

PREPARING URBAN SCHOOL LEADERS: WHAT WORKS?

Extant research, though limited in quantity, increasingly demonstrates the critical connection between quality preparation experience, candidates' leadership capacity, and their subsequent instructional and transformation leadership practices. Using mixed methods, this study builds on the current knowledge base and aims to further verify the link between program preparation and participant learning by examining the various program features and their effect on participant learning as measured by both candidates' self-reported preparedness in key leadership areas and their objective performance from NASSP's Assessment Center. Descriptive analyses indicated that the program in question was comparable to many exemplary preparation programs in terms of participants' perception in curriculum content quality and self-reported preparedness in key leadership areas. Correlation analyses resulted in a consistent positive relationship between cohort structure and participants' self-reported preparedness. NASSP's Assessment Center demonstrated less optimistic results in candidates' leadership skills. However, this did not seem to impact participants' overall positive experience with the program. Qualitative data, while confirming many of the exemplary features of the program (e.g., cohort, mentoring and coaching) demonstrated that quick and episodic changes that characterize many urban school districts seemed to exert a great influence on the consistent implementation of the program, in particular, in internship placement, one of the most important learning blocks in the studied program. The study calls for more systematic support from the district and more seamless collaborations among the partners.

The primary goal of leadership preparation programs is to develop the knowledge, skills, and dispositions of future school leaders. Although empirical data on specific leadership preparation program policies, practices, and outcomes have been slim historically (Perez, Uline, Johnson, James-Ward, & Basom, 2011), a growing interest in advancing program improvement and further establishing the link between preparation and program outcomes has emerged (Orr, 2011). Extant research, though limited in quantity, increasingly demonstrates the critical connection between quality preparation experience, candidates' leadership capacity, and their subsequent instructional and transformation leadership practices (Darling-Hammond, LaPointe, Meyerson, Orr, & Cohen, 2007). This study builds on the current knowledge base and aims to further verify the link between program preparation and participant learning with explicit attention to the urban context where the program is situated. It is designed with reference to the following three strands of literature: (a) levels of evaluation, (b) effective leadership, and (c) quality leadership preparation program features and their effects.

Conceptual Framework

Four Levels of Evaluation on Training Program

Kirkpatrick (1998) proposed four levels of outcomes in assessing the effectiveness of training programs: reaction, learning, behavior, and results. Level 1 (reaction) evaluation assesses what participants think and feel about a training program; level 2 (learning) evaluation gauges the extent to which participants have improved their knowledge or skills; level 3 (behavior) measures participants' knowledge transfer on the job; and level 4 (results) measures the extent to which the training has contributed to the achievement of organizational goals (Russ-Eft & Preskill, 2001). Although it initially emerged half a century ago, Kirkpatrick's taxonomy continues to represent a state-of-the-art model of training evaluation (Kraiger, Ford, & Salas, 1993). When operationalized in the context of school leadership preparation, these four levels of evaluation have the potential to assess program outcomes related to aspiring leaders' satisfaction with their preparation programs, their learning about desired leadership knowledge and skills, their on-the-job performance once placed as school administrators, and their contribution to school improvement and student learning.

Effective School Leadership

Over the past three decades, a significant body of empirical research indicates the following: (a) principal leadership contributes significantly to school effectiveness and student performance (Waters, Marzano, & McNulty, 2003; Bryk, Sebring, Allensworth, Luppescu, & Easton, 2010); (b) principal practices primarily affect student learning indirectly through developing teacher capacity and creating positive organizational conditions (Hallinger & Heck, 1996); and (c) effective principal practices include but are not limited to establishing a focus and vision, developing the capacity of school professionals, building a student-centered learning climate and fostering parent and community trusting relationships (Leithwood & Jantzi, 2005, 2008; Leithwood, Louis, Anderson, & Wahlstrom, 2004). These findings have significant implications for school leadership training. On the one hand, they confirm the strategic importance of leadership development. On the other, they signal important domains of knowledge and skills that preparation programs should focus on in order to develop effective school leaders.

Quality Program Features and Their Effects

With the realization of the principal's role as the change agent and the leverage of school improvement, another body of literature has emerged to identify preparation program features that are effective in cultivating leaders who exemplify the aforementioned leadership behaviors. These program features include an authentic university-district partnership in the rigorous

recruitment and selection of candidates and program delivery (Davis, Darling-Hammond, Meyerson, & LaPointe, 2005; Jackson & Kelley, 2002; Orr & Barber, 2007) and a standards-based curriculum grounded in well-tested theories on instructional leadership, transformational leadership, and organizational learning (Darling-Hammond et al., 2007; Leithwood, Jantzi, Coffin, & Wilson, 1996; Orr, 2011; Orr & Orphanos, 2011). Also a particularly noteworthy finding reveals that active learning strategies with emphasis on rigorous internship, quality mentoring and coaching (Perez et al., 2011), the use of a cohort structure, and data-driven program improvement are critical to the quality of preparation experience. (Darling-Hammond et al., 2007; Davis et al., 2005). This body of literature also suggests that innovative program features are associated with graduates' increased satisfaction with the program, improved knowledge and skills, faster advancement into leadership positions, and more effective school improvement practices once placed.

Although the results of these studies converge on the positive influence of high-quality program features on graduates' learning and their subsequent leadership practices, the finding is less clear regarding how authentically such quality preparation program features can be implemented in urban school districts confronted with unstable central administration and lessened desirability of principalship positions resulting from high turnover of building administrators and teachers and other difficult issues that typically characterize large urban school districts (Morris, 2009; Noguera, 2003; Payzant, 2011). These include a challenging student population, and disengaged parent and school community (Obiakor & Beachum, 2005). Additionally, current literature is short on answers to questions regarding what program features are more effective in facilitating participant learning and strengthening their capacity in turnaround leadership. The present study aims to address these pressing questions.

The purpose of this article is to present findings on the nature of leadership preparation experiences of two cohorts of students enrolled in a federally funded leadership preparation program in a large urban school district from 2009 to 2011. Further, the article aims to investigate the relationships among the participants' characteristics, preparation experiences, leadership learning, and graduate preparedness. Specifically, it addresses the following questions:

- What are the characteristics of program graduates (gender, race/ethnicity, years of teaching experience, prior leadership experience) at the program entry?
- How well does the preparation program reflect the core quality preparation program features, such as authentic district-university partnership, rigorous selection of candidates, research-based curriculum, active instructional strategies, rigorous internship enhanced by quality mentoring, cohort structure, and high quality faculty?
- How well do the graduates do as a result of the program in terms of

satisfaction, participant learning, and sense of preparedness at the exit of the program?

- What design features are most conducive to developing the capacity of urban school leaders as measured by participants' self-perception of learning in key leadership domains and the ratings from the National Association of Secondary School Principals (NASSP) Assessment Center?

Method

Program Background

The leadership preparation program studied here is a partnership program between the Center for Developing Urban Educational Leaders (CDUEL) at Lehigh University, the School District of Philadelphia (SDP), and the National Association of Secondary School Principals (NASSP). The Institute for School & Society (ISS) at the Temple University College of Education serves as the external evaluator. It is funded by a federal competitive grant. The program uses a rigorous selection process to recruit exemplary teacher leaders or non-instructional staff with teaching experience and leadership potential and prepare them to lead the district's most disadvantaged high schools, often classified as not having achieved Adequate Yearly Progress (AYP) under the No Child Left Behind (NCLB) law. The program implementation leverages each partner's strength and features a highly organic and authentic collaborative process. CDUEL serves as the knowledge base, designs the program curriculum, takes primary charge of the academic content delivery, and provides the certification to the principal interns. The SDP coordinates recruitment, selection, internship arrangement, and graduate placement. NASSP provides an assessment tool to measure participant leadership skills at various points of the program and trains principal mentors.

The cohort-based program represents an integrated continuum of leadership preparation and development that spans the trajectory from aspiring school leaders to early career principals and assistant principals. It consists of three discrete groups: (a) Aspiring Leaders (AL) program that provides leadership preparation and principal certification through rigorous coursework and two years of intensive site experiences split in both the home school and the host school guided simultaneously by an high-performing practicing principal and a retired principal mentor; (b) a Developing Leader (DL) program that provides intensive leadership development to individuals with principals' certification by using intensive site experiences also guided by an experienced practicing principal and a retired principal mentor; (c) an Emerging Leaders (EL) program that provides intensive leadership development that is aligned to state mandated job-embedded induction support, and mentoring for those in their first two years as an assistant principal or principal. Participants from both the AL and DL groups

are asked to work in teams and design a school restructuring plan based on their research on the real data of one of the most challenging schools in the district as a culminating project. The three groups are related in that participants who complete the AL and DL programs and have been placed as principal or assistant principal will automatically advance to the EL program.

In addition to an authentic partnership and tailored curriculum for participants with varied credentials and experiences, some other program features are worth mentioning. The Philadelphia High School Leadership Program (PHSLP) utilizes an intensive set of recruitment strategies to attract a large and diverse pool of applicants. It also uses formalized processes for screening and intake to ensure high quality program candidates. The coursework, designed and taught by Lehigh faculty, provides research-based practice-oriented knowledge and skills applicable to turn around low-achieving high schools in a large urban school district. Classes are delivered in the school district during after-school hours and in the summer at 50% subsidized tuition rate to participants. Off-campus classes with the purpose of developing creative and innovative thinking are also provided periodically. One example of such classes is the visit with the residential teaching artists at the Metropolitan Museum of Art in New York City. Since a fulltime year-round internship is not possible, a structured alternative (100 days over two years for ALs and 50 days in one year for DLs) was created to immerse participants in a variety of urban educational leadership settings and put them in contact with trained and experienced host principals and mentors. A standardized 360 degree assessment of educational leadership skills linked to NASSP's 21st Century Principal Skill Dimensions is used to identify participants' areas of strength and weakness with specific emphasis on assignments tailored to individualized leadership plans (ILP) to bolster identified areas in need of improvement, and to evaluate program effectiveness. To keep a coherent experience for participants, mentors and host principals also receive training from NASSP to help aligning their design of coaching and learning experience for principal interns with the 21st Century Principal Skill Dimensions.

Research Design

The study uses a triangulation mixed methods design (a QUANTUAL Model) (Creswell, 2009). We used survey instruments with demonstrated theoretical foundations and high quality psychometric properties to collect information on program participants' demographic and professional information, program attributes and participants' perceived learning in several key leadership areas. Specifically, our survey measures were developed based on Educational Leadership Constituent Council (ELCC) standards and the survey design work of the Stanford University study of exemplary leadership preparation programs (Darling-Hammond et al, 2007). In conjunction with candidates' self-reported learning, candidates'

performance data from the NASSP's Assessment Center were used to test the link between preparation experience and candidate learning. Finally, we complemented the quantitative data with in-depth interviews with selected program participants and their host principals and mentors to gain a realistic picture of their internship experience and learning. Additionally, other qualitative data collected from interviews with program staff, observations of program participants at intern sites and in class, and document analyses were used to further inform our interpretations of program effect. All quantitative and qualitative data were collected and provided by the external evaluator group from Temple University. Together with the triangulated mixed method design, the objectivity and independence of the data source strengthen the integrity and validity of the findings.

Measures.

Dependent measures. The dependent measures fell into three categories: participants' program satisfaction, leadership learning, and sense of preparedness. Using a single 5-point Likert scale item anchored by definitely yes and definitely no, we assessed program satisfaction by asking the participants the likelihood they would choose the same program if provided the opportunity to do it over again. To measure participants' leadership learning, we asked the participants to rate the effectiveness of the program in preparing them to do the following: (a) develop and sustain a learning-centered vision and lead ethically (learned to lead with vision and ethics, 4 items), (b) create a coherent educational program and provide instructional feedback and professional development opportunities to help teachers improve (learned to lead instruction, 3 items), (c) engage staff in school decision making, use data to lead change and monitor school progress, and create a collaborative learning organization (learned to lead organizational learning, 5 items), (d) manage various physical resources and handle disciplinary issues (learned to lead management and operations, 4 items), and (e) work with parents and community (learned to lead parent and community engagement, 2 items). All items were constructed using a 5-point Likert type effectiveness scale anchored by not at all and very well. The validity and reliability of these measures have been assessed previously by other scholars (Darling-Hammond et al., 2007; Orr, 2011).

In additional to the candidates' self report, leadership learning was also measured by the candidates' performance in the NASSP's Assessment Center. The NASSP's Assessment Center uses various simulation activities to measure participants' authentic performance in the following 10 leadership skill dimensions: setting instructional directions, teamwork, sensitivity, judgment, organizational ability, results orientation, oral communication, written communication, development of others, and understanding own strengths and weaknesses. Early validation studies (Schmitt & Cohen, 1990; Schmitt, Noe, Merrit, Fitzgerald, & Jorgensen, 1981) indicated that the content validity and criterion validity of the assessment

were satisfactory. A conceptual congruence (see Table 1) has been found between the assessment center skill dimensions and ISLLC standards (P. Reed, personal communication, June 12, 2012). Therefore, there is reasonable evidence to believe that NASSP’s Assessment Center can serve as a valuable tool to measure aspiring leaders’ learning outcome in terms of effective leadership behaviors.

Table 1

Alignment between ISLLC 2008 & NASSP’s 21st Century Leadership Skills

Educational Leadership Policy Standards ISLLC 2008	NASSP 21ST Century Leadership Skills									
	SI D	T	S	J	RO	OA	OC	W C	DO	US W
Standard 1: An educational leader promotes the success of every student by facilitating the development, articulation, implementation, and stewardship of a vision of learning that is shared and supported by all stakeholders.	√	√	√		√	√	√	√		
Standard 2: An educational leader promotes the success of every student by advocating, nurturing, and sustaining a school culture and instructional program conducive to student learning and staff professional growth.	√	√	√		√		√	√	√	√
Standard 3: An educational leader promotes the success of every student by ensuring management of the organization, operation, and resources for a safe, efficient, and effective learning environment.	√	√		√		√				
Standard 4: An educational leader promotes the success of every student by collaborating with faculty and community members, responding to diverse community interests and needs, and mobilizing community resources.		√	√			√	√	√		
Standard 5: An educational leader promotes the success of every student by acting with integrity, fairness, and in an ethic manner.		√	√	√		√				

(continued)

Table 1 (continued)

Educational Leadership Policy Standards ISLLC 2008	NASSP 21ST Century Leadership Skills									
	SI D	T	S	J	RO	OA	OC	W C	DO	US W

Standard 6: An educational leader promotes the success of every student by understanding, responding to, and influencing the political, social, economic, legal, and cultural context.

√ √ √ √

Note. SID = Setting Instructional Direction, T = Teamwork, S = Sensitivity, J = Judgment, RO = Results Orientation, OA = Organizational Ability, OC = Oral Communication, WC = Written Communication, DO = Development of Others, USW = Understanding Own Strengths and Weaknesses.

The measure of sense of preparedness was assessed using two Likert scale items. Participants were asked how strongly they agree or disagree that they could secure a principal job and could perform well.

Independent measures. Independent measures include graduate characteristics at program entry, program organizational structure (i.e., cohort structure), and leadership preparation program features consisting of curriculum focus and use of active learning strategies (i.e., internship and coaching).

Graduates’ characteristics at program entry. Information on nine participant characteristics was collected: (a) gender, (b) minority (ethnicity), (c) age, (d) highest degree, (e) principal certification, (f) years of teaching experience, (g) administrative positions held, (8) other professional certifications, and (h) other professional development activities. All nine items were single-item measures.

Program organizational structure. Program organizational structure was a single-item measure, assessing the participants’ perception of support from their cohort colleagues using a 5-point Likert scale with 1 indicating not useful at all and 5 very useful.

Leadership preparation program features. Based on ELCC standards and Stanford study (Darling-Hammond et al., 2007), we constructed six measures that assessed the program content foci. Specifically, we asked participants, using a 5-point Likert scale (1 not at all and 5 very useful) to rate the extent to which the program was useful in addressing the following six domains of principal work: (a) leading with vision and ethics, (b) leading instruction, (c) leading a learning organization, (d) leading management and operations, (e) leading parent and community involvement, and (f) leading to influence the larger context. Leading with vision and ethics scale consisted of four items that measured the extent to which the program emphasized creating a learning-centered vision, thinking strategically, and leading in an

equitable manner and honoring diversity; leading instruction scale consisted of four items measuring the extent to which the program emphasized creating a culture of learning, using data to assess school programs, and helping leaders to confront and remedy inadequate practices; leading organizational learning scale included seven items that measured the program's emphasis on leading change, creating a collaborative culture and facilitating school wide team planning, using data to inform decision making and initiate change, coaching faculty and staff at various development levels; using two items, leading management and operations scale assessed the program's emphasis on school resource management and operational matters; leading parent and community involvement was a single item measure that assessed the program emphasis on how to build a beneficial, respectful and collaborative relationships with parents and community members; finally, leading to influence the larger context scale used one item to measure the program's emphasis on developing leadership capacity in advocating for children and public education in the larger context.

Other program attributes such as internship and mentoring and coaching were measured using two items in terms of participants' perceived general support from host principals and mentors. Considering the two items could only paint a broad stroke of the internship experience and the quality of mentoring and coaching, qualitative data from in-depth interviews with participants and the host principals and mentors were rigorously analyzed to assess program effectiveness.

Sample. A total of 19 program graduates from two cohorts and a sample of their mentors and host principals participated in the study. Return rate for surveys that measured program experiences and participant learning approached 95 percent.

Data collection. Participants from each cohort were asked to complete the survey upon their exit of the program. Observations of the program participants in class and at internship sites were conducted frequently in the first year of the program and at least three to four times in the following years of the program. Interviews with host principals and mentors were carried out simultaneously with the survey administration. The external evaluators, a team of researchers from the Institute for Schools and Society (ISS) at Temple University collected all the data. The standard-based candidates' performance score in the 10 leadership skill dimensions was generated from the NASSP's Assessment Center.

Data analysis. The survey data were analyzed using SPSS to generate descriptive measures (means, standard deviations, and percentage distributions). Additionally correctional analyses were conducted to gauge how much participants' satisfaction, leadership learning including both self-reported and the NASSP's Assessment Center data, and sense of preparedness were related to program curriculum foci and other program features such as cohort structure and support from principal mentors. The the-

matic analysis and the constant comparative method (Glaser, 1978) were used to analyze qualitative data.

Quantitative Results

Participant Characteristics at the Program Entry

PHSLP program participants reflected diversity in terms of both gender and racial distributions. Seventy-five percent of the participants in the program were females. Sixty percent of them were African American, five percent Asian, five percent multiethnic, and 40 percent Caucasian. The group also featured rich teaching experience with an average of 15 years. All participants had multiple prior leadership experiences ranging from roster chair and literacy coach to department chair and academy leader. All these participant characteristics compare favorably with regular university-based leadership preparation programs. For example, in her study of 17 university-based leadership preparation programs in 13 institutions, Orr (2011) found that females represented 75% and minority candidates 33% in five programs. Four other programs were represented by no diverse graduate respondents at all and another five programs had fewer than 10% of their graduates self-identified as either African American or Hispanic. The predominant representation of minority candidates, African American candidates in particular, in PHSLP carries potential long-term benefits to students. As representative bureaucracy theory projects, empirical evidence has emerged showing that the presence of African American principals in a predominant Black community is more likely to generate higher teacher job satisfaction and prevent turnover among African American teachers (Grissom & Keiser, 2011). Further, other evidence demonstrates that random assignment to an own-race teacher significantly increased the math and reading achievement of African American students (Dee, 2004). These findings are particularly relevant to the context of the School District of Philadelphia which is challenged with high teacher turnover and a large disadvantaged African American student population. The rich diversity of the program candidates will potentially add value to the equation of improved student achievement in the long run.

Quality Program Features

PHSLP program design seems to reflect almost all of the effective program features highlighted in the literature (Darling-Hammond et al., 2007; Davis et al., 2005). The authentic partnership leverages the strength of a research university as the knowledge base and the advantage of the school district as dependable coordinating source of talent pool, the site for field practice, and the ultimate link to graduate placement. Adding to the solid link between theory and practice is the expertise of a national profes-

sional association (NASSP) in the professional development of mentors and coaches and objective evaluation of program participants' learning. Further, the critical insights provided by the external evaluator make possible the rigorous self-reflection of program staff and ongoing program adjustment.

The cohort-based structure, defined as a group of individuals who began the program together and stayed together throughout their courses (Darling-Hammond et al., 2007), has also been rated as very useful. On average, program participants rated the usefulness of support from their cohort colleagues at 4.47 out of 5. The quality of the cohort has a lot to do with the program's rigorous selection process. Aiming to identify candidates with strong instructional competence and a deep understanding of the unique challenges present in urban public high schools, the program required applicants to undergo a rigorous written application, oral interview, and instructional observation by a panel consisting of university program staff, district central office administrators, and retired principal mentors. Admission decisions were jointly made among the panel members.

In addition to laying a solid infrastructure for the program through authentic partnership and getting the right people on the bus through a rigorous selection process, the curriculum content seemed to reflect the majority of content dimensions considered key to effective principal practices. On average, the program participants perceived that the program content emphasized: how to lead instruction ($M = 4.22$ out of 5), how to lead with vision and ethics ($M = 4.39$), how to lead organizational learning ($M = 4.27$), how to lead management and operations ($M = 4.18$), how to engage parents and community ($M = 4.39$), and how to advocate for children and public education in the larger political and social context ($M = 4.37$). The average ratings of all of the core program features are comparable to the highly rated programs identified by Orr (2011) who used a benchmark of 4.26 or higher on a 5-point scale in evaluating similar core program features. However, when compared to graduates' ratings from the nation's exemplary preparation programs (Darling-Hammond et al., 2007), the rating of curriculum emphasis on instructional leadership is slightly lower. However, the program fares better on measures of management and operations and leadership advocacy role in the larger social context.

Similarly, host principals ($M = 4.11$, $SD = 1.05$), and mentors ($M = 4.05$, $SD = 1.03$) also contributed to participants' positive learning experience. However, this support did not seem to be evenly distributed among all program participants as demonstrated by the comparatively large values of standard deviations for both ratings of the host principal and the mentor.

Graduate Outcomes

Program graduates felt fairly well prepared for virtually every aspect of effective principal practice, ranging from readiness to lead with vision and ethics ($M = 4.12$), readiness to lead instruction ($M = 3.91$) and or-

ganizational learning ($M = 4.05$) to engaging parents and community ($M = 3.87$), and managing school operations ($M = 3.68$). Overall these ratings of PHSLP program were comparable to the exemplary programs and surpassed regular traditional university-based preparation programs (Darling-Hammond et al., 2007). It is noticeable that all of these core areas measured in our program represent leadership practices empirically demonstrated critical to school effectiveness and improved student achievement.

Although graduates generally felt well prepared for the various key aspects of leadership work, their performance in NASSP's Assessment Center did not seem to validate their self-perception. Except for oral communication ($M = 20.03$), participants were in need of great improvement in almost all the other areas. Two skill areas that warranted greatest attention were setting instructional direction ($M = 11.63$) and the development of others ($M = 10.25$). Also noticeable was the great difference present across the participants. For example, in the area of written communication, the gap between the highest (Max = 24.50) and lowest score (Min = 8.00) was as high as 16.50. The large variation of performance among the graduates was present across all skill areas and greatest in areas of written communication ($SD = 5.32$).

Table 2

Graduates' Performance at NASSP's Assessment Center (N = 16)

	Mean	SD	Maximum	Minimum	Range
Setting instructional direction	11.63	3.81	19.00	6.00	13.00
Teamwork	12.69	3.82	19.00	6.00	13.00
Sensitivity	13.75	3.24	19.00	8.00	11.00
Judgment	15.00	4.12	21.00	6.00	15.00
Results orientation	13.88	4.43	20.00	6.00	14.00
Organizational ability	14.25	3.43	19.00	8.00	11.00
Oral communication	20.03	2.66	24.50	15.00	9.50
Written communication	13.47	5.32	24.50	8.00	16.50
Development of others	10.25	2.40	15.00	7.00	11.00
Understanding strengths and weaknesses	13.19	3.92	20.00	6.00	14.00

Note. The highest score for each skill dimension is 30. The lowest possible score for each is 6.

Although the NASSP's Assessment Center data made them more aware of their need for improvement, the graduates expressed their general satisfaction with the overall program. When asked whether they would choose the same program given the opportunity, the majority responded they probably would ($M = 3.74$). Similarly, they expressed their confi-

dence in securing a principal position ($M = 3.58$) and performing well if hired ($M = 3.84$). Although graduates tended to have positive views on their ability in securing principalship ($M = 3.58$) and leading competently ($M = 3.84$), an apparent reservation was present as demonstrated by the rating. Compared to their ratings on other aspects of the program, the ratings on these measures reflect a rather hesitant optimism. The program staff from SDP suggested that the hesitation might have more to do with the District's hiring process than the graduates' self-efficacy.

Link Between Program Features & Graduates' Sense of Preparedness

Bivariate correlations were conducted to test the relationships between program content emphases, usefulness of cohort structure, usefulness of support from host principals and mentors, and the dependent measures: graduate sense of preparedness in core leadership dimensions, graduates' performance in the NASSP's Assessment Center, and their level of satisfaction with the program and their confidence in job placement and work performance.

As Table 3 indicated, with the exception of cohort structure, none of the program features was consistently associated with participants' sense of preparedness in the core leadership areas, or confidence in placement and performance. Cohort structures stood out as the most appreciated program feature in relation to program outcome features. For example, whether a participant felt confident that he/she could perform well as a principal was significantly related to how useful this participant perceived peer support from the cohort ($r = .42, p < .05$). A similar pattern of relationship exists between cohort structure and participants' satisfaction measured by the likelihood a participant would choose the same program if given the opportunity ($r = .43, p < .05$). In addition, a much stronger relationship emerged between cohort and participants' sense of preparedness in core leadership dimensions: lead with vision and ethics ($r = .63, p < .01$), lead instruction ($r = .62, p < .01$), lead organizational learning ($r = .53, p < .01$), and lead management ($r = .55, p < .01$). This cohort effect on positive learning outcome is consistent with previous empirical evidence that cohorts can foster improved academic learning and program completion rates among administrative credential candidates (Davis et al., 2005). Further, the positive effect of cohort on participant program satisfaction may have much to do with the fact that cohort structure enhances feelings of group affiliation and acceptance, social and emotional support, motivation, persistence, group learning, and mutual assistance (Davis et al., 2005). Because teachers tend to give higher ratings to the leadership practices of principals who participated in cohort training structures, it is quite likely that cohorts will not only benefit aspiring principals, but the faculty and students in the schools they ultimately lead (Leithwood et al., 1996).

Table 3
Correlations of Program Curriculum Foci and Graduates' Perceived Learning and Overall Program Effect

Self-reported learning outcome	Curriculum emphasis on how to lead						
	Instruction	With vision & ethics	Organizational learning	Management	Cohort support	Host principal support	Mentor support
1. Learned to lead with vision & ethics	.20	.23	.26	.41	.63**	-.34	-.04
2. Learned to lead instruction	.42	.41	.45*	.25	.62**	.04	.12
3. Learned to lead organizational learning	.29	.32	.37	.41	.53**	-.28	-.11
4. Learned to lead management	.23	.28	.27	.39	.55**	-.44	-.16
5. Learned to lead parent and community	.25	.32	.24	.46*	.28	-.07	.00
6. Overall satisfaction	.13	-.10	.04	.03	.43*	-.29	-.04
7. Confidence in placement	.17	.28	.07	.09	.15	-.65**	-.25
8. Confidence in performance	.43	.35	.29	.14	.42*	-.49**	-.12

* $p < .05$. ** $p < .01$.

In terms of the effect of the content foci, the emphasis on how to lead management had a moderate association ($r = .46, p < .05$) with how well graduates felt prepared to lead parent and community engagement. Similar magnitude of association was also found between the content focus on organizational learning and participants' self-perceived preparedness in leading instruction ($r = .45, p < .05$). Also two negative correlations emerged between how useful the graduates viewed the support from their host principals and how confident they felt about their own replacement and future performance as a principal. These negative relationships were interesting and could probably be attributed to the fact that the gap that the graduates observed between their host principals and themselves. The comparison made more salient the self-perceived needs for improvement. They became more reflective learners and were aware more clearly of their own strengths and weaknesses.

The bivariate analysis between curriculum content and other program features (e.g., cohort structure, support from mentors and host principals), and graduate performance at the NASSP's assessment center revealed only three significant relationships. First, there was a significant relationship between the support from mentors and how well the graduate felt prepared to set instructional direction ($r = .52, p < .05$). This significant relationship suggested that there was a positive relationship between mentors' work and mentees' abilities in developing a vision and establishing clear goals, providing direction in achieving stated goals, encouraging others to contribute to goal achievement, and securing commitment from stakeholders. This is a critical finding considering "establishing goals and expectations" (Robinson, Lloyd, and Rowe, 2008, p. 656) is one of the most effective leadership strategies that leaders can employ to bring student improved learning to fruition. With an effect size of 0.42 as shown in the meta-analysis of Robinson et al., the impact of setting instructional direction was considered moderately large and educationally significant. A positive relationship also emerged between mentor support and graduates' level of judgment ($r = .61, p < .05$). This positive relationship indicated that the more supportive the mentors were perceived, the more likely the mentees were able to seek out relevant data, analyze and interpret complex information, and reach logical and high-quality decisions. Third, a spurious negative relationship emerged between the usefulness of cohort support and how well the participant understood his/her own strengths and weaknesses ($r = -.64, p < .01$). Currently, we do not find any rational explanations for this spurious relationship.

Qualitative Evidences: A Thematic Analysis

The thematic analysis of the in-depth interviews with principal interns (5), host principals (3), home principals (2), and mentors (5) revealed two main themes; the first is related to program strengths, the second to

suggestions for program improvement in issues related to time, internship, and communication. The resulting data was analyzed by utilization of the constant comparative method (Glaser, 1978).

This procedure calls for coding data from interviews, observations, or documents under headings that appear to capture the theoretical properties of that category. Each coded category is then described as succinctly as possible to capture the meanings inherent in it. Then, using that category, all new and existing data are constantly compared to determine the descriptive adequacy of the category. (Haller & Kleine, 2001, p. 201)

Program Strengths and Successes

Features of exemplary principal preparation programs include a cohort structure, access to mentoring from experienced practitioners, sustained internship time in schools, active learning, coherent and aligned curriculum, standards-based assessment, ample time for reflection, and a program defined around leadership for school improvement (Darling-Hammond et al., 2007; Davis et al., 2005; Orr & Orphanos, 2011). Participants noted a variety of program strengths consistent with the aforementioned features: the cohesiveness of the cohort; the practical experience provided by the internship; access to experienced mentors; program emphasis on vision and teacher supervision; hands-on tasks assigned (such as planning a three day orientation for over 100 students); class activities such as instructional rounds; experienced principals who visited the class throughout the year to offer practical advice; and the reflective nature of the program.

Systematic analysis of home, host, and mentor principal qualitative data using Kirkpatrick's (1998) first two levels of outcomes (reaction and learning) revealed an overall positive reaction to participating in the program. All reported enjoying the role and the opportunity to build trust and supportive relationships with interns. One mentor reported "... I enjoy mentoring to begin with, that's probably the thing I enjoy the most." Several principals appreciated the chance to give back to the profession. One home/host principal noted "... I really, truly enjoy it because I wouldn't be where I am now if people hadn't mentored me, so I really take it seriously." Another principal said "... It's something that I feel really is my reasonable service as an administrator. It's something I would do gladly anyway in the absence of this program." Several principals commented positively on their own learning as afforded by mentor training provided by NASSP and praised the NASSP self-evaluation tool used by interns.

In addition, participants elaborated on specific examples of program success (interns' behavior—Kirkpatrick's third level of evaluation) that ranged from changes in mentality to cohesiveness of the cohort. One home principal remarked (referring to a participant):

Well, he's stepped up, so to speak, administratively. He's always available. I think his mindset is changing. I can see his mindset changing from a teacher's mindset to an administrator's mindset. And that's a process. It's all those little details that a teacher doesn't have to bother with so, as an administrator, you have to think of all those details in order for something to run smoothly. So that's a change in mindset.

In reference to the cohort, a participant noted, "One success with this particular cohort I think is the cohesiveness of the cohort group." These reflections are important with regard to morale and positive feedback needed for program improvement. In addition, these comments represent small wins (Kotter, 1999), which will subsequently generate excitement, energy and commitment (Kouzes & Posner, 2007).

Finally, home, host and mentor principals commented on Kirkpatrick's fourth level of program evaluation, results, in noting the extent to which the program prepared interns to take on administrative positions and contribute to the achievement of organizational goals. Commenting with pride on two interns who had been placed, one principal said "They've established themselves as instructional leaders in the building."

Suggestions for Program Improvement

At the same time, participants offered a number of ways in which they thought the program could be improved, most notably in the structuring of the internship, the way in which lack of time impacted the program quality, and the need for clear communication among stakeholders.

Time. To varying degrees, all participants noted that time was an issue: the demands of working full time while participating in the program; the timing of meetings with mentors and course assignments; the reduced amount of time spent in host schools from the first to second year of the program; and the fragmented structure of the time spent in host schools due to the District's budgetary concerns. Many expressed frustration with the structure and configuration of the internship, noting that the amount of time spent in host schools was not enough to allow for sufficient leadership skill development. Many noted that interns experienced competing demands on their time: a full time teaching position, fall and winter classes and assignments, as well as meetings with host principals and mentors. The use/misuse of time seemed to be a common issue across participants. According to one host principal, "I just wish I could have more time with them, more of a free, the way I was freed up to be a quasi-administrator in a school." This administrator was referring to their full-time internship experience. Another host principal remarked, "You know, you had several weeks here, several weeks there and whatever they say the time should be, you need to give us that time all up front. So that it can be more meaningful, both for

the intern and for the administrator (mentor), the host administrator.” Here we see the reiteration of time and especially time dedicated to the experience of building leadership. Time is an incredibly important factor in modern organizational life (Cox, 2001). It is also one that leaders must structure and manage properly to optimize effectiveness (Covey, 1989).

Internship. Ideally, internships allow aspiring leaders to experience “in situ” the real life day-to-day demands of school administration, admittedly an often frustrating blend of managerial and leadership responsibilities. It is by being in the schools and working through the myriad of situations that characterize school life that candidates integrate the theory learned in class with the often messy practice experienced in school. It is by being placed in the kinds of complex situations that characterize school interactions that aspiring leaders begin to develop the skills they will need to assume full responsibility for leading a school.

It is important to note that the structure of the internship changed over the course of the program. The internship was originally conceived to include four ten-day work periods in host schools: at the beginning of the school year, the middle of the fall term, the middle of the winter term, and the end of the school year. The rationale for this internship structure was that interns would experience the planning and preparation involved in beginning and ending the school year, as well as observing and working in the school administratively during normal periods of operation. In the first year of the program, the internship operated as originally conceived. Despite some drawbacks, participants expressed satisfaction with having this opportunity to work closely with host principals and work full-time in the host schools.

However, district directives called for changes in the internship structure in the second year of the program. The work periods in host schools were reduced from four to two. Interns were still in host schools at the beginning and end of the school year; however, the mid-term work periods were eliminated and replaced by interning in a month-long summer program at host sites. To compensate for the reduction in direct-contact time in schools, interns were given tasks/assignments which required them to meet with their host principals and mentors in order to talk through situations or dilemmas. However, internship design changes made to accommodate district directives appeared to have negatively affected intern morale, recruitment to the program, internship quality, placement outcomes, perceptions about the program, and support for the program.

The changed internship model during the school year was frustrating for all participants. One host principal commented that the reduced contact meant that he was unable to provide the direct experience in a variety of situations, such as principal meetings, hearings, and mandatory conferences that he was accustomed to giving his interns. “So they [my interns] really had a broad experience of leadership last year and because of the constraints this year, I’m not able to provide that for my in-

terns this year.” Another expressed concern that the interns were not being adequately prepared given the reduced amount of time spent in the host schools. “I’m not able to give them the type of experiences that I know they need in order to grow.”

From the interns’ perspective, the newly configured internship meant that they were not able to experience directly the types of situations that would prepare them to manage the day-to-day operations of school life. Many interns expressed frustration over the demands of full-time teaching combined with the demands of the program. One intern commented that “... during the school year, between the [evening] courses, all the other things that are going on and being expected of us, the work from the courses becomes a little bit overwhelming in terms of trying to put in real, good quality work together.”

The complexity of the principal’s job formed a common thread in the home/host/mentor principal qualitative data. All expressed concern that interns would not be on site often enough to experience the full array of responsibilities that comprise the role of the principal. From working with an employee to negotiating Family Medical Leave to building relationships with maintenance staff to working with teachers to improve instruction, principals perform a myriad of duties not often visible in the short amounts of time given to interns.

Communication. A variety of suggestions for improvement centered on improved communication about the goals and activities of the program among all stakeholders: clear support of the program from the school district; invitations to mentors, home and host principals to all program activities; and access to descriptions of the program’s curriculum to allow for better integration of theory and practice. In large urban school districts, change can be quick, random, and/or episodic (Payzant, 2011; Obiakor & Beachum, 2005). A host principal noted, “Well, right now, the concern about the program is that the district seems to be moving in another direction ...” Unfortunately, this direction is not always well communicated and shared throughout the organization, especially with a change in leadership at the top. A mentor stated that there needed to be better information sharing between partners in the program. Communication is an important component for success in schools (Fullan, 2004). Beachum, McCray, and Huang (2010) agreed, “Without credible communication, the hearts and minds of people are never won” (p. 58). Unfortunately, the sheer size and complexity of schooling in some of America’s largest cities poses some serious challenges to effective communication strategies.

Conclusions

The socioeconomic adversities present at large urban school communities are likely to have thwarted many potential candidates’ pursuit of serving as a school leader. The PHSLP program employed multiple meth-

ods to reach the most capable and eligible candidates. The final candidates in the first two cohorts were rich in diversity with 75% female and 60% of minority candidates. Additionally, these aspiring leaders had exemplary teaching and quasi-administrative experience.

The program also reflected multiple innovative features, such as university-district partnership, cohort structure, research-based content, structured internship augmented by quality mentoring and coaching. Among them, cohort structure was viewed as the most supportive program feature. It was associated with all aspects of leadership learning among the graduates. However, these exemplary features in a single program did not seem to produce consistent singular learning outcome among all candidates. In general, program participants expressed great satisfaction with the program and perceived themselves well prepared in all areas of effective leadership (e.g., lead with vision and ethics, lead instruction, lead organizational learning). However, this self-perception did not seem to be in line with their actual performance in NASSP's Assessment Center, a more objective form of performance evaluation. Overall, program participants had the greatest needs for improvement in the areas of setting instructional direction and the development of others, two of the most effective domains of leadership that showed greatest impact on student achievement with an effect size of 0.42 and 0.84 respectively (Robinson et al., 2008). It is also of note that great variations existed across all 10 dimensions of leadership skills among all candidates. This variation in performance suggested not only that every candidate experienced the learning curve differently, but also that each of them might have gained different levels of support through their internship and coaching experience. This has great implications for program implementations. A well-designed program infrastructure (e.g., cohort, teaming of interns with high-performing host principals and mentors) is not enough to guarantee seamless quality execution. Some measures have to be taken at the program level to ensure every participant gains quality learning experience. Future research pertaining to consistent quality experience across all principal interns will help solve this challenge.

Themes emerged from qualitative data analysis that suggested the challenges associated with quick, random, and episodic changes often taking place in large urban school districts (Payzant, 2011; Obiakor & Beachum, 2005). The sudden change in internship model, one of the most important learning blocks in leadership learning, had necessarily brought about great frustrations in both the interns and the host principals. Leadership matters and it matters most in places that need it most, the majority of urban school districts (Leithwood et al., 2004). Unless urban school districts demonstrate consistent and unconditional system support to leadership development programs, the quality of leadership learning will be likely compromised.

Limitations and Recommendations

Caution has to be taken when interpreting the quantitative relationship between program input variables and program outcome variables considering the sample size ($N = 19$). The connection of program features with program outcomes was further challenged by the use of a single program. Future research can expand the investigation by comparison across multiple programs. Utilization of multiple programs will not only eliminate the challenges associated with sample size but also afford scholars more powerful statistical methods to detect the within- and between-program differences and the use of techniques such as multiple regression to determine how much leadership learning can be attributed to various program features. Additionally the scales used to measure curriculum foci and participants' sense of preparedness warrant further factor analyses to test their validity and reliability. However, due to our small sample size, such analyses were not feasible in this study. Finally, our current qualitative data did not provide sufficient evidence in assessing how effectively the off-campus activities (e.g., the art of observation at the Metropolitan Museum of Art) developed participants' capacities for imaginative thinking. Similarly, we were short on evidence in assessing the value of the culminating restructuring project. A revision of our research protocols to be used for the third cohort is underway. Follow-up studies will focus more on the uniqueness of our program in addition to exemplary program features highlighted in extant literature.

Author Note

The authors would like to thank Jon Drescher and Marcia Schumann for their helpful comments and suggestions on this paper.

References

- Beachum, F. D., McCray, C. M., & Huang, T. (2010). Administrators and teachers sharing leadership: Utilizing C.A.R.E. in urban schools. *Making Connections*, 12(1), 51–62.
- Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago: University of Chicago Press.
- Cox, T., Jr. (2001). *Creating the multicultural organization: A strategy for capturing the power of diversity*. San Francisco: Jossey-Bass.
- Covey, S. R. (1989). *The 7 habits of highly effective people: Powerful lesson in personal change*. New York: Free Press.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed-methods approaches* (3rd, ed.). Thousand Oaks, CA: Sage.

- Darling-Hammond, L., LaPointe, M., Meyerson, D., Orr, M. T., & Cohen, C. (2007). *Preparing school leaders for a changing world: Lessons from exemplary leadership development programs*. Stanford, CA: Stanford University, Stanford Educational Leadership Institute.
- Davis, S., Darling-Hammond, L., Meyerson, D., & LaPointe, M. (2005). *Review of research. School leadership study: Developing successful principals*. Palo Alto, CA: Stanford University, Educational Leadership Institute.
- Dee, T. S. (2004). Teacher, race, and student achievement in a randomized experiment. *Review of Economics and Statistics*, 86(1), 195–210.
- Fullan, M. (2004). *Leading in a culture of change: Personal action guide and workbook*. San Francisco: Jossey-Bass.
- Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. Mill Valley, CA: Sociology Press.
- Grissom, J. A., & Keiser, L. R. (2011). A supervisor like me: Race, representation, and the satisfaction and turnover decisions of public sector employees. *Journal of Policy Analysis and Management*, 30(3), 557–580.
- Haller, E. J., & Kleine, P. F. (2001). *Using educational research: A school administrator's guide*. New York: Longman.
- Hallinger, P., & Heck, R. (1996). Reassessing the principal's role in school effectiveness: A review of empirical research, 1980–1995. *Educational Administration Quarterly*, 32, 5–44.
- Jackson, B. L., & Kelley, C. (2002). Exceptional and innovative programs in educational leadership. *Educational Administration Quarterly*, 38, 192–212.
- Kirkpatrick, D. L. (1998). *Evaluating training programs: The four levels* (2nd ed.). San Francisco: Berrett-Koehler.
- Kotter, J. P. (1999). Leading change: The eight steps to transformation. In J. A. Conger, G. M. Speitzer, & E. E. Lawler III, (Eds.), *The leader's change handbook: An essential guide to setting direction & taking action* (pp. 87–99). San Francisco: Jossey-Bass.
- Kouzes, J. M., & Posner, B. Z. (2007). *The leadership challenge* (4th ed.). San Francisco: Wiley & Sons.
- Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology*, 78, 311–328.
- Leithwood, K., & Jantzi, D. (2005, April). *A review of transformational school leadership research*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.
- Leithwood, K., & Jantzi, D. (2008). Linking leadership to student learning: The contribution of leader efficacy. *Educational Administration Quarterly*, 44(4), 496–528.
- Leithwood, K., Jantzi, D., Coffin, G., & Wilson, P. (1996). Preparing school leaders: What works? *Journal of School Leadership*, 6, 316–342.

- Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning*. Toronto, Canada: Center for Applied Research and Educational Improvement and Ontario Institute for Studies in Education.
- Morris, J. E. (2009). *Troubling the waters: Fulfilling the promise of quality public schooling for Black children*. New York: Teachers College Press.
- Noguera, P. (2003). *City schools and the American Dream: Reclaiming the promise of public education*. New York: Teachers College Press.
- Obiakor, F. E., & Beachum, F. D. (2005). Urban education: The quest for democracy, equity, and excellence. In F. E. Obiakor & F. D. Beachum (Eds.), *Urban education for the 21st century: Research, issues, and perspectives* (pp. 3–19). Springfield, IL: Charles C. Thomas.
- Orr, M. T. (2011). Pipeline to preparation to advancement: Graduates' experiences in, through, and beyond leadership preparation. *Educational Administration Quarterly*, 47(1), 114–172.
- Orr, M. T., & Barber, M. E. (2007). Collaborative leadership preparation: A comparative study of innovative programs and practices. *Journal of School Leadership*, 16, 709–739.
- Orr, M. T., & Orphanos, S. (2011). How graduate-level preparation influences the effectiveness of school leaders: A comparison of the outcomes of exemplary and conventional leadership preparation programs for principals. *Educational Administration Quarterly*, 47, 18–70.
- Payzant, T. (2011). *Urban school leadership*. San Francisco: Jossey-Bass.
- Perez, L. G., Uline, C. L., Johnson, J. F., James-Ward, C., & Basom, M. R. (2011). Foregrounding fieldwork in leadership preparation: The transformative capacity of authentic inquiry. *Educational Administration Quarterly*, 47, 217–257.
- Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly*, 44, 635–674.
- Russ-Eft, D., & Preskill, H. (2001). *Evaluation in organizations: A systematic approach to enhancing learning, performance, and change*. Cambridge, MA: Perseus.
- Schmitt, N., & Cohen, S. A. (1990). *Criterion-related validity of the National Association of Secondary School Principals' Assessment Center*. Reston, VA: NASSP.
- Schmitt, N., Noe, R., Meritt, R., Fitzgerald, M., & Jorgensen, C. (1981). *Criterion-related and content validity of the NASSP Assessment Center*. Reston, VA: NASSP.
- Waters, J. T., Marzano, R. J., & McNulty, R. A. (2003). *Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement*. Aurora, CO: McREL.

Tiadan Huang is a doctoral student in the Department of Education and Human Services at Lehigh University, Bethlehem, Pennsylvania.

Floyd D. Beachum is an Associate Professor in the Department of Education and Human Services at Lehigh University, Bethlehem, Pennsylvania.

George P. White is a Professor of Educational Leadership and Director of the Center for Developing Urban Educational Leaders at Lehigh University, Bethlehem, Pennsylvania.

Girija Kaimal is a Senior Researcher in the Institute for Schools and Society at Temple University, Philadelphia, Pennsylvania.

Anne Marie FitzGerald is a doctoral student in the Department of Education and Human Services at Lehigh University, Bethlehem, Pennsylvania.

Peter Reed is the Director of Professional Development at the National Association of Secondary School Principals, Reston, Virginia.