Effects of the “High Impact Leadership for School Renewal” Project on Principal Leadership, School Leadership, and Student Achievement

Xin Ma 1,*, Jianping Shen 2, Patricia L. Reeves 2, Huang Wu 3, LaSonja Roberts 2, Yunzheng Zheng 2 and Qianqian Chen 1

1 Department of Educational, School, and Counseling Psychology, University of Kentucky, Lexington, KY 40506, USA; qianqian.chen@uky.edu
2 Department of Educational Leadership, Research and Technology, Western Michigan University, Kalamazoo, MI 49008, USA; jianping.shen@wmich.edu (J.S.); patricia.reeves@wmich.edu (P.L.R.); lasonja.roberts@wmich.edu (L.R.); yunzheng.zheng@wmich.edu (Y.Z.)
3 Division of Educational Leadership, Policy, and Foundations, University of Missouri-Kansas City, Kansas City, MO 64110, USA; hwu@umkc.edu

* Correspondence: xin.ma@uky.edu

Abstract: In this study, we designed an educational experiment to evaluate the High Impact Leadership for School Renewal project’s impact on principal leadership, school leadership, and student achievement. Principals in the experimental group reported statistically significantly greater improvement in principal leadership than their counterparts in the control group (Hedge’s g of 0.73), but teachers in the experimental and control groups showed no difference in their ratings of principal leadership. Teachers in the experimental group reported statistically significantly more improvement in school leadership than their counterparts in the control group (Hedge’s g of 0.53 on the scale of Orientation to School Renewal and 0.58 on the scale of Learning-Centered School Leadership), but principals in the two groups reported no difference in the growth in school leadership. Schools in the experimental group showed statistically significantly larger annual growth in English language arts proficiency rate (2.05 percentage points more annually). The implications of the findings were discussed in the context of leadership development programs for school improvement.

Keywords: high impact leadership; principal leadership; school leadership; student achievement

1. Literature Review and Study Justification

How to improve our schools through the leadership approach is among the most important questions for educational research, policy, and practice. A preponderance of research on this topic is of correlation (i.e., association rather than causation), and neither longitudinal designs nor educational experiments are very common as a way to study the effects (i.e., causation rather than association) of leadership development programs on school outcomes (e.g., student achievement). In this article, we report the efficacy results of an educational experiment, a federally funded project titled High Impact Leadership for School Renewal.

1.1. Principal Leadership

Principal leadership is a well-known leadership concept, and a vast body of literature addresses whether there is any association between principal leadership and school outcomes related to students, teachers, and schools. First, numerous empirical studies indicate associations between principal leadership and a wide range of teacher outcomes. Principal leadership is associated with teachers’ sense of well-being at school [1,2], self-efficacy [3,4], collaboration [5], commitment to teaching [6–8], and turnover and attrition [9,10]. Principal leadership is also associated with teachers’ professional learning [11–13] and classroom...
practice [14,15]. Burkhauser’s longitudinal analysis of North Carolina public schools showed that principals have significant effects on teachers’ satisfaction with their working conditions [16]. In another longitudinal study in Tennessee, Grissom and Bartanen found that schools with more effective principals demonstrated a lower rate of teacher turnover [9]. A recent meta-analysis by Liebowitz and Porter estimated the effect size of principal leadership was 0.38 SD for teacher well-being and 0.35 SD for teacher instructional practices [17].

Second, recent studies have begun to highlight the ways in which principals shape organizational (school) outcomes. A significant amount of empirical literature associates principal leadership with school climate and culture [1,18,19]. McCarley et al. found that transformational leadership is positively associated with a supportive and engaged school climate [20]. While principal leadership is positively associated with family involvement and engagement [21], Smith et al., in a study based on data from 18 schools in a Midwest state in the United States (U.S.), demonstrated the substantial effects of principal leadership on both baseline and improvement in family engagement after controlling for background variables [22].

Third, another extensive set of empirical studies attempted to establish the association between principal leadership and student achievement [23]. In their highly referenced review, Leithwood et al. argued that “leadership is second only to classroom instruction among all school-related factors that contribute to what students learn at school” (p. 5) [24]. Based on six rigorous studies that utilized longitudinal data, Grissom et al. showed that, on average, one standard deviation increase in principal leadership can increase students’ mathematical achievement by 0.13 SD and their reading achievement by 0.09 SD [25]. By summarizing results from 12 previous meta-analyses performed from 2000 onwards, Wu and Shen confirmed the positive association between principal leadership and student achievement (effect size = 0.34 SD) [26]. Furthermore, principal leadership is associated not only with average student achievement but also with the growth of student learning [27]. Based on data from 24 schools, Shen et al. revealed that improvement in principal efforts in promoting parental involvement had significant effects on growth of student achievement in both reading and math [28].

Finally, there is also a body of literature on how principal leadership is associated with school outcomes, which, in turn, are positively associated with student achievement. These studies fall into three categories. First, principal leadership is directly associated with outcomes. Distributed leadership practices such as building collaboration and providing teachers with instructional support, intellectual stimulation, and opportunities to participate in school decision-making are associated with teacher job satisfaction and commitment [2,29]. Some principal leadership practices are also associated with student achievement, such as engaging in formal or informal discussions with students [30]. Second, since school administrators, teachers, students, families, and communities are interconnected, principals’ associations with student achievement may manifest through their direct efforts in promoting school and teacher outcomes. Leithwood et al. suggested rational, emotional, organizational, and family paths through which principal leadership is associated with student learning [31]. Third, principal leadership may interact with school background and conditions on student learning (i.e., the association of principal leadership with student learning is mediated by school background and learning conditions) [32,33]. In summation, the wide range of empirical evidence seems to suggest that the association of principal leadership with school outcomes is both ubiquitous and versatile.

1.2. Teacher Leadership

In recent years, teacher leadership has received growing attention as an essential ingredient in school improvement from educational policymakers, practitioners, and researchers. The notion of teacher leadership refers to teachers’ formal and informal influence over instructional practice and organizational decision-making processes in efforts to transform and improve schools [34]. Other important leadership concepts such as collaborative
leadership [27], distributed leadership [35], and teacher empowerment [36] carry similar meanings, all of which emphasize that school leadership is a collective activity that, not restricted to designated formal positions, involves teachers and other school personnel [37].

Existing research on teacher leadership and professional standards for teachers recognizes various domains of practice in which teachers can show their leadership. York-Barr and Duke summarized literature into seven domains of teacher leadership practice: (a) coordination and management, (b) school or district curriculum work, (c) professional development of colleagues, (d) participation in school change/improvement, (e) parent and community involvement, (f) contribution to the profession, and (g) preservice teacher education [38]. In a more recent meta-analysis of the literature, Shen et al. identified seven dimensions of teacher leadership: (a) promoting a shared vision, mission, and goals of student learning, (b) coordinating and managing beyond the classroom, (c) facilitating improvements in curriculum, instruction, and assessment, (d) promoting teachers’ professional development, (e) engaging in policy and decision making, (f) improving outreach and collaboration with families and communities, and (g) fostering a collaborative culture in school [39].

There is some overlap between the two lists of dimensions. York-Barr and Duke’s list of dimensions include themes that emerge from a content analysis of the literature, while Shen et al.’s list of dimensions is based on (a) 10 frameworks of teacher leadership, including York-Barr and Duke’s work, (b) various operationalized elements in the empirical studies, and (c) the dimensions supported by the results of meta-analyses. In other words, Shen et al.’s list can be considered an update, refinement, expansion, and validation of York-Barr and Duke’s list.

Studies suggest that teacher leadership has significant positive associations with student, teacher, and school outcomes [34,38,40]. In their meta-analysis of 21 empirical studies, Shen et al. found positive associations between student achievement and teacher leadership in (a) facilitating improvements in curriculum, instruction, and assessment (0.21 in correlation), (b) promoting teachers’ professional development, (c) engaging in policy-and decision-making (correlation of 0.19 in both cases), (d) promoting shared vision and mission, (e) coordinating and managing beyond the classroom (correlation of 0.18 in both cases), (f) fostering a collaborative culture (0.17 in correlation), and (g) improving outreach and collaboration with families and communities (0.15 in correlation) [39]. They concluded that the positive association between teachers’ leadership and students’ academic achievements has an average correlation of 0.19 (specifically, 0.18 for reading and 0.24 for math). These measures equate to roughly 8.1 months of learning in reading and 8.5 months in math at fourth grade in the standard nine-month school year [41,42].

1.3. School Leadership: Connecting Principal Leadership and Teacher Leadership

In reality, teacher leadership and principal leadership are connected. For example, distributed leadership is a construct (and practice) that connects teacher leadership and principal leadership [43]. Leadership is shared among stakeholders, relying on school-wide sources of leadership [44]. The experimental study by Jacob and her colleagues revealed the effects of the Balanced Leadership Program (https://www.mcrel.org/balancedleadership (accessed on 10 April 2024)) on leadership, principal efficacy, instructional climate, educator turnover, and student achievement [45]. Although principals participating in the program self-reported more positive outcomes than their counterparts in the control group, data collected from teachers indicated no effects on their schools’ instructional climate. There were also no effects on student achievement. When Jacob et al. discussed their research findings, one of the key discussion points was that leadership training for principals without contextual and sustained implementation was not enough. One important element concerning this implementation is to work with teachers. In our educational experiment, our focus was on both teacher leadership and principal leadership; this involved working with a team consisting of the principal and three teacher leaders from each school.
1.4. Experimental Studies on Leadership Development

There have been some experimental inquiries into the effects of leadership development programs for aspiring or practicing principals, but few significant effects have been identified, especially concerning student learning [46,47]. Clark et al. and Corcoran et al. assessed the effects of New York City’s Aspiring Principals Program (https://wallacefoundation.org/sites/default/files/2023-07/Taking-Charge-of-Principal-Preparation.pdf (accessed on 10 April 2024)) on student achievement [46,48]. There was no significant difference between the experimental and comparison schools in terms of student achievement in any year when the program was implemented [46]. The National Institute for School Leadership program (https://ncee.org/nisl-program (accessed on 10 April 2024)) was evaluated in Pennsylvania [49,50], Massachusetts [51,52], and Wisconsin [53]. These evaluations did suggest significant, though small, student achievement gains.

A few experimental inquiries did reveal some positive effects of leadership development programs. Studies focusing on the New Leaders program (https://www.newleaders.org/programs (accessed on 10 April 2024)) found that students in schools led by New Leaders principals showed significantly larger achievement gains on average than their counterparts in schools led by other principals [54]; however, the effects varied significantly across districts. In a follow-up evaluation, Gates et al. reported that (a) K-8 students in schools led by first-year New Leaders principals outperformed students in comparison schools led by first-year principals trained through other avenues, (b) the attendance rates in elementary and middle schools with first-year New Leaders principals were higher, and (c) first-year New Leaders principals were more likely to stay as principals at their schools for a second year than other first-year principals [55]. Similarly, Steinberg and Haisheng evaluated Pennsylvania’s Inspired Leadership program (https://www.education.pa.gov/Teachers-%20Administrators/PA%20Inspir%20Leaders/pages/default.aspx (accessed on 10 April 2024)) and reported that (a) the program improved student achievement in math (effect size = 0.10 SD) but not in reading, (b) the effects were concentrated among low-achieving and economically disadvantaged schools, and (c) the program had no significant effects on teacher turnover [56]. Steele et al. examined the first three years of implementation of the Pathways to Leadership in Urban Schools residency program (https://phillyplus.org (accessed on 10 April 2024)) in an urban school district, and they found positive effects in high schools in terms of student math achievement, graduation rate, and suspension rate, but the results were sensitive to school level, residents’ placement, and model specification [47].

Most experimental inquiries on the effects of leadership development programs are quasi-experimental in nature, and there are a very limited number of experimental inquiries based on more vigorous experimental design (i.e., randomized controlled trials) [57]. While investigating 126 economically disadvantaged rural schools in Michigan, Jacob et al. examined whether the Balanced Leadership Program has effects on leadership, principal efficacy, principal turnover, instructional climate, teacher turnover, and student achievement [45]. Their results suggested that the program only has potentially positive effects on teacher and principal retention and principals’ self-reported measures. In Texas, Freyer estimated the effects of principal leadership training on student achievement. Participants were provided with 300 h of training on lesson planning, data-driven instruction, and teacher observation and coaching across two years [58]. The results showed significant positive effects on student achievement in the first year but not in the second year. Moreover, the University of Washington Center for Educational Leadership’s program (https://k-12leadership.org (accessed on 10 April 2024)) provided about 200 h of professional development over 2 years for principals in 100 elementary schools from eight districts in five states, with a focus on instructional leadership, human capital management, and organizational leadership. Herrmann et al. indicated that the program did not have any effects either on student achievement or on teacher or school outcomes (e.g., teachers’ perception of school climate, teacher retention, principal retention) [59].

Our study made an effort to assess the effects of the High Impact Leadership for School Renewal project (the HIL Project in short) (see below for further details). Our initial design
was a randomized controlled trial. Because of the difficulties in retaining participating schools in a multi-year, longitudinal design, some schools dropped out of our project. We refrained from using the term randomized controlled trial to reflect this fact. But our project had a very unique strength in that we addressed Jacob et al.’s recognition of the need for teacher involvement in any leadership training programs [45]. Our project worked with a team, which consisted of the principal and three teachers from each school.

Overall, the above literature review has indicated that although a large body of literature points to the positive associations of teacher leadership and principal leadership with (and sometimes the positive effects of teacher leadership and principal leadership on) various school outcomes, there has not yet been any consensus concerning the effectiveness of leadership development programs in the literature (i.e., research findings are inconclusive). This lack is surely a call for more experimental inquiries to accumulate empirical evidence based on research designs that are as vigorous as possible. Our study aimed to inform the literature in such a spirit.

2. Rationale for and Implementation of the HIL Project

2.1. School Renewal as the Underlying Construct for the HIL Project

Underlying the HIL Project is the construct of school renewal. This project addresses a persistent issue faced by schools in their attempt to achieve transformative change leading to significantly improved school outcomes. Principals are expected to lead teachers and others to sustain second-order (transformative) change and address a wide range of issues related to student achievement [60]. Principals and teachers work in a complicated policy environment [61] that requires strategic and systemic responses [62]. Yet, studies also show that transformative change happens at the school level [63], and principals play a significant role in shaping and leading that change [64] with leadership practices that are highly distributed and inclusive [27]. Therefore, based on our team’s previous funded projects and research, the fundamental idea underlying the HIL Project is school renewal, a construct that emphasizes the collective role of the principal and teachers in balancing internal and external influences, using a “dialogue, decision, action, and evaluation” process for change initiative in the school with a focus on student achievement, implementing the change initiative with both integrity and fidelity, and engaging in a continuous renewal process as part of the intrinsic responsibility and professionalism for both principals and teachers.

2.2. HIL Principles and Practices and the Renewal Cycle

To operationalize the construct of school renewal, HIL’s 4 principles and 16 practices were framed to grow a school’s capacity and motivation for sustained, strategic, and systemic change (see Figure 1). To further guide the school renewal process in each school, we developed training activities that coach school leadership teams to adopt specific leadership behaviors associated with the 16 HIL leadership practices and complete a full HIL school renewal cycle to implement evidence-based literacy practices (see Figure 2). Data-informed decision-making [65,66] ran through the entire renewal cycle. HIL (designated) facilitators trained principals and teacher leaders to guide, engage, and support their schools through ongoing cycles of school renewal to identify needed change, monitor progress, and achieve priority growth targets for student literacy success. Schools were asked to produce critical deliverables associated with implementation integrity and fidelity, including artifacts for each phase of the school renewal cycle as follows:

1. Student performance profile, including both post hoc and real-time data to identify student strengths and growth edges (i.e., zones of proximal development).
2. School performance profile, including real-time data to describe the current state of school practices and conditions to support student growth targets and determine growth opportunities for classroom, leadership, and school-wide support.
3. Student priority growth targets (PGTs) and the classroom, leadership, and school-wide support practices. PGTs for students are based on students’ zones of proximal
development. PGTs for teachers, school leaders, and other support personnel are based on student PGTs.

4. Gold Standard Implementation Guides (GSIGs) for the classroom, leadership, and school-wide support practices to achieve student PGTs. GSIGs for each school renewal cycle include (a) the student PGTs to be achieved; (b) the adult PGTs to be achieved; (c) the critical features (i.e., the fidelity elements) of each evidence-based classroom practice; (d) the gold-standard teacher behaviors (i.e., the fidelity behaviors) to implement each critical feature; (e) the integrity adaptations that teacher teams develop to implement the evidence-based practice (first, in ways that make best use of school and classroom resources, and second, in ways that fit the background, circumstances, and learning characteristics of the students); (f) real-time measurements of student PGTs, and (g) measurements of full, sustainable implementation of the classroom, leadership, and school-wide support practices.

5. Implementation Monitoring Plans that provide the process, timelines, progress benchmarks, roles and responsibilities for data collection, analysis, and reporting for each school renewal cycle (based on the measures identified in the GSIGs).

6. Implementation Dashboards that provide school-wide and stakeholder access to the key indicators and measures of progress meeting student and adult PGTs for the focus area of school renewal over school renewal cycles. The dashboard includes post hoc and real-time data with progress monitoring benchmarks for sustainable growth over time.

Data-informed decision-making [65,66] is an important aspect of the HIL Project. A high degree of shared leadership, responsibility, and interdependence between principals and teachers and among teachers as professional co-creators of the school renewal process is critical to completing each school renewal cycle. For the 2.5 years of intervention, each school team would complete at least one full cycle, with some schools already in the second cycle.


2.3. More Details on the Implementation Process

A HIL facilitator, trained on the HIL model, worked with a school leadership team (consisting of the principal and three teacher leaders) in each school. For at least an equivalent of 32 full days a year, the facilitator worked with the school through the leadership team, using a gradual release approach when interacting with the school team. As soon as the HIL Project got started, the facilitator began to engage with a school in the experimental group, with a focus on gaining understanding of the school context and building a rapport with the principal and the teacher leadership team. The facilitator typically employed an appreciative, strengths-based approach to engaging with the school’s leadership team and staff, meeting weekly with the school leadership team. The tasks of the facilitator included organizing school renewal rounds (an adaptation of the Harvard Instructional Rounds process), assisting staff in developing school performance profiles to identify growth edges related to student literacy and adult practice, and helping the principal and teacher leaders guide the school-wide decision-making process to determine priority growth targets for the school’s first school renewal cycle. From that point forward, the facilitator guided the school leadership team through the five phases of school renewal at a pace appropriate to the school to achieve their priority growth targets. A gradual release approach was enacted with the school leadership team in order to achieve a high degree of shared ownership and efficacy among principals and teachers for the school renewal process.

During the 2.5 years, facilitators worked with schools in the experimental group, helping school leadership teams learn and apply high impact leadership practices to (a) assess their current state of student literacy development and classroom literacy practice, (b) target specific areas for growth, (c) align school-wide leadership and resources to achieve growth targets, (d) monitor implementation and student growth with real-time data (including additional sessions of school renewal rounds), and (e) monitor changes in classroom and leadership practice to ensure sustainability. In addition to the school-based renewal work, each spring, fall, and summer through March 2020, all facilitators and school
leadership teams in the experimental group came together to network around the theme of progress monitoring and celebration summits.

2.4. Intervention Fidelity

Intervention fidelity was monitored via both outputs and outcomes. Outputs included (a) the amount of time for which each facilitator worked with each school and the focus of the interactions and (b) school leadership teams’ participation in summits and follow-up activities. More importantly, outcomes included various artifacts produced along the five foci of the school renewal cycle. A database on work records was established and updated biweekly to monitor the outputs and outcomes and ensure the intervention fidelity.

2.5. Research Questions

Based on the literature review and the characteristics of the HIL Project, we formulated three questions for the current study. What is the HIL Project’s impact on principal leadership? What is the HIL Project’s impact on school leadership? What is the HIL Project’s impact on student achievement?

3. Methods

3.1. Experimental Design and Empirical Data

The HIL Project started in October 2017 with the participation of 152 schools from the U.S. Midwest state of Michigan. We focused on elementary grades, and all schools had a student free and reduced-price lunch rate higher than the state’s average. Most of these schools were in urban and rural areas. These schools were recruited when the grant proposal was submitted, and reconfirmed their commitment when the grant award was announced. Initially, the external evaluation team randomly and evenly assigned the 152 schools to the experimental group (with intervention) and the control group (business as usual) in October 2017. All 152 schools were recruited from the Reading Now Network (RNN). The membership in RNN was a constant and membership in experimental or control groups was the experimental factor. After the HIL Project began, some schools dropped out and replacement schools were sought to maintain the size of the experimental group (these replacement schools did not take part in data analysis for the lack of baseline data).

The intervention lasted 2.5 years, from October 2017 to March 2020. The first cycle of data collection was carried out in the spring semester of 2018 (as the baseline) (the collection and application of all data were approved by the Institutional Review Board or IRB). We invited all principals and teachers in the experimental and control groups to complete principal and teacher questionnaires (see Appendices A and B for sample information). The second cycle of data collection was carried out in the spring semester of 2020, when experimental schools had been treated for 2.5 years. Again, we invited all principals and teachers in the participating schools to take part in the data collection (see Appendices A and B for sample information). Different sets of principals and teachers responded to questionnaires between the two cycles of data collection. Overall, 71 (unique) principals and 1250 (unique) teachers (from 123 schools) participated in the experiment (i.e., provided data for both data collections), with overall return rates of 47.7% for principals and 45.2% for teachers. Such attrition rates were compatible in the literature. Both principals and teachers were well balanced between the two experimental conditions (see the notes of the Tables for specific information). Finally, there was no unique principal who oversaw more than one school; and among the unique teachers, we were restricted from identifying teacher leaders from each school (i.e., we were not able to calculate the number of teacher leaders except for the maximum number of $3 \times 62 = 186$ in the experimental group).

We obtained public, school-level achievement data from the Michigan Student Test of Educational Progress (M-STEP), a computer-based student assessment program designed to measure the extent to which students are mastering statewide (curriculum) content standards. These standards are developed for educators by educators and aim to ensure that students are prepared for the workplace, career training, and higher education, with a
detailed prescription of what students need to know and be able to do in certain school subjects across grade levels. We focused on state testing data (measuring school-level academic achievement) as well as data on school contextual characteristics. M-STEP took place in Spring 2017, 2018, and 2019 (see Appendix C).

3.2. Outcome Measures

Previously validated principal and teacher questionnaires were used to measure various aspects of principal leadership and school leadership. Principals responded to issues of leadership in three scales: (a) Principal Leadership, (b) Principal Data-Informed Decision-Making, and (c) school’s Orientation to School Renewal. Teachers responded to issues of leadership in three scales: (a) Principal Leadership (same as the one for principals), (b) school’s Orientation to School Renewal (same as the one for principals), and (c) Learning-Centered School Leadership. Principal Leadership measures principal leadership in general, and Principal Data-Informed Decision-Making gauges principals’ data-informed decision-making in particular. Orientation to School Renewal and Learning-Centered School Leadership measures various aspects of school leadership. Please see online supplemental materials for the items for the four scales.

All four scales were previously validated, with psychometrics published in the literature. Specifically, the scale of Principal Leadership came from the Programme for International Student Assessment (PISA), and its psychometrics was published in a technical report [67]. Studies have been published in peer-reviewed journals that validate the other three scales: Principal Data-Informed Decision-Making [68], Learning-Centered School Leadership [69], and Orientation to School Renewal [63,70,71]. These three scales were developed in various funded U.S. research projects. All these three scales were examined for construct validity via vigorous literature synthesis and confirmatory factor analysis (CFA) or multilevel CFA. The CFA results were all satisfactory by means of multiple fitting indices for each scale (see those references mentioned above). Each scale was also examined for reliability via Cronbach’s alpha, with a high reliability reported for each scale (greater than 0.90). For any given scale, construct validity remains stable because it is a result of content analysis of the literature. When a scale is applied to a similar population in the same manner as the validation studies, reliability may change because reliability is a characteristic of data at hand. Therefore, we performed reliability analysis as a part of our data analysis in the current study to ensure that our data indicate sound reliability for any scale that we employed.

We chose to use these four scales because they were aligned with the HIL Project’s emphasis on principal leadership, data-informed decision-making, learning-centered school leadership, and the school renewal process. Both principals and teachers responded to the scales of (a) Principal Leadership and (b) Orientation to School Renewal, allowing us to compare the two groups’ perceptions on both principal leadership and school leadership and yielding a full picture on the effects of the HIL Project. Administering three scales to principals and teachers, respectively, balanced between the need for more data and the burden on the respondents. Collecting data from both principals and teachers was consistent with the research practice in the literature. For example, principal leadership as reported by principals themselves [72,73] and by teachers [31,40] were both found to be related to student achievement, among other school outcomes.

The measures we established for student academic achievement came from M-STEP. M-STEP covers English Language Arts (ELA), Math, Science, and Social Studies. For each school subject, M-STEP produces for each school (by grade) a measure of proficiency rate defined as the proportion of students within a school who achieve (or perform) at or above the proficient level (M-STEP calculates a proficiency rate on each school subject for each school and stores the information in a public database on Michigan schools referred to as MI School Data). In our case, we focused on ELA because the project partnered with the RNN to recruit schools and these schools focused on literacy-related school renewal initiatives during the HIL intervention. M-STEP provided three waves of data on school
proficiency rates. The baseline data were calculated for the spring semester of 2017, the second wave of data was calculated for the spring semester of 2018, and the final wave of data was calculated for the spring semester of 2019. We originally aimed to utilize data from the spring semester of 2020 as our final wave to align with data on principals and teachers and the end of the educational experiment, but the state did not conduct testing due to COVID-19. To use the Spring 2021 data to study the carryover effects was not feasible because most of the 5th graders in 2019-2020 became 6th graders in other middle schools for the 2020-2021 school year.

Given that the educational experiment was conducted from October 2017 to March 2020 (a intervention of 2.5 years) and the student achievement data were for Spring 2017, 2018, and 2019, the test on student achievement captured the intervention effects of 1.5 years (from October 2017 to March 2019 as the intervention period, with Spring 2017 M-STEP assessment as the baseline and Spring 2019 M-STEP assessment as the end point). As to the effects on principal leadership and school leadership, the current study captured the effects of 2.5 years as designed because the second survey cycle of principals and teachers in both experimental and control groups was conducted in Spring 2020.

3.3. Control Variables

When we assessed the intervention effects on student outcomes (based on M-STEP data), we adjusted the intervention effects for time-varying characteristics of schools that students attended. These school characteristics, functioned as control variables, included school (enrollment) size, proportion of male students, proportion of minority students, and proportion of students eligible for free or reduced-price lunch. They are often considered as the essential school contextual variables [74]. This information pertaining to school context is publicly available. To align with M-STEP data in our case (i.e., 2017, 2018, and 2019), we utilized school contextual information from the same years. Because school contextual characteristics may change from year to year (e.g., school size), the above school-level variables are time-varying.

3.4. Statistical Analyses

As discussed earlier, the principal questionnaire contained three scales of (a) Principal Leadership and (b) Principal Data-Informed Decision-Making for measuring principal leadership as well as (c) Orientation to School Renewal for measuring school leadership. For each principal, the score for each scale was computed using the method of valid average. Therefore, the intervention effects were evaluated in terms of the three outcome measures reported by principals. Analysis of covariance (ANCOVA) was performed by means of multiple regression analysis to compare experimental schools with control schools on each outcome measure. In such a model, Cycle 2 data were used as the posttest measures, and Cycle 1 data were used as the pretest measures. Because we performed a total of three statistical significance tests based on the same principal data, we carried out Bonferroni adjustment to an alpha level of 0.05 in order to control for family Type I error.

Meanwhile, the teacher questionnaire contained three scales of (a) Principal Leadership for measuring principal leadership, and (b) Orientation to School Renewal and (c) Learning-Centered School Leadership for measuring school leadership. For each teacher, the score for each scale was computed using the method of valid average. Therefore, the intervention effects were evaluated in terms of the three outcome measures reported by teachers. Because a number of teachers might come from the same school, hierarchical linear modeling (HLM) was applied to address the data hierarchy of teachers (at level 1) nested within schools (at level 2). In our HLM model (which is a multilevel ANCOVA model), the teacher-level model adjusted the posttest measures with the pretest measures; and the school-level model compared experimental schools with control schools on each outcome measure. Similar to the case of principal data, we performed a total of three statistical significance tests based on the same teacher data and carried out Bonferroni adjustment to an alpha level of 0.05 in order to control for family Type I error.
Finally, we used the three waves of data on the school proficiency rate to compare the experimental and control groups as a way to evaluate the intervention effects on student outcome measures. Analytically, a two-level HLM growth model was developed with repeated measures (at level 1) nested within schools (at level 2) to examine the differences between experimental and control schools in the rate of growth in schools’ ELA academic achievement. This analysis was based on grade 3 data (in Spring 2017), grade 4 data (in Spring 2018), and grade 5 data (in Spring 2019). This design closely resembled a panel, with the following model:

$$SEL_{ii} = \pi_{0i} + \pi_{1i}Year_{ii} + \sum_{m=1}^{4} SC_{mti} + \epsilon_{ti}$$

$$\pi_{0i} = \gamma_{00} + \gamma_{01}Treatment_i + u_{0i}$$

$$\pi_{1i} = \gamma_{10} + \gamma_{11}Treatment_i + u_{1i}$$

The repeated-measures model (at the first level) set up the growth model (i.e., $SEL_{ii}$ = school ELA proficiency rate on M-STEP, $\pi_{0i}$ = initial status, and $\pi_{1i}$ = rate of growth) with adjustment over time-varying (by year) school characteristics ($SC_{mti}$, $m = 1, 2, 3, 4$) of school (enrollment) size, proportion of male students, proportion of minority students, and proportion of students eligible for free or reduced-price lunch. At the school (second) level, the model compared experimental and control schools ($Treatment$) in terms of the rate of growth (and also the initial status) on each school’s ELA proficiency rate.

4. Results

4.1. Results Based on Data Collected from Principals

When examining the intervention effects of the HIL Project based on data collected from principals, we focused on principals who had both pretest and posttest measures (Appendix A). Because of this, apart from the overall return rate of 47.7% for principals, which we reported earlier (which excluded some principals from data analysis), there were no missing data on principals who were included in our data analysis. Table 1 presents descriptive information of these principals on pretest (Cycle 1) and posttest (Cycle 2) measures on the three scales administered to principals: (a) Principal Leadership (with a measurement scale of 1 to 6), (b) Data-Informed Decision-Making (with a measurement scale of 1 to 4), and (c) Orientation to School Renewal (with a measurement scale of 1 to 6).

Table 1. Descriptive statistics on outcome measures reported by principals, Cycles 1 and 2.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Cycle 1</th>
<th>Cycle 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Principal Leadership</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>4.29</td>
<td>0.76</td>
</tr>
<tr>
<td>Control</td>
<td>4.30</td>
<td>0.68</td>
</tr>
<tr>
<td>Principal Data-Informed Decision-Making</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>2.87</td>
<td>0.48</td>
</tr>
<tr>
<td>Control</td>
<td>2.79</td>
<td>0.32</td>
</tr>
<tr>
<td>Orientation to School Renewal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>4.51</td>
<td>0.75</td>
</tr>
<tr>
<td>Control</td>
<td>4.53</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Note. Descriptive results are based on 71 schools (principals) with two cycles of data. Experimental group = 36 schools. Control group = 35 schools.

Table 2 presents the intervention effects of the HIL Project based on the three scales. First, the reliability measure was both satisfactory and consistent across data cycles for each scale. Bonferroni adjustment was applied to an alpha level of 0.05. Hedge’s $g$ was used as a measure of effect size to quantify the magnitude or strength of the intervention effects. The criteria for three thresholds are small = 0.20, moderate = 0.50, and large = 0.80.
Table 2. Intervention effects on gains in outcomes between Cycles 1 and 2 based on data collected from principals.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Effect</th>
<th>SE</th>
<th>ES</th>
<th>$R^2$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal Leadership</td>
<td>0.44 *</td>
<td>0.12</td>
<td>0.73</td>
<td>0.44</td>
<td>0.87</td>
<td>0.86</td>
</tr>
<tr>
<td>Principal Data-Informed Decision-Making</td>
<td>0.19</td>
<td>0.08</td>
<td>0.46</td>
<td>0.38</td>
<td>0.94</td>
<td>0.95</td>
</tr>
<tr>
<td>Orientation to School Renewal</td>
<td>0.22</td>
<td>0.14</td>
<td>0.37</td>
<td>0.14</td>
<td>0.94</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Note. * $p < 0.05$. Bonferroni adjustment is applied to an alpha level of 0.05. Analytical results are based on 71 schools (principals) with two cycles of data. Experimental group = 36 schools. Cohort group = 35 schools. SE = standard error. ES = effect size (Hedge’s $g$). $R^2$ = proportion of variance in each measure explained by the regression model including posttest scores as the dependent variable and pretest scores and intervention condition (experimental vs. control) as the independent variables. $\alpha_1$ and $\alpha_2$ are Cronbach’s alpha on pretest scores and posttest scores, respectively.

Based on data collected from principals, principals in the experimental group reported statistically significantly greater improvement in principal leadership (effects = 0.44 on a six-point scale) than principals in the control group. The intervention effects were nearly large (effect size = 0.73), and thus bore important practical importance. However, principals in the experimental group did not report statistically significantly greater improvement than principals in the control group either in their leadership in data-informed decision-making or in their perceptions of orientation to school renewal.

Finally, Table 2 also provides information about model performance ($R^2$ = proportion of variance in posttest measures explained by pretest measures and intervention conditions of the experimental or control group). For example, in the case of principal leadership, 44% of the variance in posttest measures was accounted for. Based on the 10% standard in evaluating $R^2$ in social sciences, all estimates of $R^2$ were considered acceptable [75].

4.2. Results Based on Data Collected from Teachers

Similarly, when we examined the intervention effects of the HIL Project on outcomes based on data collected from teachers, we focused on teachers who had both pretest and posttest measures (Appendix B). As a result, similar to the case of principal data, apart from the overall return rate of 45.2% for teachers, which we reported earlier (which excluded some teachers from data analysis), there were no missing data on teachers who were included in our data analysis. Table 3 presents descriptive information of these teachers on pretest (Cycle 1) and posttest (Cycle 2) measures on the three scales in the teacher questionnaire (Principal Leadership, Orientation to School Renewal, and Learning-Centered School Leadership, all with a measurement scale of 1 to 6). Table 4 presents comprehensive information about the intervention effects on all three scales. The reliability measure in the table was both satisfactory and consistent across data cycles for each scale. Bonferroni adjustment was applied to an alpha level of 0.05. Hedge’s $g$ was used as a measure of effect size to quantify the magnitude or strength of the intervention effects.

Based on data collected from teachers, teachers in the experimental group reported statistically significantly greater improvement than teachers in the control group in Orientation to School Renewal (effects = 0.42 on a six-point scale). The intervention effects were moderate (effect size = 0.53), and thus bore notable practical importance. Teachers in the experimental group also reported statistically significantly more improvement than teachers in the control group in Learning-Centered School Leadership (effects = 0.45 on a six-point scale), with a moderate effect size of 0.58, thus bearing notable practical importance. There was no statistically significant difference between teachers in the experimental and control groups concerning the amount of improvement in principal leadership. It should be noted that Principal Leadership measured principals’ leadership, while Orientation to School Renewal and Learning-Centered School Leadership measured school leadership. In other words, teachers perceived that the intervention had a statistically significant positive impact on school leadership, but not on principal leadership. Finally, Table 4 indicates that all estimates of $R^2$ for the models examining the intervention effects among teachers were
considered acceptable. For example, in the case of Orientation to School Renewal, 26% of
the variance in posttest measures was accounted for.

Table 3. Descriptive statistics on teacher outcome measures, Cycles 1 and 2.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Cycle 1</th>
<th>Cycle 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Principal Leadership</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>3.75</td>
<td>1.08</td>
</tr>
<tr>
<td>Control</td>
<td>3.74</td>
<td>1.18</td>
</tr>
<tr>
<td>Orientation to School Renewal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>4.42</td>
<td>0.77</td>
</tr>
<tr>
<td>Control</td>
<td>4.47</td>
<td>0.88</td>
</tr>
<tr>
<td>Learning-Centered School Leadership</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>4.08</td>
<td>0.75</td>
</tr>
<tr>
<td>Control</td>
<td>4.17</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Note. Descriptive results are based on 1250 teachers with two cycles of data. Experimental group = 648. Control group = 602.

Table 4. Intervention effects on gains in outcomes between Cycles 1 and 2 based on data collected from teachers.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Effect</th>
<th>SE</th>
<th>ES</th>
<th>R²</th>
<th>α₁</th>
<th>α₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal Leadership</td>
<td>0.26</td>
<td>0.12</td>
<td>0.23</td>
<td>0.18</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>Orientation to School Renewal</td>
<td>0.42 *</td>
<td>0.08</td>
<td>0.53</td>
<td>0.26</td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td>Learning-Centered Leadership</td>
<td>0.45 *</td>
<td>0.07</td>
<td>0.58</td>
<td>0.32</td>
<td>0.96</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Note. * p < 0.05. Bonferroni adjustment is applied to an alpha level of 0.05. Analytical results from hierarchical linear modeling (HLM) are based on 1250 teachers from 123 schools with two cycles of data. Cohort A (intervention) = 648 teachers from 62 schools. Cohort B (control) = 602 teachers from 61 schools. SE = standard error. ES = effect size (Hedge’s g). R² = proportion of overall variance in each measure explained by the HLM model with posttest scores as the dependent variable, pretest scores as the independent variable at Level 1, and intervention condition (Cohort A vs. Cohort B) as the independent variable at Level 2. α₁ and α₂ are Cronbach’s alpha on pretest scores and posttest scores, respectively.

4.3. Student Outcomes

Based on school-level student academic achievement data for the (pseudo) panel of the 3rd grade in the Spring of 2017, the 4th grade in the Spring of 2018, and the 5th grade in the Spring of 2019, we compared schools in the experimental group with schools in the control group in terms of the rate of growth in school proficiency rate across grades 3 to 5 in ELA. Table 5 presents descriptive information on student outcome measures across the three time points.

Table 5. Descriptive statistics on student measures (school-level ELA proficiency rate), 2017 to 2019.

<table>
<thead>
<tr>
<th>Measure</th>
<th>2017 Mean</th>
<th>2017 SD</th>
<th>2018 Mean</th>
<th>2018 SD</th>
<th>2019 Mean</th>
<th>2019 SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>18.95</td>
<td>7.38</td>
<td>19.69</td>
<td>7.21</td>
<td>25.95</td>
<td>9.65</td>
</tr>
<tr>
<td>Control</td>
<td>20.77</td>
<td>7.96</td>
<td>19.71</td>
<td>7.75</td>
<td>21.34</td>
<td>8.84</td>
</tr>
</tbody>
</table>

Note. Descriptive results are based on school proficiency rate calculated from M-STEP data.

Table 6 presents the intervention effects of the HIL Project on student outcomes in ELA. Experimental schools grew statistically significantly more in terms of their school proficiency rate than control schools in ELA. Annually, the experimental schools grew 2.05 points more than the control schools in terms of the percentage of students who were proficient in ELA (we did not covert the intervention effects to measures of effect size in the case of student outcomes because a percentage is usually considered to be a measure of
effect size by itself). Overall, the experimental schools grew 4.10 percentage points more than the control schools in school proficiency rate in ELA for the 1.5 years of intervention under the HIL Project. Finally, 24% of the variance among schools in the rate of growth in school proficiency rate across grades 3 to 5 in ELA was accounted for.

Table 6. Intervention effects on rate of growth in school ELA proficiency rate across years from 2017 to 2019 (grade 3 in 2017, grade 4 in 2018, and grade 5 in 2019).

<table>
<thead>
<tr>
<th>Effect</th>
<th>SE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Status (in 2017)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental vs. Control</td>
<td>−1.52</td>
<td>1.06</td>
</tr>
<tr>
<td>Rate of Growth (from 2017 to 2019)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental vs. Control</td>
<td>2.05 *</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Note. * p < 0.05. Estimates are adjusted for time-varying (by year) school characteristics of school (enrollment size, proportion of male students, proportion of minority students, and proportion of students eligible for free or reduced-price lunch. R² = proportion of variance in each parameter explained by the HLM model with repeated measures as the dependent variable and grade level as the independent variable at Level 1, and intervention condition (experimental vs. control) as the independent variable at Level 2.

5. Summary and Discussion

5.1. Summary

Principals in the experimental group reported statistically significantly more improvement in Principal Leadership than their counterparts in the control group, but there was no statistically significant difference between the two principal groups in the improvement of either Principal Data-Informed Decision-Making or Orientation to School Renewal (a school leadership measure). Teachers in the experimental group reported statistically significantly greater improvement in the two measures of school leadership (Orientation to School Renewal and Learning-Centered School Leadership), but there was no statistically significant difference between the two teacher groups in their ratings on the improvement of Principal Leadership. The above results, based on data collected from principals and teachers, captured the effects of the HIL Project in its entire duration of 2.5 years of intervention, as designed. In addition, on an annual basis, the school-level ELA proficiency rate measured by M-STEP increased by 2.05 percentage points more for schools in the experimental group than those in the control group, a result that was statistically significant. The result for the school-level ELA proficiency rate captured the effects of the HIL Project in the first 1.5 years of intervention because the administration of M-STEP at the end point of the design (Spring 2020) was canceled by the state due to COVID-19. The design of the current study and the overall findings provide some insights for the effectiveness of the leadership development programs. We now discuss what we have learned from the HIL Project about using leadership development programs as a vehicle to improve our schools.

5.2. Different Perspectives between Principals and Teachers

We confirmed the phenomenon of “different perspectives between principals and teachers” in the literature. When deciding on the scales (measures) to be used, the project team intentionally administered two identical scales (Principal Leadership and Orientation to School Renewal) to the teachers and principals. This provided us with the opportunity to compare teachers’ and principals’ perspectives in order to understand the impact of the project on Principal Leadership (at the individual level) and Orientation to School Renewal (a form of school leadership). Jacob et al.’s study found that principals in the experimental group reported statistically significantly higher scores for principal efficacy and school instructional climate than their counterparts in the control group, but there was no statistically significant difference between teachers in the experimental and control groups in rating the school instructional climate [45]. Our study has the same pattern of results as Jacob et al. There appear to be different perspectives between teachers and principals. Citing Barnes et al. [76], Jacob et al. explained the phenomenon that principals report a significant change, but their teachers do not, as follows: In leadership development
programs, principals tend to refine their professional practices rather than making big changes [45]. Therefore, teachers do not perceive changes in principal leadership. The feedback from our project’s school renewal facilitators (who were assigned by the project to each school) and the project’s formative evaluation data suggest that the explanations by Barnes et al. and Jacob et al. are plausible for our result [45,76].

It is interesting to examine the phenomenon of different perspectives between teachers and principals from teachers’ and principals’ perceptions on Orientation to School Renewal. We found that teachers in the experimental group perceived more improvement in Orientation to School Renewal than their counterparts in the control group, while there was no difference concerning improvement in Orientation to School Renewal as reported by principals in the experimental and control groups. The feedback from school renewal facilitators indicated that principals faced many demands on their time and attention, and they tended to delegate in terms of developing and implementing the literacy-focused renewal activities in their schools. Our finding once again raised the question posed in Wallace Foundation’s report [77], entitled *Searching for a Superhero: Can Principals Do It All*? There appears to be a dilemma: on the one hand, our project staff ponder about the additional mechanisms that could involve principals even more to avoid the phenomenon of different perspectives between teachers and principals regarding school leadership; on the other hand, the project staff also wonder if that is even possible, given the extensive demands on principals’ time and attention. For our field, to ponder on this dilemma is important in designing leadership development programs for school improvement.

Finally, we offer some further “educated speculation” on the reasons behind the different perspectives between teachers and principals. We consider it critical to separate a principal’s educational leadership into principal leadership and school leadership (as we did in the HIL Project). Principal leadership is individual and thus tends to be personal. Personal changes are often hidden, starting with changes in attitudes and then permeating into changes in behaviors; this occurs gradually, to the point that others can observe the changes. During such a process, perhaps only principals themselves know what has been changing, but the gradual process is likely difficult for teachers to observe. In contrast, school leadership is often collective or distributed. Collective changes are often open, with teachers involved via delegation or encouragement from principals. Such changes can be more easily experienced by teachers. But from the perspective of principals, they may not consider themselves really doing anything rather special, which effectively brings us back to the conclusions of Barnes et al. and Jacob et al. [45,76]. If empirical scrutiny supports the merits of this educated speculation, then school improvement may simply focus on school leadership for school-wide changes involving every member of a school, while leaving principals to run their course when it comes to changes in principal leadership.

The current study has important implications for educational researchers or evaluators and for designers of professional development. For developers, it appears necessary to involve both parties (principals and teachers) in the same professional development effort. Obviously, any educational reform must empower both principals and teachers if a school-wide educational change is expected. For researchers, it appears to be appropriate not to overly highlight the difference between principals and teachers in their ratings of school policies and practices. If both parties say no change, it is likely there is no change. But if one party says yes, the change is likely occurred, it is just not yet noticeable enough for both parties to observe. Researchers may want to focus on where and why one party says yes rather than pushing a narrative of differences between the two parties.

### 5.3. Improvement in ELA: The Teacher Factor

We found that schools in the experimental group had a faster growth rate in the school proficiency level in ELA than their counterparts in the control group, a finding that seems to point to the importance of involving teachers in leadership development. The study by Jacob et al. and Herrmann et al. involved principals only [45,59]; and neither study found statistically significant positive effects on student achievement. Both studies, in their
discussion sections, indicated the need to involve teachers. In contrast, the HIL Project worked with a team consisting of the principal and three teacher leaders from each school and had statistically significant positive effects on ELA achievement. Both Jacob et al.’s and Herrmann et al.’s findings and their recommendations, as well as the results of positive effects on school ELA proficiency rate and the team design feature of the current study, seem to affirm the importance of involving teachers in leadership development, not just providing leadership development to principals.

There has been a growing consensus that the association between principal leadership and student achievement is mostly indirect, or that this association is mediated by various school processes [25,31]. The findings on principals’ indirect effects in the literature, the results from studies by Jacob et al. and Herrmann et al. [45,59], and the current study’s findings (the positive effects on school ELA proficiency rate and the positive effects on teachers’ responses on the two school leadership scales for the experimental group) all point to the importance of simultaneously engaging both principals and teacher leaders in the same school in leadership development. After all, it is reasonable to consider the association between principal leadership and school outcomes to be mediated by teachers’ behaviors and practices, among other school process variables.

This section generates one implication, a big one, for policymakers. If the goal of any educational reform is to eventually improve the academic achievement of all students, then both principals and teachers must be involved. Although similar arguments were made earlier, the emphasis here is on student achievement. Teachers interact with students on a daily basis, so much so that principals’ interaction with students becomes trivial in comparison. It is important to emphasize that teachers, not principals, are the agents who directly impart knowledge and skills to students. Such a notion is certainly not new but is worthy of repeating.

5.4. Importance of the Process of Application

The finding of positive effects on school proficiency rate in ELA also seems to point to the importance of the process of application in leadership development (specifically the school renewal process in the HIL Project). As mentioned in the sampling section, schools in both the experimental and control groups were from RNN, a 20-county consortium in a Midwest state that emphasized the five best evidence-based practices in literacy education. Since the experimental and control groups of schools for the current study were both derived from RNN school districts with formal commitments to prioritize literacy improvement based on sound literacy practices, focusing on literacy was not the reason why the experimental schools grew faster in their school proficiency rate in ELA. Instead, it was the experimental schools’ emphasis on the process of application (i.e., the school renewal process) that made the difference. As described in the section on intervention, the HIL Project featured the five phases of the school renewal cycle, emphasizing the use of data to guide the development of renewal activities and ensure implementation with integrity and fidelity. Through the renewal cycle, the HIL Project focused on a sustained, long-term (2.5 years) application of practice. From this viewpoint, one possible explanation for the non-significant findings on student achievement in the works of Jacob et al. and Herrmann et al. [45,59] is that the leadership development programs did not put sufficient emphasis on application. The former study focused on training principals on 21 responsibilities, and the latter study on training in three areas of instructional leadership, human capital management, and organizational leadership. Without the emphasis on the application of the 21 responsibilities and the three focus areas, their programs might lack sufficient momentum to improve student achievement.

In the literature, there are some instances in which the process of application for a leadership development program is emphasized. For example, Vanderbilt Assessment of Leadership in Education is a scale for assessing principal leadership. It has six dimensions of core components and six dimensions of processes, generating a six-by-six framework [78,79]. The six dimensions of the process are planning, implementing, supporting, advocating,
communicating, and monitoring. In fact, as Jacob et al. admitted themselves, it is possible that leadership training for principals without contextual and sustained implementation is not enough to lead to either positive effects that can be perceived by teachers or positive effects at the student level (i.e., student achievement) [45]. The emphasis on application in practice while engaging in leadership development is consistent with the notion of “Dissertation in Practice” advocated in the Carnegie Project on the Education Doctorate [80].

The implication for designers of professional development is clear: the process of application may need to become a major pillar of any reform program. It appears that gaining knowledge and skills is one thing, but putting said knowledge and skills into practice is quite another. The key is likely how to help principals and teachers both effectively and efficiently apply what they have learned in their practice (i.e., process of application) in a sustainable way.

5.5. Enhancing Principals’ Capacity for Working with Data

The scale (measure) of Principal Data-Informed Decision-Making was based on the 11-factor framework in Marzano’s work [81], What Works in Schools. It initially surprised us that principals in the experimental group did not report statistically significantly greater improvement in their own data-informed decision-making, because one of the design features of the HIL Project was to use data to inform decision-making throughout the cycles of school renewal. Among the factors, principals in the experimental group reported more improvement in the areas of collegiality and professionalism and instructional strategies only. We thought about why. We believe that when principals were engaging in the HIL Project, they tended to focus on a body of work that seemed to be more immediate (urgent) and more feasible (achievable). Because the 11 factors do cover a wide terrain (e.g., parent and community involvement; a safe and orderly school environment; collegiality and professionalism; and instructional strategies), principals might be forced to focus on specific factors (i.e., choose those urgent and achievable). We were more convinced of this after the feedback from school renewal facilitators who indicated that principals had many demands on their time and attention.

We have a major implication to offer here. There is a strong need to enhance principals' capacity for handling data (for data-informed decision-making). We may have oversimplified the technical complicity that principals face in collecting (generating) data and applying (analyzing) data to inform decision-making. The link between data and decision may not be a simple straight line. Further programs promoting data-informed decision-making may need to pay much more attention to data generation and data application. Teacher leaders may play a critical role here as well, in that they may need to shoulder the burden of data generation and data application with principals. Distributed school leadership may have a case of real application here. School districts may also provide technical assistance for schools to develop a comprehensive school monitoring system for data generation and data application.

5.6. Caution and Limitation

The current study has a major limitation. Although the intervention period was planned for 2.5 years (from October 2017 to March 2020), the end point of testing for student achievement, which originally scheduled by the state for April 2020, was canceled due to COVID-19. Therefore, the testing for the effects on student achievement used the data up to April 2019, and we were not able to capture the effects on student achievement up to the end of the project. Meanwhile, there is also a major area of caution in the current study. Essentially, the HIL Project was carried out entirely within the context of an emphasis on literacy (because we used RNN schools). Therefore, the application of HIL principles and practices was all geared towards literacy. This caution becomes necessary when adopting HIL principles and practices for future school renewal initiatives.
Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/educsci14060600/s1, File S1: Online supplemental file for questionnaires.


Funding: This research was funded by the U.S. Department of Education, grant number U423A170077. Opinions expressed in this article reflect those of the authors not necessarily those of the granting agency.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of the University of Kentucky (protocol code: 48374 and date of approval: 9 December 2022).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The datasets presented in this article are not readily available because of the consent agreement with the study participants.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Principal Questionnaire Sample Statistics.

<table>
<thead>
<tr>
<th></th>
<th>Cycle 1</th>
<th>Cycle 2</th>
<th>Common</th>
<th>Return Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group</td>
<td>76</td>
<td>67</td>
<td>36</td>
<td>47.4%</td>
</tr>
<tr>
<td>Control group</td>
<td>73</td>
<td>72</td>
<td>35</td>
<td>47.9%</td>
</tr>
<tr>
<td>Total</td>
<td>149</td>
<td>139</td>
<td>71</td>
<td>47.7%</td>
</tr>
</tbody>
</table>

Appendix B

Appendix B.1

Table A2. Teacher Questionnaire Sample Statistics (Teacher as the Unit).

<table>
<thead>
<tr>
<th></th>
<th>Cycle 1</th>
<th>Cycle 2</th>
<th>Common</th>
<th>Return Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group</td>
<td>1447</td>
<td>1288</td>
<td>648</td>
<td>44.8%</td>
</tr>
<tr>
<td>Control group</td>
<td>1318</td>
<td>1313</td>
<td>602</td>
<td>45.7%</td>
</tr>
<tr>
<td>Total</td>
<td>2765</td>
<td>2601</td>
<td>1250</td>
<td>45.2%</td>
</tr>
</tbody>
</table>

Appendix B.2

Table A3. Teacher Questionnaire Sample Statistics (School as the Unit).

<table>
<thead>
<tr>
<th></th>
<th>Cycle 1</th>
<th>Cycle 2</th>
<th>Common</th>
<th>Return Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group</td>
<td>76</td>
<td>68</td>
<td>62</td>
<td>81.6%</td>
</tr>
<tr>
<td>Control group</td>
<td>72</td>
<td>72</td>
<td>61</td>
<td>84.7%</td>
</tr>
<tr>
<td>Total</td>
<td>148</td>
<td>140</td>
<td>123</td>
<td>83.1%</td>
</tr>
</tbody>
</table>
Appendix C

Table A4. School Sample Statistics for School-Level ELA Proficiency.

<table>
<thead>
<tr>
<th></th>
<th>E</th>
<th>C</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of schools with 3rd grade Prof. rate in Spring 2017</td>
<td>62</td>
<td>60</td>
<td>122</td>
</tr>
<tr>
<td>No. of schools with 4th grade Prof. rate in Spring 2018</td>
<td>60</td>
<td>56</td>
<td>116</td>
</tr>
<tr>
<td>No. of schools with 5th grade Prof. rate in Spring 2019</td>
<td>34</td>
<td>32</td>
<td>66</td>
</tr>
<tr>
<td>Retention rate</td>
<td>45.2%</td>
<td>46.7%</td>
<td>45.9%</td>
</tr>
</tbody>
</table>

Note: E = experimental. C = control.

References

51. Nunnery, J.A.; Ross, S.M.; Yen, C. An Examination of the Effects of a Pilot of the National Institute for School Leadership's Executive Development Program on School Performance Trends in Massachusetts; The Center for Educational Partnerships, Old Dominion University: Norfolk, VA, USA, 2010. Available online: https://www.odu.edu/content/dam/odu/offices/tec/p/docs/16582_3_MAFinal_interim_report_08_23_2010.pdf (accessed on 10 April 2024).


53. Corcoran, R.P. Preparing principals to improve student achievement. Child Youth Care Forum 2017, 46, 769–781. [CrossRef]


68. Shen, J.; Cooley, V.; Ma, X.; Reeves, P.; Burt, W.; Rainey, J.M.; Yuan, W. Data-informed decision-making on high-impact strategies: Developing and validating an instrument for Principals. J. Exp. Educ. 2012, 80, 1–25. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.