in administrative data. An alternative approach would be to intentionally place principals into schools using a set of rules that could allow researchers to fully account for the placement process when conducting the research. For example, if principals were ranked based on their scores on the selection instrument and matched, in order, to schools that are also ranked based on a measured characteristic, such as the poverty rate of the community, we could be much more confident that accounting for the poverty rate would address this concern and result in an accurate measure of predictive validity. Such a process may not always be feasible, however. Last, measuring outcomes in schools before a leader might influence outcomes so that they can be accounted for in an analysis can improve the measure of predictive validity. Fall test scores were measured, but additional information about outcomes in the schools would be valuable, such as measures of school climate or teachers' instructional practices. Measuring these prior outcomes was a particular challenge for the Fisher Fellowship program because the principals were leading new schools, so outcomes from the years before the Fisher Fellowship principal, such as student retention in the same school, teacher retention, and the use of effective leadership practices, do not exist. An alternative is to measure proxies such as the poverty rate in a school's community.

This page has been left blank for double-sided copying.

Appendix A.
Additional findings from the survey of KIPP school leadership program participants

This page has been left blank for double-sided copying.

This appendix provides additional information about the surveys and additional details on the results from the analysis of the surveys.

Appendix A.1. Survey response options for why respondents participated in the programs

Boxes 1 and 2 list the options presented to program participants in the survey when asked about why they participated in a program.

Box 1. List of survey response options for the Successor Prep program participants on why they participated in the program

What were the **three most important** reasons you participated in the Successor Prep program?

SELECT UP TO THREE

- ☐ I wanted to become a school leader of an existing school.
- ☐ I wanted to establish a new school.
- ☐ I wanted a leadership role (other than school leader) within a school.
- ☐ I wanted to have an impact on more children.
- ☐ It was required for my current or future position.
- ☐ I was looking for a challenge.
- ☐ I did it to increase my earnings.
- ☐ It was a steppingstone to achieve a career goal.
- ☐ My school leader or supervisor encouraged me to do it.
- ☐ I wanted to improve my leadership skills.
- ☐ I wanted to move to a different school.
- ☐ I wanted to train and prepare school leaders more effectively.
- ☐ I wanted to learn about how KIPP trains and prepares school leaders.
- ☐ Something else *(specify)*

Box 2. List of survey response options for the KLDF program participants on why they participated in the program

What were the **three most important** reasons you participated in the KIPP Leadership Design Fellowship?

SELECT UP TO THREE

- ☐ I wanted to train and prepare school leaders more effectively.
- ☐ I wanted to learn about how KIPP trains and prepares school leaders.
- ☐ My superintendent, executive director, academic officer, or supervisor encouraged me to participate.
- ☐ I wanted to have an impact on more children.
- ☐ It was required for my position or was a stepping stone to achieve a career goal.
- ☐ I was looking for a challenge.
- ☐ I did it to increase my earnings.
- ☐ I wanted to improve my leadership skills.
- ☐ Something else *(specify)*

Appendix A.2. Details on respondents' satisfaction with the program content and delivery

Figures A.1 and A.2 report how program participants responded when asked about their satisfaction with certain aspects of the programs.

Figure A.1. Overall satisfaction with content and delivery of the Successor Prep program

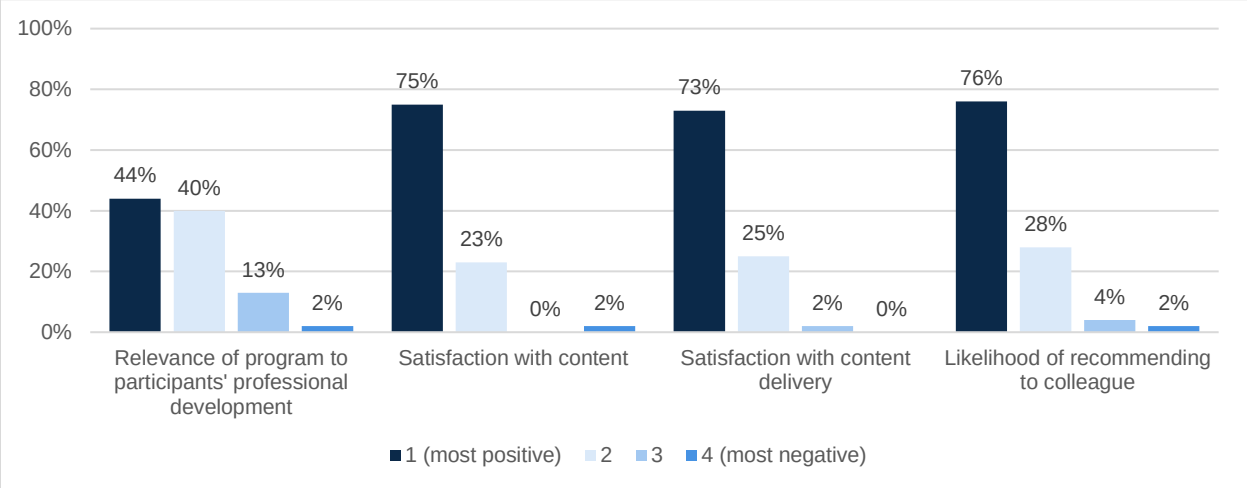
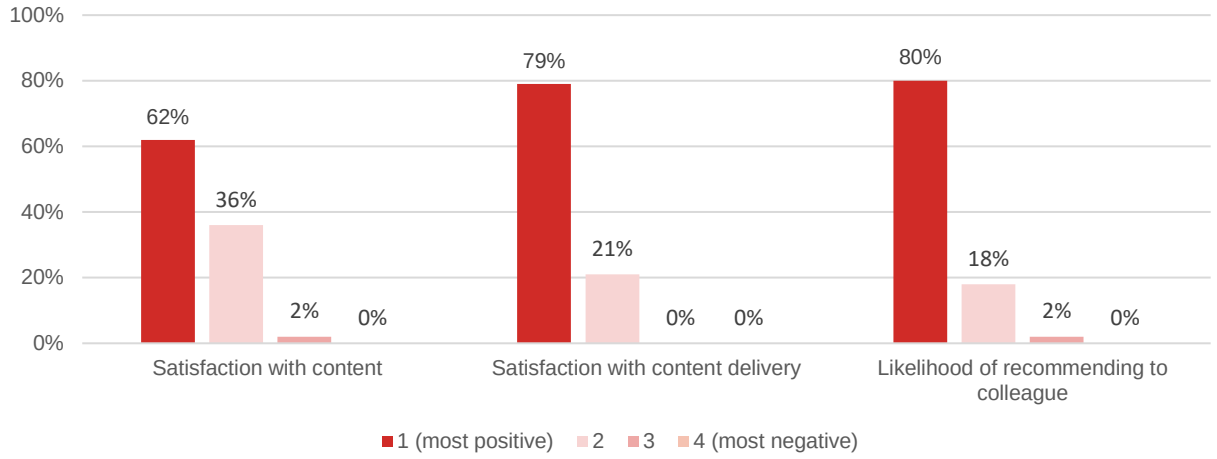


Figure A.2. Overall satisfaction with content and delivery of the KLDF program



Appendix A.3. Details on respondents' views on Successor Prep program effectiveness

We asked Successor Prep principals to rate the effectiveness of the program on building skills in each of three program domains. Tables A.1-A.3 provide the average response for each item within a program domain. Figures A.3-A.5 provide detail on how often each response option (from 1 to 4) was selected for these same survey items.

We also asked Successor Prep principals to rate how important each of the skills in these domains was to success as a school principal. These responses are displayed in Figures A.6-A.8.

Finally, we asked about the degree to which Successor Prep participation increased the participants' abilities to support teacher practice and improvements in student outcomes. These responses are displayed in Figure A.9.

Table A.1. Successor Prep program average effectiveness at building outcome management skills

	Number of responses	Mean
How effective was the program in helping you… [1 = very effective, 2 = effective, 3 = somewhat effective, and 4 = not at all effective]		
Provide continuous instructional performance support	46	2.0
Provide data-driven leadership	45	2.0
Monitor school performance metrics	45	1.9
Engage in continuous learning	46	1.8
Collect data for monitoring school performance	43	1.8

Table A.2. Successor Prep program average effectiveness at building strategic planning skills

	Number of responses	Mean
How effective was the program in helping you… [1 = very effective, 2 = effective, 3 = somewhat effective, and 4 = not at all effective]		
Set and pursue operational goals	46	2.3
Support staff learning and growth	46	2.1
Establish decision-making processes	46	2.0
Set and pursue instructional goals	46	1.9
Plan, execute, and commit	46	1.8
Provide mission-driven leadership	46	1.7
Provide visionary leadership	46	1.6

Table A.3. Successor Prep program average effectiveness at building equity leadership skills

	Number of responses	Mean
How effective was the program in helping you… [1 = very effective, 2 = effective, 3 = somewhat effective, and 4 = not at all effective]		
Strive to dismantle systemic inequities	46	2.2
Provide equity leadership	46	2.1
Identify guiding principles for culture systems	46	2.1
Encourage constructive dialogue	46	2.0
Set direction and model expectations	46	1.9
Demonstrate cultural competence	46	1.9
Encourage team leadership	46	1.8

Figure A.3. Successor Prep program effectiveness at building outcome management skills, by response option

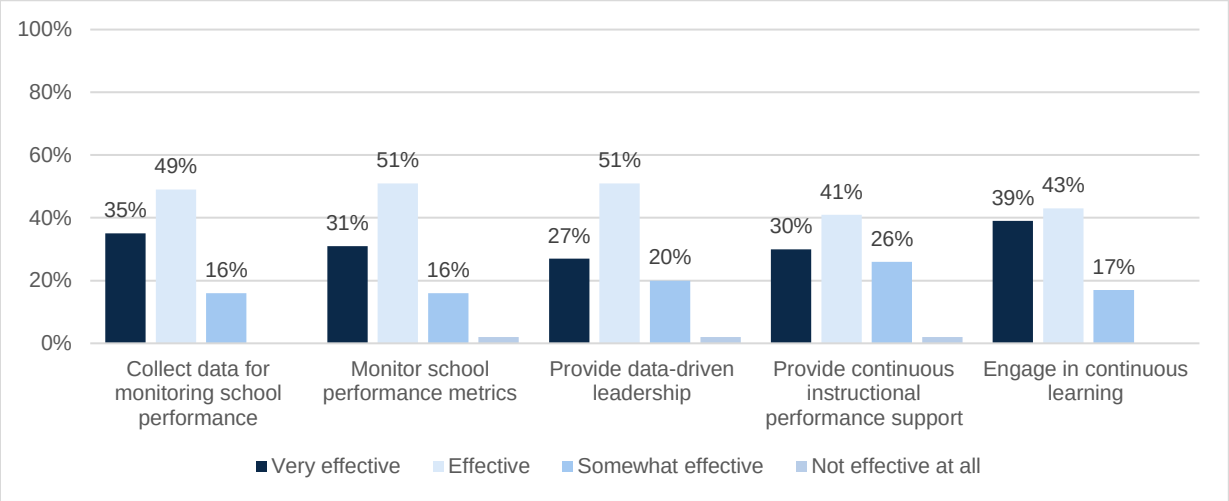


Figure A.4. Successor Prep program effectiveness at building strategic planning skills, by response option

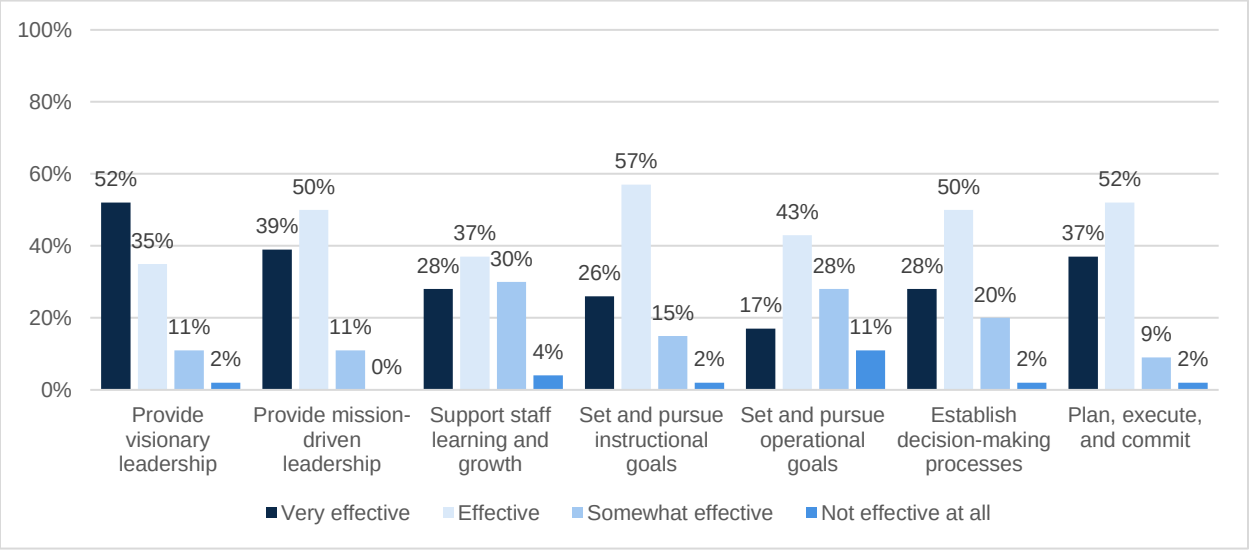


Figure A.5. Successor Prep program effectiveness at building equity leadership skills, by response option

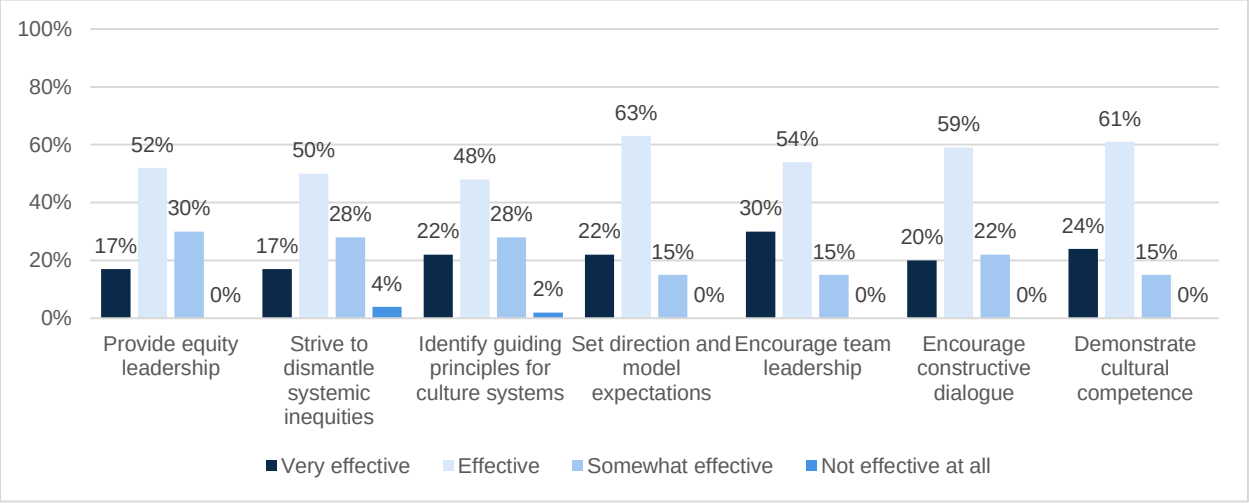


Figure A.6. Importance of outcome management skills to success according to Successor Prep participants, by response option

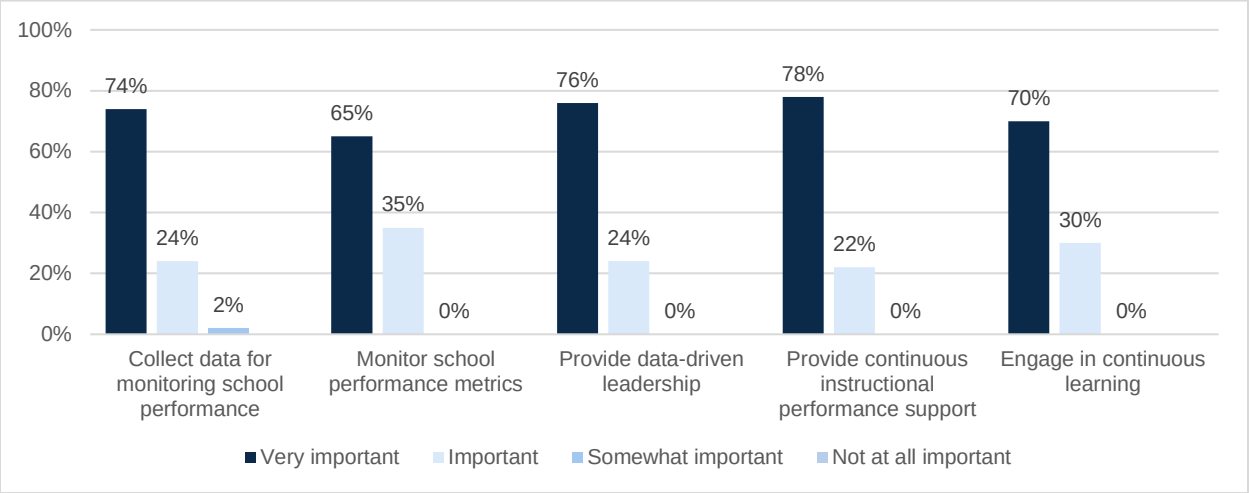


Figure A.7. Importance of strategic planning skills to success according to Successor Prep participants, by response option

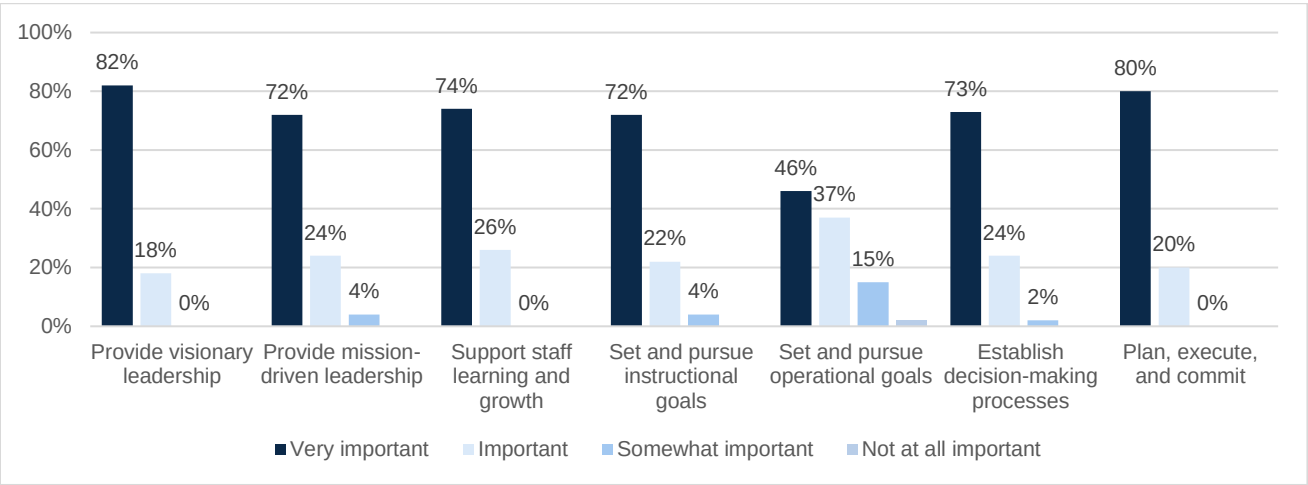


Figure A.8. Importance of equity leadership skills to success according to Successor Prep participants, by response option

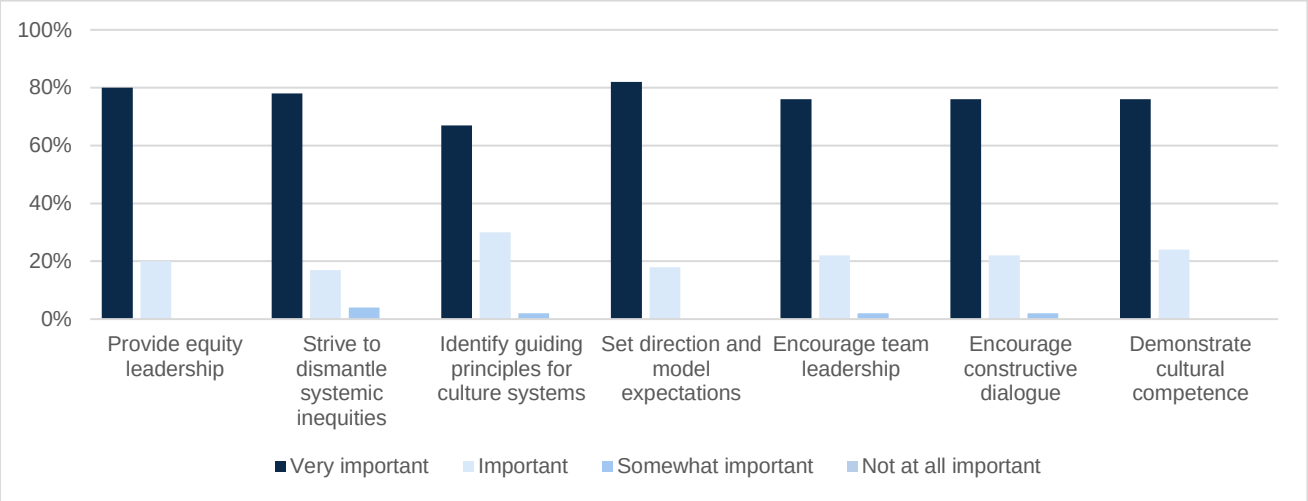
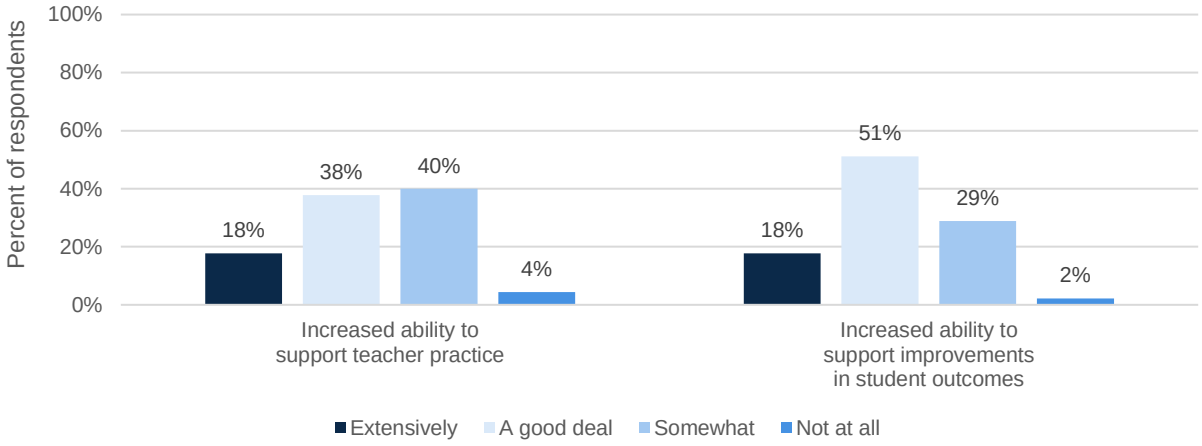


Figure A.9. Degree to which Successor Prep participation increased abilities to support teacher practice and improvements in student outcomes



Appendix A.4: Details on respondents' views on KLDF program effectiveness

We asked KLDF participants to rate the effectiveness of the program on building skills in each of two program domains. Tables A.4 and A.5 provide the average response for each item within a program domain. Figures A.10 and A.11 provide detail on how often each response option (from 1 to 4) was selected for these same survey items.

We also asked KLDF participants to rate how often they focus on building skills from each program domain in potential or current leaders. These responses are displayed in Figures A.12 and A.13.

Table A.4. KLDF program average effectiveness at building talent development skills of others:

	Number of responses	Mean
How effective was the KIPP Leadership Design Fellowship in building your capacity to develop the following skills of potential or current leaders … [1 = very effective, 2 = effective, 3 = somewhat effective, and 4 = not at all effective]		
Recruit and select highly effective staff	48	2.1
Provide on-the-job development	44	2.3
Retain talented and valued employees	42	2.4
Manage staff who do not meet expectations or fit the organizational culture	41	2.6
Provide coaching and one-on-one support to staff	46	2.2
Establish standards for effective leadership	52	1.8
Develop leaders and prepare successors	51	2.0
Currently, how often do you focus on building the following skills of potential or current leaders ... [1=Very often, 2 = Often, 3 = Sometimes, 4=Never]		
Recruit and select highly effective staff	55	2.1
Provide on-the-job development	55	1.9
Retain talented and valued employees	55	2.2
Manage staff who do not meet expectations or fit the organizational culture	55	2.4
Provide coaching and one-on-one support to staff	55	1.7
Establish standards for effective leadership	55	2.1
Develop leaders and prepare successors	55	2.3

Table A.5. KLDF program average effectiveness at building equity leadership skills of others:

	Number of responses	Mean
How effective was the KIPP Leadership Design Fellowship in building your capacity to develop the following skills of potential or current leaders … [1 = very effective, 2 = effective, 3 = somewhat effective, and 4 = not at all effective]		
Provide equity leadership	45	2.2
Strive to dismantle systemic inequities	44	2.5
Identify equitable organizational policies	45	2.4
Demonstrate cultural competence	46	2.3
Currently, how often do you focus on building the following skills of potential or current leaders... [1=Very often, 2 = Often, 3 = Sometimes, 4=Never]		
Provide equity leadership	55	2.0
Strive to dismantle systemic inequities	55	2.2
Identify equitable organizational policies	55	2.2
Demonstrate cultural competence	54	1.9

Figure A.10. KLDF program effectiveness at building capacity to develop talent development skills of others, by response option

Figure A.11. KLDF program effectiveness at building capacity to develop equity leadership skills of others, by response option

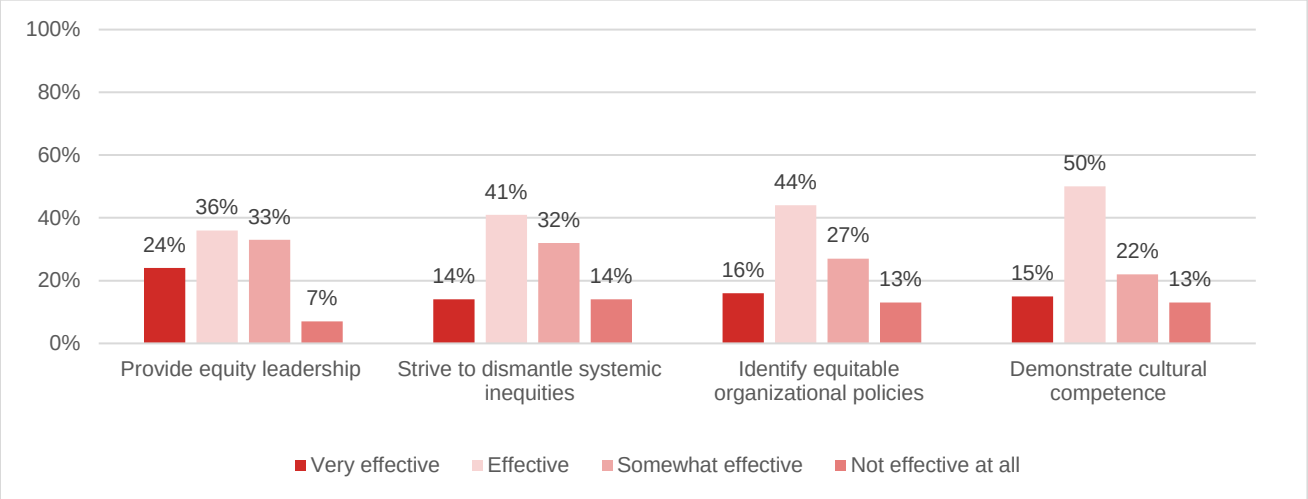


Figure A.12. Frequency KLDF participants focus on building talent development skills of others in their current role, by response option

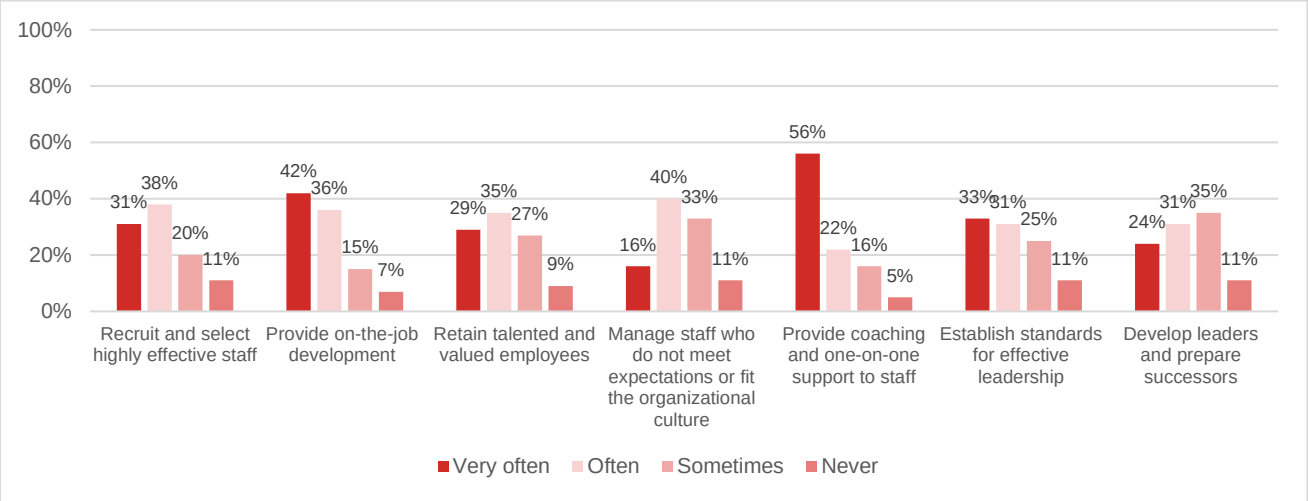
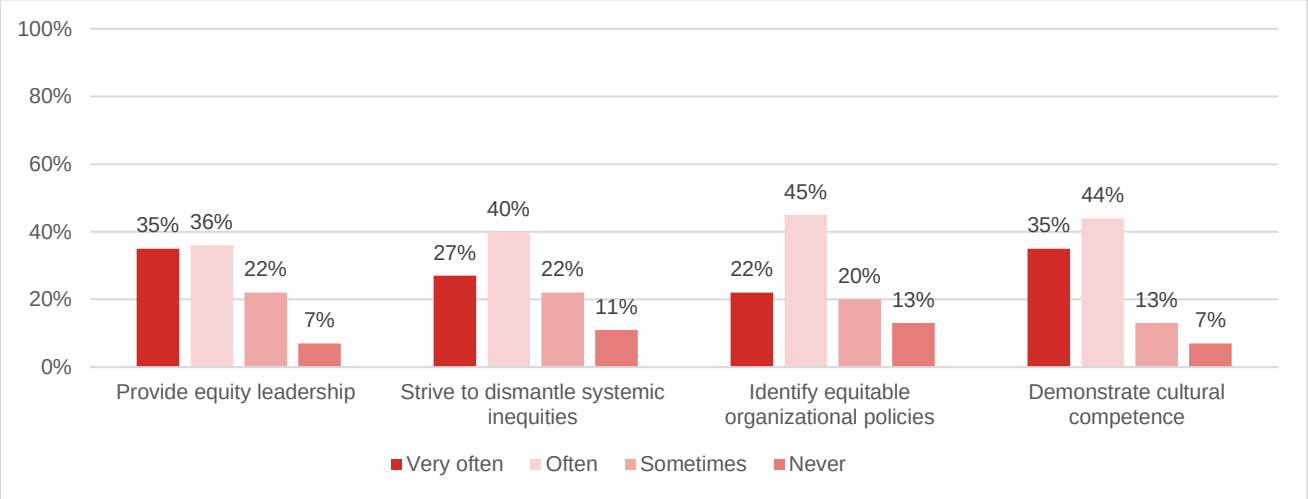


Figure A.13. Frequency KLDF participants focus on building equity leadership skill of others in their current role, by response option



Appendix B.
Additional details about the analysis of outcomes in Successor Prep schools

This page has been left blank for double-sided copying.

This appendix includes additional description of the study sample for the analysis of student and teacher outcomes in Successor Prep schools and technical details about the analysis methods.

A. Additional details on the study samples

1. Data on all KIPP schools during the study time period

Across the study time period, the number of KIPP schools grew substantially. Table B.1 reports school-level means and standard deviations by year for the outcomes in our analyses. Across years, the mean MAP scores differ from 0 and the MAP score standard deviations differ from 1 because we standardized aggregated scores within grade, subject, and year, but we averaged these data across grades to the school level for analysis. Because we standardized the MAP test scores within each year, the means do not measure changes in the test scores over time. Across the study period, student retention rates decreased by about 2 percentage points and teacher retention rates decreased by 3 or 4 percentage points.

Table B.1. Outcomes for all KIPP schools by school year

Outcome and school year	Number of schools	Mean	Standard deviation
Math MAP scores			
2011–2012	81	-0.005	0.838
2012–2013	105	-0.007	0.864
2013–2014	120	-0.014	0.899
2014–2015	139	-0.011	0.928
2015–2016	158	-0.035	0.942
2016–2017	171	0.003	0.943
2017–2018	181	0.001	0.929
2018–2019	178	0.032	0.936
Reading MAP scores			
2010–2011	64	0.047	0.909
2011–2012	81	-0.016	0.839
2012–2013	105	-0.008	0.920
2013–2014	120	0.000	0.913
2014–2015	139	0.000	0.913
2015–2016	158	-0.028	0.932
2016–2017	172	-0.001	0.922
2017–2018	181	-0.006	0.935
2018–2019	180	0.025	0.935
Percentage of students remaining at the school the following year			
2010–2011	81	87.6	8.3
2011–2012	88	87.1	7.3
2012–2013	104	87.7	7.6
2013–2014	115	87.9	7.8
2014–2015	136	87.3	8.4
2015–2016	154	86.7	7.8
2016–2017	163	85.5	8.5

Outcome and school year	Number of schools	Mean	Standard deviation
2017–2018	167	85.6	8.3
2018–2019	168	85.8	8.8
Percentage of teachers still teaching at the school the following year			
2010–2011	81	68.5	14.9
2011–2012	88	68.4	15.9
2012–2013	104	68.0	15.2
2013–2014	115	68.4	14.0
2014–2015	136	64.4	14.4
2015–2016	154	64.7	15.2
2016–2017	163	63.9	14.5
2017–2018	167	62.5	13.8
2018–2019	165	64.5	13.3
Percentage of teachers in any position in any KIPP school the following year			
2010–2011	81	74.4	15.0
2011–2012	88	74.4	14.5
2012–2013	104	75.5	14.0
2013–2014	115	75.2	13.5
2014–2015	136	72.7	13.6
2015–2016	154	72.4	14.0
2016–2017	163	71.7	14.0
2017–2018	167	71.1	13.9
2018–2019	165	70.8	11.9

Source: Administrative data from KIPP.

Note: The means and standard deviations for the test scores are measured in units of school-level standard deviations. For other measures, the standard deviations are calculated based on data aggregated to the school level.

2. Successor Prep principals' prior experience as school leaders

In general, participants of the Successor Prep program are expected to become a principal of their Successor school in the fall following the start of the program. While most of the 82 Successor Prep participants who led a KIPP school—the sample for research question 1—became a principal of their school for the first time on the expected Successor Prep program timeline, some had begun leading their schools up to three years before they began the Successor Prep program (Table B.2).

Table B.2. Number of years Successor Prep principals had been leading their schools before the expected first year as a Successor Prep principal

Years of prior experience leading their school	Number of leaders
Zero years	63
One year	17
Two years	1
Three years	1

Source: Administrative data from KIPP.

3. Unweighted sample characteristics for the Successor Prep and comparison group schools

As shown in Table B.3, without the propensity score weights, the schools led by Successor Prep principals included in the analysis of the principal retention outcome (research question 2) were generally similar to those led by the comparison group principals. This may be because both groups of principals in this analysis were new to their schools. However, there were substantively large differences in the baseline outcomes for the Successor Prep and comparison schools in the analytic samples for each of the outcome measures used in research question 3.

Table B.3. Unweighted baseline data for each analysis sample

	Successor Prep			Comparison group			
Sample and baseline measure	Number of observations	Mean	Standard deviation	Number of observations[a]	Mean	Standard deviation	Effect size
Principal retention analytic sample (research question 2)							
Math MAP scores, 3 years prior to a transition	44	-0.105	0.794	75	0.018	0.829	-0.15
Math MAP scores, 2 years prior to a transition	44	-0.067	0.867	75	-0.025	0.819	-0.05
Math MAP scores, 1 year prior to a transition	44	0.021	0.960	75	-0.111	0.804	0.15
Reading MAP scores, 3 years prior to a transition	44	-0.111	0.814	75	-0.061	0.865	-0.06
Reading MAP scores, 2 years prior to a transition	44	-0.062	0.919	75	-0.068	0.870	0.01
Reading MAP scores, 1 year prior to a transition	44	0.003	0.953	75	-0.117	0.836	0.14
Middle school	44	59.1	--	75	64.0	--	-0.10
Math test score analytic sample							
Math MAP scores, 3 years prior to a transition	40	-0.085	0.781	68	0.274	0.773	-0.46
Math MAP scores, 2 years prior to a transition	40	-0.012	0.849	68	0.330	0.759	-0.43
Math MAP scores, 1 year prior to a transition	40	0.067	0.959	68	0.368	0.794	-0.35
Math MAP scores, end of transition year	40	0.051	0.979	68	0.395	0.870	-0.37
Math MAP scores, 2 years after a transition	40	0.144	0.929	68	0.446	0.905	-0.33
Middle school	40	60.0	--	68	60.3	--	-0.01
Reading test score analytic sample							
Reading MAP scores, 4 years prior to a transition	36	0.098	0.836	53	0.192	0.850	-0.11
Reading MAP scores, 3 years prior to a transition	36	-0.115	0.832	53	0.085	0.877	-0.23
Reading MAP scores, 2 years prior to a transition	36	-0.035	0.919	53	0.121	0.823	-0.18
Reading MAP scores, 1 year prior to a transition	36	0.015	0.951	53	0.225	0.807	-0.24
Reading MAP scores, end of transition year	36	-0.065	1.066	53	0.257	0.875	-0.33
Reading MAP scores, 2 years after a transition	36	0.108	0.946	53	0.291	0.950	-0.19

Sample and baseline measure	Successor Prep			Comparison group			
	Number of observations	Mean	Standard deviation	Number of observations[a]	Mean	Standard deviation	Effect size
Middle school	36	63.9	--	53	66.0	--	-0.04
Student and teacher retention analytic sample							
Percentage of students remaining at the school the following year, 4 years prior to a transition	33	87.6	7.7	50	89.6	6.7	-0.28
Percentage of students remaining at the school the following year, 3 years prior to a transition	33	87.4	8.1	50	90.2	6.5	-0.39
Percentage of students remaining at the school the following year, 2 years prior to a transition	33	88.0	8.3	50	89.9	6.3	-0.27
Percentage of students remaining at the school the following year, 1 year prior to a transition	33	87.8	6.8	50	90.3	5.6	-0.40
Percentage of students remaining at the school the following year, end of transition year	33	87.9	7.8	50	89.7	5.8	-0.27
Percentage of students remaining at the school the following year, 2 years after a transition	33	87.5	7.5	50	89.4	5.8	-0.29
Percentage of teachers still teaching at the school the following year, 4 years prior to a transition	33	65.8	15.1	50	72.4	13.7	-0.46
Percentage of teachers still teaching at the school the following year, 3 years prior to a transition	33	70.5	12.4	50	70.0	12.3	0.04
Percentage of teachers still teaching at the school the following year, 2 years prior to a transition	33	64.8	13.5	50	64.7	12.2	0.01
Percentage of teachers still teaching at the school the following year, 1 year prior to a transition	33	61.7	11.8	50	66.7	12.9	-0.40
Percentage of teachers still teaching at the school the following year, end of transition year	33	63.7	13.3	50	65.9	12.6	-0.17
Percentage of teachers still teaching at the school the following year, 2 years after a transition	33	60.8	16.2	50	67.5	15.4	-0.42
Percentage of teachers in any position in any KIPP school the following year, 4 years prior to a transition	33	72.7	15.7	50	78.9	11.9	-0.45
Percentage of teachers in any position in any KIPP school the following year, 3 years prior to a transition	33	77.4	10.2	50	76.5	10.1	0.09
Percentage of teachers in any position in any KIPP school the following year, 2 years prior to a transition	33	72.9	12.5	50	73.0	12.2	-0.02
Percentage of teachers in any position in any KIPP school the following year, 1 years prior to a transition	33	70.7	13.1	50	73.5	12.5	-0.22
Percentage of teachers in any position in any KIPP school the following year, end of transition year	33	71.3	12.2	50	73.8	11.6	-0.21

Sample and baseline measure	Successor Prep			Comparison group			
	Number of observations	Mean	Standard deviation	Number of observations[a]	Mean	Standard deviation	Effect size
Percentage of teachers in any position in any KIPP school the following year, 2 years after a transition	33	70.0	17.2	50	73.9	14.0	-0.26
Math MAP scores, 3 years prior to a transition	33	-0.041	0.817	50	0.228	0.820	-0.33
Math MAP scores, 2 years prior to a transition	33	-0.012	0.834	50	0.265	0.809	-0.33
Math MAP scores, 1 year prior to a transition	33	0.053	0.909	50	0.335	0.861	-0.32
Reading MAP scores, 4 years prior to a transition	33	0.131	0.851	50	0.160	0.843	-0.03
Reading MAP scores, 3 years prior to a transition	33	-0.109	0.864	50	0.072	0.900	-0.20
Reading MAP scores, 2 years prior to a transition	33	-0.057	0.945	50	0.116	0.842	-0.19
Reading MAP scores, 1 year prior to a transition	33	-0.001	0.956	50	0.242	0.821	-0.28
Middle school	33	66.7	--	50	64.0	--	0.05

Source: Administrative data from KIPP.

Note: The number of observations reflects the number of principals for research question 2, and schools for the other outcomes. The means and standard deviations for the test scores are measured in units of school-level standard deviations. Whether a school is a middle school is a dichotomous measure, so the standard deviations are not reported. Effect sizes are Hedges' *g* effect sizes representing school-level standard deviations with an adjustment for small sample sizes (WWC 2022), including for the dichotomous middle school measure.

[a] For the comparison condition, the 75 observations used in research question 2 reflect 74 distinct principals who led 49 distinct schools, including schools led by more than one principal in a given year or that had multiple leadership transitions during this time period. For the other outcomes, the same comparison school can be counted multiple times because it can appear in the comparison group for multiple Successor Prep cohorts. For example, the number of unique comparison schools for the analysis of math scores is just 42.

After applying the propensity score weights, the differences in the baseline outcomes for Successor Prep and comparison group schools for each outcome's analytic sample were generally small (Table B.4). All but one of the baseline outcome differences were within 0.25 standard deviations (last column of Table B.4) and many were within 0.05 standard deviations. The one baseline difference that exceeded 0.25 standard deviations in absolute value is the percentage of teachers still teaching at the school the following year, measured 2 years prior to a transition. Although the WWC requires that baseline differences only from the immediate prior year to the intervention are within 0.25 standard deviations, this baseline difference is suggestive that the sample of intervention and comparison group schools for this outcome may be dissimilar, so we suggest some caution in interpreting the findings in Table IV.6 for this outcome. In particular the baseline effect size in Table B.4 of -0.37 standard deviations suggests that the estimated impact of -0.17 standard deviations 2 years after the transition might be too negative because the Successor Prep schools had lower rates of teacher retention 2 years before the transition.

The table also reports the weighted outcome means for the analytic samples used in research question 3 in the first and second years the Successor Prep principals led their schools. The means reported in Table B.4 differ from the findings in Table IV.6 because they are not adjusted for the covariates included in the regression model.

Table B.4. Propensity score weighted baseline and outcome data for each analysis sample

	Successor Prep			Comparison group			
Sample and baseline measure	Number of observations	Mean	Standard deviation	Number of observations[a]	Mean	Standard deviation	Effect size
Principal retention analytic sample (research question 2)							
Math MAP scores, 3 years prior to a transition	44	-0.072	0.794	75	-0.047	0.829	-0.03
Math MAP scores, 2 years prior to a transition	44	-0.115	0.867	75	-0.076	0.819	-0.05
Math MAP scores, 1 year prior to a transition	44	-0.122	0.960	75	-0.102	0.804	-0.02
Reading MAP scores, 3 years prior to a transition	44	-0.104	0.814	75	-0.097	0.865	-0.01
Reading MAP scores, 2 years prior to a transition	44	-0.109	0.919	75	-0.096	0.870	-0.01
Reading MAP scores, 1 year prior to a transition	44	-0.119	0.953	75	-0.111	0.836	-0.01
Middle school	44	63.5	--	75	63.9	--	-0.01
Math test score analytic sample							
Math MAP scores, 3 years prior to a transition	40	0.101	0.781	68	0.064	0.773	0.05
Math MAP scores, 2 years prior to a transition	40	0.178	0.849	68	0.140	0.759	0.05
Math MAP scores, 1 year prior to a transition	40	0.223	0.959	68	0.186	0.794	0.04
Math MAP scores, end of transition year	40	0.199	0.979	68	0.196	0.870	0.00
Math MAP scores, 2 years after a transition	40	0.318	0.929	68	0.247	0.905	0.08
Middle school	40	59.6	--	68	59.8	--	0.00
Reading test score analytic sample							
Reading MAP scores, 4 years prior to a transition	36	0.146	0.836	53	0.142	0.850	0.00
Reading MAP scores, 3 years prior to a transition	36	-0.048	0.832	53	-0.028	0.877	-0.02
Reading MAP scores, 2 years prior to a transition	36	0.004	0.919	53	0.014	0.823	-0.01
Reading MAP scores, 1 year prior to a transition	36	0.089	0.951	53	0.109	0.807	-0.02
Reading MAP scores, end of transition year	36	0.031	1.066	53	0.148	0.875	-0.12
Reading MAP scores, 2 years after a transition	36	0.218	0.946	53	0.188	0.950	0.03
Middle school	36	66.9	--	53	66.2	--	0.02
Student retention analytic sample							
Percentage of students remaining at the school the following year, 4 years prior to a transition	33	87.2	7.7	50	88.4	6.7	-0.17
Percentage of students remaining at the school the following year, 3 years prior to a transition	33	87.9	8.1	50	89.3	6.5	-0.19
Percentage of students remaining at the school the following year, 2 years prior to a transition	33	88.1	8.3	50	89.2	6.3	-0.16

Sample and baseline measure	Successor Prep			Comparison group			
	Number of observations	Mean	Standard deviation	Number of observations[a]	Mean	Standard deviation	Effect size
Percentage of students remaining at the school the following year, 1 year prior to a transition	33	88.5	6.8	50	89.3	5.6	-0.12
Percentage of students remaining at the school the following year, end of transition year	33	88.6	7.8	50	89.1	5.8	-0.08
Percentage of students remaining at the school the following year, 2 years after a transition	33	88.4	7.5	50	88.3	5.8	0.01
Math MAP scores, 3 years prior to a transition	33	0.003	0.817	50	0.074	0.820	-0.09
Math MAP scores, 2 years prior to a transition	33	0.054	0.834	50	0.116	0.809	-0.07
Math MAP scores, 1 year prior to a transition	33	0.139	0.909	50	0.177	0.861	-0.04
Reading MAP scores, 4 years prior to a transition	33	0.135	0.851	50	0.130	0.843	0.01
Reading MAP scores, 3 years prior to a transition	33	-0.106	0.864	50	-0.038	0.900	-0.08
Reading MAP scores, 2 years prior to a transition	33	-0.028	0.945	50	0.007	0.842	-0.04
Reading MAP scores, 1 year prior to a transition	33	0.080	0.956	50	0.108	0.821	-0.03
Middle school	33	64.0	--	50	66.3	--	-0.05
Teacher retention in the same school analytic sample							
Percentage of teachers still teaching at the school the following year, 4 years prior to a transition	33	66.5	15.1	50	69.7	13.7	-0.22
Percentage of teachers still teaching at the school the following year, 3 years prior to a transition	33	69.6	12.4	50	70.4	12.3	-0.07
Percentage of teachers still teaching at the school the following year, 2 years prior to a transition	33	58.6	13.5	50	63.4	12.2	-0.37
Percentage of teachers still teaching at the school the following year, 1 year prior to a transition	33	63.8	11.8	50	65.5	12.9	-0.14
Percentage of teachers still teaching at the school the following year, end of transition year	33	63.7	13.3	50	64.8	12.6	-0.08
Percentage of teachers still teaching at the school the following year, 2 years after a transition	33	62.6	16.2	50	67.0	15.4	-0.28
Math MAP scores, 3 years prior to a transition	33	-0.044	0.817	50	0.055	0.820	-0.12
Math MAP scores, 2 years prior to a transition	33	0.013	0.834	50	0.084	0.809	-0.09
Math MAP scores, 1 year prior to a transition	33	0.086	0.909	50	0.142	0.861	-0.06
Reading MAP scores, 4 years prior to a transition	33	0.087	0.851	50	0.089	0.843	0.00
Reading MAP scores, 3 years prior to a transition	33	-0.076	0.864	50	-0.029	0.900	-0.05

Sample and baseline measure	Successor Prep			Comparison group			
	Number of observations	Mean	Standard deviation	Number of observations[a]	Mean	Standard deviation	Effect size
Reading MAP scores, 2 years prior to a transition	33	-0.037	0.945	50	-0.032	0.842	-0.01
Reading MAP scores, 1 year prior to a transition	33	0.026	0.956	50	0.060	0.821	-0.04
Middle school	33	66.8	--	50	61.3	--	0.11
Teacher retention in any KIPP school analytic sample							
Percentage of teachers in any position in any KIPP school the following year, 4 years prior to a transition	33	75.0	15.7	50	76.7	11.9	-0.13
Percentage of teachers in any position in any KIPP school the following year, 3 years prior to a transition	33	76.9	10.2	50	76.9	10.1	0.00
Percentage of teachers in any position in any KIPP school the following year, 2 years prior to a transition	33	70.6	12.5	50	72.5	12.2	-0.15
Percentage of teachers in any position in any KIPP school the following year, 1 years prior to a transition	33	72.1	13.1	50	74.1	12.5	-0.16
Percentage of teachers in any position in any KIPP school the following year, end of transition year	33	72.6	12.2	50	73.4	11.6	-0.07
Percentage of teachers in any position in any KIPP school the following year, 2 years after a transition	33	72.9	17.2	50	73.7	14.0	-0.05
Math MAP scores, 3 years prior to a transition	33	0.041	0.817	50	0.065	0.820	-0.03
Math MAP scores, 2 years prior to a transition	33	0.082	0.834	50	0.090	0.809	-0.01
Math MAP scores, 1 year prior to a transition	33	0.171	0.909	50	0.159	0.861	0.01
Reading MAP scores, 4 years prior to a transition	33	0.148	0.851	50	0.107	0.843	0.05
Reading MAP scores, 3 years prior to a transition	33	-0.019	0.864	50	-0.034	0.900	0.02
Reading MAP scores, 2 years prior to a transition	33	0.041	0.945	50	-0.015	0.842	0.06
Reading MAP scores, 1 year prior to a transition	33	0.121	0.956	50	0.084	0.821	0.04
Middle school	33	61.5	--	50	60.8	--	0.01

Source: Administrative data from KIPP.

Note: The means shown in the table have been adjusted using the propensity score weights calculated for each analysis. Standard deviations are unweighted. The number of observations reflects the number of principals for research question 2, and schools for the other outcomes. The means and standard deviations for the test scores are measured in units of school-level standard deviations. Whether a school is a middle school is a dichotomous measure, so the standard deviations are not reported. Effect sizes are Hedges' *g* effect sizes representing school-level standard deviation units, with an adjustment for small sample sizes (WWC 2022), including for the dichotomous middle school measure.

[a] For the comparison condition, the 75 observations used in research question 2 reflect 74 distinct principals who led 49 distinct schools, including schools led by more than one principal in a given year or that had multiple leadership transitions during this time period. For the other outcomes, the same comparison school can be counted multiple times because it can appear in the comparison group for multiple Successor Prep cohorts. For example, the number of unique comparison schools for the analysis of math scores is just 42.

4. Criteria to identify Successor Prep leader placements and measure principal tenure

To determine which schools had been led by Successor Prep principals and measure all principals' leadership tenure, we needed to identify the principals of schools in each school year. We used additional criteria to determine which years principals led their schools. The school leader data included the start and end dates of each principal's tenure in each school they led, and not all principals were indicated as leading a school for complete school years. To avoid incorrectly assigning outcomes to principals who were present in their schools for only a small portion of a school year, we required that principals had led a school for at least three weeks that year to be counted as the school's leader that year, and if they led the school between three and nine weeks, that at least one of the weeks was in September or April so that the principals were in the school at the beginning or the end of the year.

5. Criteria to include a school's test scores in the analysis

We also included two criteria to include a school's test scores in the study. We received student-level test score data by year, grade, and subject, which we aggregated to the school-level by year and subject. Some schools had few students in a given year, grade, and subject. To limit the influence of measurement error from scores from so few students, we dropped a school's test scores in a grade and subject when fewer than 30 students from that school had test scores in that grade and subject that year. In addition, in order to accurately standardize the scores within year, grade, and subject, we dropped all scores from a test administration if there were not at least five schools with test scores in that year, grade, and subject.

B. Additional details on the study methods

1. Research question 2

We used weighted least squares regressions to examine whether Successor Prep principals were more likely to remain in their placement schools for a second or third year than non-Successor Prep principals. Because some Successor Prep principals had previously led their placement schools before participating in the program, we included only Successor Prep principals who were new to leading their schools. We also included non-Successor Prep principals who were new to leading their schools in one of the same years that a cohort of Successor Prep principals first took over leadership of their schools. We assigned the non-Successor Prep principals to the cohorts that matched the year they started leading their schools. For example, we assigned non-Successor Prep principals who first took over leadership of their schools in the 2014–2015 school year to the January 2014 cohort.

In order to allow us to control for the possible impact of school performance on principal retention, we included only the Successor Prep and non-Successor Prep principals who led schools with test scores in each of the four years before they took over leadership of the schools, or three years for the math MAP score outcome. This requirement does prevent us from including some newer KIPP schools in the analysis that would have fewer than four years of prior scores. However, prior research on principal transitions (particularly, Miller 2013) has shown that it is potentially important to measure prior trends in student achievement over multiple years. This is because schools that experience principal transitions might experience a decline in student achievement *before* the transition occurs. This might happen if KIPP sometimes decides to replace a principal because the current principal is struggling.[1] If so, findings

[1] This is an example of an "Ashenfelter Dip," which referred originally to falsely attributing wage gains to a training program that may only have returned participants to the wage rate they would have obtained without the program.

that adjust for only the immediate prior year of test scores would be biased. Including multiple years of prior test score data allows us to assess whether there is a pre-transition decline in test scores and perform a more accurate adjustment for differences in test scores.

To achieve a comparison group of principals in schools with similar characteristics to the Successor Prep schools, we calculated propensity score weights to account for prior-year outcomes, cohort, and a middle school indicator. To calculate the propensity score weights, we estimated probit models of a Successor Prep school indicator on prior-year outcomes, a set of cohort indicators, and a middle school indicator. We estimated the propensity scores separately by outcome analytic sample. We included a middle school indicator and the cohort indicators for the January 2015, January 2016, January 2017, and January 2018 cohorts. We also included the three prior years of MAP reading and math scores.

We used the following probit model to calculate the propensity scores:

$$SP_{ijt} = \alpha + \beta X_{jt} + \gamma_{jt} + \mu_{jt} + \epsilon_{ijt}$$

where SP_{ijt} is an indicator for whether principal *i* in school *j* and cohort *t* participated in the Successor Prep program, X_{jt} is the set of variables for school *j* in cohort *t's* baseline test scores, γ_{jt} is a middle school indicator, μ_{jt} is the set of four cohort indicators, and ϵ_{ijt} is a random error term that reflects the influence of unobserved factors on being a Successor Prep principal.

We used the following formula to calculate the propensity score weights (*w*), where *p* is each school's propensity to be a Successor Prep school based on the probit model:

For Successor Prep schools: $w(x) = \frac{1}{p(x)}$

For comparison schools: $w(x) = \frac{1}{1 - p(x)}$

Using these weights, principals of Successor Prep schools that have characteristics that are more similar to the schools led by comparison group principals contribute more to the analysis. Similarly, principals of comparison schools that have characteristics that are more similar to Successor Prep schools contribute more to the analysis. We then used weighted least squares regressions to estimate whether the principals in Successor Prep schools continued to lead their schools longer than the comparison principals. We weighted each regression using the propensity score weights. We also adjusted the standard errors to account for clusters by school because some schools appear in the comparison group multiple times.

To conduct the analysis, we used the following regression model:

$$y_{ijt} = \alpha + \delta SP_{ijt} + \beta X_{ijt} + \gamma_{ijt} + \mu_{ijt} + \vartheta_{ijt}$$

where y_{ijt} is one of two binary outcome variables, one for whether principal *i* in school *j* and cohort *t* was still leading school *j* for a second year, and one for whether principal *i* in school *j* in cohort *t* was still leading school *j* for a third year. SP_{ijt} is an indicator for whether principal *i* in school *j* and cohort *t* participated in the Successor Prep program. X_{ijt} is a set of baseline outcome variables for school *j*'s MAP reading and math test scores in each of the three years before cohort *t*'s principals first took over leadership of their schools. γ_{ijt} is a middle school indicator. μ_{ijt} is a set of four cohort indicators for the

The apparent gains arose because workers who had experienced a dip in wages were the ones who chose to participate in the program (Ashenfelter 1978; Jacobson, LaLonde, and Sullivan 1993).

January 2015, January 2016, January 2017, and January 2018 cohorts. ϑ_{ijt} is a random error term that reflects the influence of unobserved factors on the outcome. Finally, $\hat{\delta}$ estimates the difference in the retention rates between the Successor Prep and non-Successor Prep principals.

2. Research question 3

We also used weighted least squares regressions to examine whether the Successor Prep program resulted in better student and teacher outcomes after the first two years of the transitions than might have occurred had their never been principal transitions. The regressions compared outcomes between Successor Prep principals' placement schools and a comparison group of schools that never received a Successor Prep principal. For each Successor Prep cohort, we identified a comparison group of schools that never received Successor Prep principals and that had outcome data in the several years before and after the Successor Prep principals in that cohort first took over their schools. Because comparison schools only needed to have outcome data in the several years before and after each cohort took over their schools, the same schools could be in the comparison group for multiple cohorts.

To achieve comparison groups that were similar to the Successor Prep schools, we calculated propensity score weights to account for prior-year outcomes, KIPP region, cohort, and a middle school indicator. To calculate the propensity score weights, we estimated probit models of a Successor Prep school indicator on prior-year outcomes, a set of region indicators, a set of cohort indicators, and a middle school indicator. We estimated the propensity scores separately by outcome analytic sample. For each outcome analytic sample, we included a set of region indicators for each of the five most common KIPP regions in the study data (Austin, DC, Houston, New Orleans, and New York City). The remaining KIPP regions were grouped together in the excluded category for the indicator. We also included indicators for the January 2015, January 2016, and January 2017 cohorts, and a middle school indicator. For the reading test score analytic sample, we also included school-level MAP reading scores in each of the four prior years. For the math test score analytic sample, we also included school-level MAP math scores in each of the three prior years. We did not include a fourth prior year of MAP math scores because the data were unavailable in the 2011–2012 school year, which would have prevented us from including the January 2014 cohort. For each of the three retention outcome analytic samples, we also included the four prior years of MAP reading scores and the three prior years of MAP math scores, as well as four prior years of that retention rate outcome.

We used the following probit model to calculate the propensity scores:

$$SP_{jt} = \alpha + \beta X_{jt} + \gamma_{jt} + \mu_{jt} + \theta_j + \epsilon_{jt}$$

where SP_{jt} is an indicator for school *j* in cohort *t* being a Successor Prep school, X_{jt} is the set of variables for school *j* in cohort *t's* baseline outcomes and test scores, γ_{jt} is a middle school indicator, μ_{jt} is the set of four cohort indicators, θ_j is a set of region indicators. ϵ_{jt} is a random error term that reflects the influence of unobserved factors on being a Successor Prep school. We then used the same formulas to calculate the propensity score weights that we used in research question 2.

We used these weights to estimate weighted least squares regressions to understand whether the Successor Prep program resulted in better test scores and student and teacher retention rates after each of the first two years of the transitions than might have occurred had their never been principal transitions. We ran separate regressions to estimate the impacts on each outcome after the first year and after the

second year. We also adjusted the standard errors to account for clusters by school because some schools appear in the comparison group for multiple cohorts.

To estimate the impacts, we used the following regression model:

$$y_{jty} = \alpha + \delta SP_{jt} + \beta X_{jt} + \gamma_{jt} + \mu_{jt} + \theta_j + \vartheta_{jt}$$

where y_{jty} is the test score or retention outcome for school *j* in cohort *t* in year in *y* (which represents whether the outcome was measured at the end of the first or second years) and most terms are the same as those from the probit model used to estimate propensity scores. ϑ_{ijt} is a random error term that reflects the influence of unobserved factors on the outcomes. The coefficient $\hat{\delta}$ estimates the difference in outcomes between the Successor Prep and non-Successor Prep schools.

C. Additional results

1. Research question 2

We tried an alternative analysis of whether Successor Prep principals stayed in their placement schools longer than non-Successor Prep principals and found results similar to those presented in the main body of the report. Our main analysis of whether Successor Prep principals stayed in their placement schools longer than non-Successor Prep principals included only principals who were new to leading their schools and allowed the principals to come from any KIPP region. As a result, some non-Successor Prep principals were from regions that did not have Successor Prep principals in the analysis. To examine whether this impacted the results, we conducted a similar analysis that restricted the non-Successor Prep principals to the same regions as the Successor Prep principals. However, because few non-Successor Prep principals in those same regions were new to leading their schools, we included all Successor Prep and non-Successor Prep principals in this alternative analysis regardless of whether they were new to leading their schools. Our alternative analysis found results similar to those presented in the main body of the report.

2. Research question 3

In Table IV.6, we reported effect sizes for the test score outcomes representing standard deviations of school-level achievement. To facilitate comparisons with other studies that report effect sizes for test scores using student-level standard deviations, we report both sets of effect sizes in Table B.5. We calculated the student-level effect sizes by standardizing the test scores across students within each grade, subject, and year before aggregating the test scores to the school level. Using these alternative test scores, we re-estimated the regression analyses and also calculated standard deviations using student-level data for the students contributing test scores to the schools included in the analysis. The student-level math test score samples include 15,535 student-year-cohort observations in Successor Prep schools and 27,951 student-year-cohort observations in comparison schools for the first year the Successor Prep principals led their schools, and 15,235 student-year-cohort observations in Successor Prep schools and 28,204 student-year-cohort observations in comparison schools for the second year the Successor Prep principals led their schools. The student-level reading test score samples include 14,171 student-year-cohort observations in Successor Prep schools and 21,631 student-year-cohort observations in comparison schools for the first year the Successor Prep principals led their schools, and 13,861 student-year-cohort observations in Successor Prep schools and 21,485 student-year-cohort observations in comparison schools for the second year the Successor Prep principals led their schools.

Table B.5. School- and student-level effect sizes for the test score outcomes

Outcome	First year the Successor Prep principals led their schools		Second year the Successor Prep principals led their schools	
	School-level effect size	Student-level effect size	School-level effect size	Student-level effect size
Math MAP scores	-0.04	-0.02	0.02	0.00
Reading MAP scores	-0.11	-0.05	0.02	0.00

Source: KIPP administrative data.

Note: Effect sizes are Hedges' *g* effect sizes representing school- or student- level standard deviation units with an adjustment for small sample sizes (WWC 2022).

* Significantly different from zero at the .10 level, two-tailed test.

** Significantly different from zero at the .05 level, two-tailed test.

To examine whether the variables we included in our regressions affected the results of the impact of the Successor Prep program on student and teacher outcomes, we tried alternative regressions for each outcome. As principal experience may impact student outcomes, we tried controlling for principals' prior tenure in their schools. Specifically, we controlled (linearly) for the number of years of prior experience leading that school across all principals in the school in the year the Successor Prep principals first took over leadership of their schools. For the Successor Prep schools this ranged from 0 to 3 years of prior experience in that school (median of 0 years) and for the comparison schools from 0 to 9 year of prior experience in that school (median of 4 years). We also estimated specifications that removed the propensity score weights and required only three years of prior outcomes to allow a slightly larger sample of principals for the reading and retention outcomes. None of these alternative regressions changed the results presented in the main body of the report. We also analyzed each cohort of Successor Prep principals separately, but these findings were too imprecise due to the small samples for each cohort.

This page has been left blank for double-sided copying.

Appendix C.
Additional findings from the analysis of the Fisher Fellowship selection instrument

This page has been left blank for double-sided copying.

This appendix includes additional detail on the findings from the analysis of the reliability of the Fisher Fellowship selection instrument. Tables C.1-C.3 report (1) the internal consistency of scores on each of the three competencies when excluding each individual item within a competency and (2) the interrater reliability of each item within each competency. The internal consistency results are called leave-out alphas and a number closer to 1 indicates that scores on items within the competency are more similar when excluding the item than if that item were included, suggesting that the item may be dissimilar from the others in the competency. Interrater reliability is measured as the interclass correlation coefficient and a number closer to 1 indicates that the raters gave candidates scores on the item that are more similar.

Table C.1. Leave-out internal consistency and interrater reliability for the Culture & Self-Awareness competency

Question	Item	Leave-out internal consistency (overall α = 0.85)	Interrater reliability
Why do you want to lead a KIPP school in XX community?	1a. Cultural Competence	0.84	0.69
	1b. Self-Awareness	0.84	0.59
What values guide your professional work?	2. Impact and Influence	0.84	0.64
What does it mean to you to be a leader for equity and to establish an equitable school?	3. Cultural Competence	0.84	0.71
What are your greatest areas of strength and areas of growth?	4. Self-Awareness	0.85	0.77
What has been your greatest professional failure? Why did it occur and what did you learn?	5. Continuous Learning	0.86	0.65
(follow up after parent-teacher meeting role play)	6a. Critical Thinking and Problem Solving	0.83	0.68
What did you want to accomplish in the meeting? Were you successful? Why or why not?	6b. Decision Making	0.83	0.74
How do you think that felt for the parent? Would you change your approach if you did it again?	6c. Stakeholder Management	0.84	0.73
Likewise, how do you think that felt for the teacher? Would you change your approach if you did it again?	6d. Overall Communication	0.83	0.80
How would you follow up after this meeting?			

Source: Administrative data from KIPP.

Notes: The overall Cronbach's alpha is the internal consistency when accounting for all items within the competency.

Table C.2. Leave-out internal consistency and interrater reliability for the Vision & Goals competency

Question	Item	Leave-out internal consistency (overall α = 0.87)	Interrater reliability
What is your vision for the school you want to lead? Please describe both your instructional vision and your cultural vision.	1a. Direction-Setting	0.86	0.75
	1b. Student Focus	0.85	0.50
	1c. Achievement Orientation	0.85	0.57
Tell us about a time you identified a problem at your school that impacted multiple stakeholders and the initiatives you took to address it.	2a. Decision-Making	0.85	0.74
	2b. Planning and Execution	0.85	0.75
What is the most successful, recent relationship you've had with either a student or a colleague at your school – what made it successful?	3a. Student Focus	0.86	0.65
	3b. Impact & Influence	0.86	0.63
Tell me about a direct report you have coached toward a goal. What specific coaching strategies did you use to support him/her?	4. Performance Management	0.85	0.62
(follow up after parent-teacher meeting role play) What did you want to accomplish in the meeting? Were you successful? Why or why not How do you think that felt for the teacher? Would you change your approach if you did it again? How would you follow up after this meeting?	5a. Team Leadership	0.87	0.77
	5b. Overall Communication	0.86	0.70
	5c. Overall Cultural Competence	0.87	0.63

Source: Administrative data from KIPP.

Notes: The overall Cronbach's alpha is the internal consistency when accounting for all items within the competency.

Table C.3. Leave-out internal consistency and interrater reliability for the Instructional Leadership competency

Question	Item	Leave-out internal consistency (overall α = 0.84)	Interrater reliability
(follow up after regional leadership meeting role play) Did the candidate provide a correct analysis of the data with the correct core issue? What does the data tell you about what is happening instructionally in the school? What is the core issue that must be addressed in the coming semester? Given your focus for the new semester, what are the possible implications or obstacles?	1a. Correct analysis of core issue (Y/N)	0.86	0.77
	1b. Critical Thinking & Problem-Solving	0.82	0.80
	1c. Achievement Orientation	0.82	0.82
	1d. Decision Making	0.82	0.77
In reflecting on your own data, what specific success are you most proud of?	2a. Self-Awareness	0.83	0.84
If this data was from a teacher on your staff, how would you coach him/her?	2b. Talent Development	0.83	0.74
(follow up after role play of a teacher meeting following up on lesson observation) Evaluation of selected action step(s). How did you select that action step? Why was that action step the highest leverage for the teacher to focus in order to drive student learning?	3a. Highest leverage action step (Y/N)	0.86	0.90
	3b. Instructional Leadership	0.83	0.85
(follow up after schoolwide instructional needs assessment scenario) What steps will you take to address these instructional needs? Why did you select those action steps/what process did you use? As you implement those action steps, what possible consequences and/or obstacles might arise and how would you address them? What goals might you set for your team in relation to your action steps?	4a. Critical Thinking and Problem-Solving	0.82	0.77
	4b. Achievement Orientation	0.82	0.84
	4c. Decision Making	0.83	0.79
Complete with all questions in mind.	5a. Overall Communication	0.83	0.73
	5b. Overall Cultural Competence	0.84	0.77

Source: Administrative data from KIPP.

Notes: The overall Cronbach's alpha is the internal consistency when accounting for all items within the competency.

References

Ashenfelter, Orley. "Estimating the Effect of Training Programs on Earnings." *Review of Economics and Statistics*, vol. 60, no. 1, 1978, pp. 47–57.

Branch, Gregory F., Eric A. Hanushek, and Steven G. Rivkin. "Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals." NBER Working Paper No. 17803, 2012.

Chiang, Hanley, Stephen Lipscomb, and Brian Gill. "Is School Value Added Indicative of Principal Quality?" *Education Finance and Policy*, vol. 11, no. 3, 2016, pp. 283–309.

Clark, D., P. Martorell, and J. Rockoff. "School Principals and School Performance." National Center for Analysis of Longitudinal Data in Education Research Working Paper No. 38, 2009.

Coelli, M., and D.A. Green. "Leadership Effects: School Principals and Student Outcomes." *Economics of Education Review*, vol. 31, no. 1, 2012, pp. 92–109.

Cortina, J.M. "What Is Coefficient Alpha? An Examination of Theory and Applications." *Journal of Applied Psychology*, vol. 78, no. 1, 1993, pp. 98–104.

Cronbach, L.J. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika*, vol. 16, no. 3, 1951, pp. 297–334.

de Vaus, D.A. *Surveys in Social Research, 5th edition*. Crows Nest, Australia: Allen & Unwin, 2002.

Dhuey, Elizabeth, and Justin Smith. "How Important Are School Principals in the Production of Student Achievement?" *Canadian Journal of Economics*, vol. 47, no. 2, 2014, pp. 634–663.

Gates, Susan M., Laura S. Hamilton, Paco Martorell, Susan Burkhauser, Paul Heaton, Ashley Pierson, Matthew Baird, et al. "Preparing Principals to Raise Student Achievement: Implementation and Effects of the New Leaders Program in Ten Districts." Santa Monica, CA: RAND Corporation, 2014.

George W. Bush Institute. Principal Talent Management According to the Evidence: A Review of the Literature. Dallas, TX: George W. Bush Institute, 2016.

Grissom, Jason A., Demetra Kalogrides, and Susanna Loeb. "Using Student Test Scores to Measure Principal Performance." *Educational Analysis and Policy Evaluation*, vol. 37, no. 1, 2015, pp. 3–28.

Ho, A.D., and T.J. Kane. "The Reliability of Classroom Observations by School Personnel." Seattle, WA: Bill & Melinda Gates Foundation, 2013.

Ho, D., K. Imai, G. King, and E. Stuart. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis*, vol. 15, 2007, pp. 199–236.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. "Earnings Losses of Displaced Workers." *American Economic Review*, vol. 83, no. 1, 1993, pp. 685–709.

Miller, Ashley. "Principal Turnover and Student Achievement." *Economics of Education Review*, vol. 36, no. 1, 2013, pp. 60–72.

Neal, Z., J.W. Neal, and A. Piteo. "Call Me Maybe: Using Incentives and Follow-Ups to Increase Principals' Survey Response Rates." *Journal of Research on Educational Effectiveness*, vol. 13, no. 4, 2020, pp. 784-793. https://doi.org/10.1080/19345747.2020.1772423

RAND Corporation. "School Leadership Interventions Under the Every Student Succeeds Act: Evidence Review." Santa Monica, CA: RAND Corporation, 2017.

Walsh, E., and D. Dotter. "The Impact on Student Achievement of Replacing Principals in District of Columbia Public Schools." *Education Finance and Policy*, vol. 15, no. 3, 2020, pp. 518–542.

Wasserman, J.D., and B.A. Bracken. "Psychometric Characteristics of Assessment Procedures." In *Handbook of Psychology: Assessment Psychology*, vol. 10, edited by Irving B. Weiner, John R. Graham, and Jack A. Naglieri. Hoboken, NJ: John Wiley and Sons, 2003.

What Works Clearinghouse. "What Works Clearinghouse Procedures and Standards Handbook, version 5.0." U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2022. https://ies.ed.gov/ncee/wwc/Handbooks.

Mathematica Inc.

Princeton, NJ • Ann Arbor, MI • Cambridge, MA
Chicago, IL • Oakland, CA • Seattle, WA
Tucson, AZ • Woodlawn, MD • Washington, DC

Mathematica®
Progress Together

mathematica.org website

EDI Global, a Mathematica Company

Operating in Tanzania, Uganda, Kenya, Mozambique, and the United Kingdom