Article

Journal of School Leadership
2021, Vol. 31(4) 318–342

DOI: 10.1177/1052684620972048
journals.sagepub.com/home/jsl
SAGE

Cost-Effectiveness of Instructional Coaching: Implementing a Design-Based, Continuous Improvement Model to Advance Teacher Professional Development

David S. Knight[1] and Thomas M. Skrtic[2]

Abstract

Schools devote substantial resources to teacher professional development each year. Yet studies show much of this investment is directed toward ineffective short-term workshops that have little impact on instructional change or student outcomes. At the same time, more intensive job-embedded forms of professional learning, such as instructional coaching, require substantially more resources than traditional professional development. We report the results of a two-year study assessing the cost-effectiveness of instructional coaching through a design-based, continuous improvement research model. We study iterative, inquiry cycles in which educators collect data and make changes to the coaching model based on multiple rounds of implementation. We determined the effectiveness of coaching during each iteration by tracking the number of times teachers and coaches reached student-outcome based goals set during the coaching cycle. We assess the cost of implementing the coaching model for each of the three iterations by monitoring staff time allocations and other resource use. Results show that across five schools, the cost of the coaching intervention decreased substantially from the first iteration to the second iteration but increased moderately during the third iteration. Our findings suggest that coaching programs

[1]University of Washington, Seattle, WA, USA
[2]University of Kansas, Lawrence, KS, USA

Corresponding Author:
David S. Knight, University of Washington, Seattle, WA 98195-3600, USA.
Email: dsknight@uw.edu

can become more cost-effective over time, as coaches and teachers refine their work together. While specific design features of the study limit generalizability of our findings, the study demonstrates how improvement science or design-based research can be combined with cost-effectiveness research to improve practice in local settings.

Keywords
cost-effectiveness analysis, professional development, instructional coaching, design-based research, continuous improvement science

School districts devote millions of dollars to teacher professional development, representing 6%–9% of total operating budgets (Hill, 2015; Miles et al., 2004). Yet educators and researchers criticize current practices as ineffective and wasteful (e.g., Darling-Hammond et al., 2009; The New Teacher Project [TNTP], 2015). Many districts still primarily rely on traditional approaches that focus on short-term workshops disconnected from teachers' curriculum and pedagogy. District leaders report that one barrier to implementing intensive, job-embedded professional development such as instructional coaching is cost (Cortina, 2015). Surprisingly little research examines the cost-effectiveness of alternate forms of professional development and instructional coaching models (Foster et al., 2013).

Despite the large number of experimental studies demonstrating the effectiveness of professional development programs, many fall short of their intended impacts when implemented in local contexts. Another barrier to implementing effective professional development models stems from the experimental research standards that generate knowledge of best practices. Evidence of what works in education may not be sufficient if programs are not relevant to educators or if research findings are not translatable enough to facilitate implementation (Cohen-Vogel et al., 2015). An alternate research paradigm, referred to as the continuous improvement model, suggests that educational leaders need to know not only what works, but what works for whom and in what contexts (Bryk, 2009; Bryk et al., 2010; Cobb et al., 2013; Cohen-Vogel et al., 2015; Stufflebeam, 1968; Stufflebeam & Zhang, 2017). Thus, rather than the experimentalist's concern for average treatment effects, under the continuous improvement model—as in design-based research (Bell, 2004; Kelly, 2006)—educators and researchers jointly implement a prototype practice, iteratively study its implementation to identify potential areas for improvement, and then implement a refined, context-fitted version of the initial practice. A limited number of studies apply continuous improvement models to teacher professional development (Knight et al., 2012). No previous studies have used improvement models or design-based research techniques to analyze the cost of instructional coaching over time and identify mechanisms for improving the cost-effectiveness of coaching in local contexts.

We report the results of a two-year study assessing the cost-effectiveness of instructional coaching through a design-based, continuous improvement research model. Our study is based on iterative, rapid cycles (approximately 90 days) in which

educators implemented an instructional coaching model, made changes to the coaching model based on their own insights and feedback from teachers and researchers, and repeated the cycle. We focused specifically on how the cost-effectiveness changed over the course of the study across three implementation iterations (or cycles). We determined the effectiveness of coaching during each iteration by tracking the number of times teachers and coaches reached the goal that they set during the coaching cycle. Teachers and coaches worked together to set measurable student outcome-based, ambitious, yet attainable goals. We measured the cost of implementing the coaching model during each of the three iterations by monitoring staff time allocations and other resource use.

We found that across five instructional coaches working in five different schools, the cost of instructional coaching decreased substantially from the first iteration to the second iteration but increased moderately during the third iteration. The highest number of goals were reached during the first iteration (when the greatest amount of resources was allocated to coaching). Teachers and coaches met fewer goals during Iteration 2 but saw a slight increase during Iteration 3. For this study, we define the cost-effectiveness of instructional coaching as the total cost of the resources, on average, that were used to achieve each teacher-coach goal (i.e., the ratio of costs to effects).[1] Examples of goals include having 90% of students correctly answer an exit ticket question or having at least five students raise their hand in response to a specific discussion question. The average cost of reaching a goal during Iteration 1 was $12,072. Coaches made changes to the coaching model after the first iteration. During Iteration 2, the average cost per goal declined to $10,192, implying that the coaching model became more cost-effective. Additional changes to the coaching model for Iteration 3 likely resulted in deeper more engaged interactions between teachers and coaches. Our measure of the cost-effectiveness remained relatively low during Iteration 3, with an average cost of $10,393 per goal.

These cost-effectiveness estimates are useful for relative comparisons, but prior research does not provide guidance for gauging whether the costs per goal are high or low. To place cost findings in context, we estimate the total expenditure of the coaching program across all schools in the sample. In total, the five instructional coaches collaborated with an average of 50 teachers per academic year, and we estimate annual costs of the coaching program of approximately $4,800–$9,500 per teacher, or about 2%–4% of annual expenditures.[2] Our findings show that coaching programs can become more cost-effective over time, as coaches and teachers refine their work together. More broadly, our study demonstrates how improvement science or design-based research can be used to improve practice in local settings.

In what follows, we first describe how our study addresses an important gap in the literature on instructional coaching. We then present our conceptual framework, which applies a design-based, continuous improvement research approach and the tools of cost-effectiveness analysis to an instructional coaching model that emphasizes partnerships between teachers and coaches. The subsequent two sections describe our process of planning and implementing three separate coaching cycles (or iterations) and our procedures for collecting and analyzing data to measure costs and effects.

Finally, we present our findings and conclude by discussing the implications for policy, practice, and future research.

Background Research

Research shows traditional forms of professional development rarely lead to substantial changes in instructional practices or student outcomes (Desimone et al., 2002; Garet et al., 2001; Harris & Sass, 2011; Hill, 2015; TNTP, 2015). Desimone et al. (2002) describe traditional professional development as short-term workshops on new curriculum or pedagogical practices that are not connected to specific schools or students. In the early 2000s, scholars began to emphasize the need for teacher professional development that is job-embedded, on-going, and specifically tied to teachers' daily work (Darling-Hammond et al., 2009). Job-embedded professional development refers to learning experiences that take place within the classroom or school and relate directly to issues of practice (Croft et al., 2010; Hill & Desimone, 2018). Examples include professional learning communities, instructional rounds, and instructional coaching (City et al., 2009; DuFour, 2014). Professional learning communities are groups of teachers who develop an improvement strategy to solve educational problems at their schools, while instructional rounds add classroom observations and sometimes broaden networks beyond a specific school (City, 2011). Instructional coaching involves one-on-one and small group meetings between coaches and teachers to improve instruction.

A growing number of studies show that instructional coaching can have large impacts on instruction and student achievement (e.g., Kraft et al., 2018). Yet there is substantial variation in the quality of coaching models. Like most educational interventions, the impact of instructional coaching depends largely on aspects of implementation, including the effectiveness of the specific model and the skills, training, and support of those tasked with implementation (Atteberry & Bryk, 2011; Knight et al., 2012; Biancarosa et al., 2010; Marsh et al., 2009). Two large-scale randomized trials of instructional coaching found moderate effects on instruction, but no impacts on student achievement (Garet et al., 2008, 2011). Coaches in those studies received limited training prior to beginning their work. Kraft et al. (2018) review 60 studies assessing the causal impact of coaching on student achievement. They find that programs paired with larger group workshops are more likely to lead to increases in achievement, while a larger dosage of coaching is not associated with instructional or student outcomes, among the set of studies analyzed.

Research on coaching leaves several important questions unanswered. We highlight two in particular. First, how do coaching programs get better over time? As we discuss subsequently, the scientific paradigm that emphasizes randomized experiments is not well suited to answer this question. The second question relates to resources. Coaching requires additional instructional staff to work one-on-one with teachers and therefore requires significant investment for school districts. Knight (2012) finds the cost of instructional coaching programs in three schools ranges from $3,300 to $5,200 per

teacher, or 6 to 12 times more costly than short-term workshops. Little prior research addresses the question of whether the effects of coaching models are large enough to warrant the cost, or how the costs of a coaching program may change over time (see e.g., Knight et al., 2019).

The current study begins to address these gaps by adopting a new approach to evaluation focused on continuous improvement. We examine how the outcomes associated with coaching change over time, and what actions lead to those changes. We also track resource investments to estimate how the cost of instructional coaching changes over time. This approach allows us to compare the cost-effectiveness of instructional coaching across multiple cycles of implementation. The next section describes a conceptual framework that brings together three core concepts: continuous improvement research, cost-effectiveness analysis, and partnership-based instructional coaching.

Conceptual Framework

In this section we describe how design-based and continuous improvement research methodologies differ from the typical scientific approach. We then contrast typical evaluation research with an approach that emphasizes both effects and costs. Finally, we discuss the coaching model implemented for the current study, which has its own set of guiding principles about the teacher-coach relationship.

Experimental Research and Designed-Based, Continuous Improvement Models

Under the Educational Sciences Reform Act (2002), the Institute of Education Sciences (IES) championed random-assignment experimental research as the "gold standard" for evaluating education programs. Although its primary focus continues to be "scientifically based research" (Cohen-Vogel et al., 2015, p. 261), IES responded to increasing calls "for 'scientifically valid' and 'relevant' research" (p. 261) by initiating two grant competitions emphasizing research–practice partnerships and continuous improvement research, thus "opening the way for research methodologies that capture the context in which programs are implemented" (p. 261). Most who use and advocate for such alternative methodologies recognize the validity and relevance of random-assignment experiments for research on well-defined interventions when the question is one of causality—whether treatment A causes outcome B. However, given the importance of context in innovation implementation, stakeholders question the utility of experimental designs where the task is "supporting the decisions of state and local policymakers and practitioners" (Means & Penuel, 2005, p. 178).

Cohen-Vogel et al. (2015) and Barab and Squire (2004), respectively, contrast continuous improvement methods and design research with experimental research in terms of respective aims, foci, and researcher roles. Whereas experimental research design is

premised on control, aiming to "hold variables constant and … [research procedures] fixed" (pp. 267–268), design research and continuous improvement methods both foster informed adaptation, using "an iterative, flexible process wherein [innovation] design and research plans are revised [by researchers, policymakers, and practitioners] as the work progresses" (Cohen-Vogel et al., 2015, p. 268; Barab & Squire, 2004). In terms of focus, experimental research is concerned narrowly with average treatment effects, but improvement methodology "also *involves study of the design process*" (p. 268), and design research "focuses on understanding the messiness of real world practice, with context being a core part of the story and not an extraneous variable to be trivialized" (Barab & Squire, 2004, p. 3). Given these differences, both design-based and continuous improvement methods necessarily require reconceptualizing the roles of researchers and participants. Rather than keeping researchers outside the intervention being tested and treating participants as subjects, the continuous improvement approach "purposely *involves* the researchers in innovation design and revision" and "includes participants in the design process, involving them as equals in the work" (Cohen-Vogel et al., 2015, p. 268). Similarly, participants in design-based research "are not 'subjects' assigned to treatments but instead are treated as co-participants in both the design and even analysis" (Barab & Squire, 2004, p. 3).

Finally, perhaps the definitive contrast between experimental and continuous improvement or design-based research is their respective conceptions of the source of knowledge in research, whereas the former locates knowledge and outcomes in the intervention, where other entities can access them by implementing the intervention with fidelity, design-based research and the continuous improvement model acknowledge that knowledge also is located in people and systems (Cohen-Vogel et al., 2016; Fishman et al., 2013; Kelly, 2006). As such, by expanding the experimental question, what works, to the contextual question, what works for whom in what context (Bryk, 2009), the continuous improvement model moves beyond experimental methods that merely "isolate the average effect of a particular intervention across contexts" (Cohen-Vogel et al., 2016, p. 5) by providing rich descriptions of "the contextual conditions that are necessary for success" (p. 5). Likewise, design research "involves complex social interactions with participants sharing ideas… looking at multiple aspects of the design and developing a profile that characterizes the design in practice" (Barab & Squire, 2004, p. 4).

No prior studies of which we are aware apply cost-effectiveness analysis to continuous improvement. Cost-effectiveness analysis differs from most approaches to evaluation in that researchers collect information both about the effects of a policy or intervention and about the cost. We provide additional background on this evaluation method in the section that follows.

Costs Analysis in Education

Cost-effectiveness analysis ranks alternative policy options according to a measure of their effect on outcomes relative to their resource use (Levin, 2001; Levin et al., 2017;

Levin & Belfield, 2015). For example, a study might compare the relative cost-effectiveness of reducing class size to raising teacher salaries (e.g., Belfield & Levin, 2007). Cost-effectiveness analysis has been applied to a broad range of educational interventions, such as comprehensive school reform (Borman & Hewes, 2002), dropout prevention programs (Levin et al., 2012), and pre-school (Barnett & Masse, 2007; Belfield et al., 2006).

Costs are generally defined as the total value of resources used to achieve the intended outcome (Levin & McEwan, 2000). Resources are defined broadly and can include personnel time, curriculum materials, equipment, and classroom space. In most cases, the majority of costs of educational interventions are related to the additional staff, or the reallocation of staff time, even though such reallocation does not necessarily result in additional expenditures. Economists refer to opportunity cost as the value of a given resource, if used for other purposes. When educational interventions require teachers' salaried work time before or after school or during a planning period, for example, that time can no longer be used for other potentially productive uses and therefore results in a cost. Costs analysts use hourly salaries to estimate the price of personnel time. Outcomes of educational interventions can be assessed in a variety of ways, but many cost-effectiveness analyses use effect sizes on standardized exams (e.g., Aos & Pennucci, 2013; Levin et al., 1987). In some cases, studies focus solely on costs because measures of effectiveness are often difficult to assess (e.g., Parrish, 1994).

Coaching Principles

The instructional coaches participating in this study worked from a set of principles, or meta-theoretical assumptions, referred to as the "partnership principles" (Knight, 2007, 2011), which include equality, choice, voice, dialogue, reflection, praxis, and reciprocity. The framework is drawn from literature on organizational behavior and development, anthropology, the study of change, leadership, philosophy, and educational theory (Bernstein, 1983; Block, 1993; Bohm,1996; Buckingham & Coffman, 1999; Ryan & Deci, 2000; Freire, 1970; Isaacs, 1999; Killion & Todnem, 1991; Schein, 2009; Schön, 1991).

In our prior work, we find that goal setting is an important component of instructional coaching (e.g., Knight et al., 2012). Studies show that student-related goals help teachers decide how to organize and motivate their instructional lessons and evaluate their progress, which ultimately improves student outcomes (Cheatham & Ostrosky, 2013; Mansfield et al., 2016; Rosenholtz et al., 1986; Shogren et al., 2012). Bruce et al. (2010), for example, find that teacher goal setting is positively associated with student achievement because goal setting acts as a mediator between teacher self-efficacy and student outcomes. Their study found that professional learning for teachers only resulted in student achievement gains when teachers set goals. Goal setting is also a key component of several statewide teacher evaluation systems such

as the Texas Teacher Evaluation and Support System and the Washington Principal/Teacher Evaluation Program (Texas Education Agency, 2016; Washington Office of Superintendent of Public Instruction, 2016).

The coaches participating in this study received training in effective teaching practices related to (a) content planning, (b) formative assessment, (c) instructional practices to increase engagement and mastery of content, and (d) classroom management (Knight et al., 2012). Consistent with the continuous improvement approach, the coaching model is designed to be adaptive to local context, while maintaining core values. Although the coaches' actions were refined through our iterative, design-based continuous improvement model, three features of the coaching model are central to the theory of action and therefore do not change during implementation. First, coaches take a collaborative stance with teachers, adopting the partnership principles outlined earlier. Second, coaches and teachers set goals, track progress toward those goals, and reflect on success and challenges. We provide guidance on goal-setting processes and measurement but allow local actors to adjust the specific approaches as needed. Last, teachers ultimately determine the goals and coaches provide tools to help teachers reach those goals. Teachers are not limited on the types of goals they select or the general approach to goal setting, but coaches work with teachers to make sure goals are measurable and attainable. Coaches then draw on their own training, as well as the training that we provide in the four areas listed earlier, to assist teachers in reaching their goals. Coaches may engage in modeling, additional observation, or may share curricular resources. We describe next the design and timeline of the study and how changes to the coaching model took place over three cycles, or iterations.

Study Configuration, Timeline, and Implementation Cycles

Configuration of the Study and Timeline of Iterations

The study consisted of three separate time periods in which the coaching model was implemented, with slight modifications in each cycle of coaching. We refer to each cycle as an iteration. Table 1 shows when each iteration was implemented across schools and teachers. The first iteration began in September of 2013 and lasted until the beginning of the spring 2014 semester. Iteration 2 began shortly thereafter, during the initial weeks of the spring 2014 semester; however, the exact start date differed across teachers. The second iteration was generally shorter, with an average of about 16 weeks, compared to about 22 weeks during the first iteration. Finally, the third iteration began during the first week of the fall 2014 semester and lasted 17 weeks for all teacher-coach dyads. One instructional coach from each of the five schools participated in the study. As we show in Table 1, each coach worked with between six and ten teachers, although the number varied across iterations and across coaches (and some coaches collaborated with teachers outside the study). Throughout each iteration, we interviewed teachers and coaches to learn about the improvement process and

collected survey and outcome data from coaches to assess changes in the cost-effectiveness over time.

Table 1. Timeline of Iterations Across Schools and Teachers.

		Iteration 1		Iteration 2		Iteration 3	
School	Teacher	Start Date	Weeks	Start Date	Weeks	Start Date	Weeks
1	T1	Sept 6, 2013	24	Feb 26, 2014	13	n/a	
1	T2	Sept 6, 2013	19	Jan 18, 2014	18	n/a	
1	T3	Sept 6, 2013	24	n/a		n/a	
1	T4	Sept 6, 2013	24	n/a		n/a	
1	T5	Sept 6, 2013	24	n/a		n/a	
1	T6	n/a		Feb 25, 2014	13	n/a	
1	T7	n/a		n/a		Aug 29, 2014	17
1	T8	n/a		n/a		Aug 29, 2014	17
2	T9	Sept 6, 2013	25	n/a		n/a	
2	T10	Sept 6, 2013	25	Mar 5, 2014	12	n/a	
2	T11	Sept 6, 2013	17	Jan 9, 2014	20	n/a	
2	T12	n/a		Jan 9, 2014	20	n/a	
2	T13	n/a		n/a		Aug 29, 2014	17
2	T14	n/a		n/a		Aug 29, 2014	17
3	T15	Sept 6, 2013	24	n/a		n/a	
3	T16	Sept 6, 2013	21	Feb 4, 2014	16	n/a	
3	T17	Sept 6, 2013	24	n/a		n/a	
3	T18	Sept 6, 2013	21	Feb 4, 2014	16	n/a	
3	T19	Sept 6, 2013	24	n/a		n/a	
3	T20	Sept 6, 2013	24	n/a		n/a	
3	T21	n/a		Feb 3, 2014	16	n/a	
3	T22	n/a		Feb 27, 2014	13	n/a	
3	T23	n/a		n/a		Aug 29, 2014	17
3	T24	n/a		n/a		Aug 29, 2014	17
4	T25	Sept 6, 2013	19	Feb 19, 2014a	14	n/a	
4	T26	Sept 6, 2013	19	Feb 19, 2014	14	n/a	
4	T27	Sept 6, 2013	24	n/a		n/a	
4	T28	Sept 6, 2013	24	n/a		n/a	
4	T29	n/a		n/a		Aug 29, 2014	17
4	T30	n/a		n/a		Aug 29, 2014	17

(*continued*)

Table 1. continued

		Iteration 1		Iteration 2		Iteration 3	
School	Teacher	Start Date	Weeks	Start Date	Weeks	Start Date	Weeks
5	T31	Sept 6, 2013	23	n/a		n/a	
5	T32	Sept 6, 2013	23	n/a		n/a	
5	T33	Sept 6, 2013	23	n/a		n/a	
5	T34	Sept 6, 2013	22	n/a		n/a	
5	T35	Sept 6, 2013	17	n/a		n/a	
5	T36	n/a		Feb 10, 2014	15	n/a	
5	T37	n/a		Jan 15, 2014	19	n/a	
5	T38	n/a		n/a		Aug 29, 2014	17
5	T39	n/a		n/a		Aug 29, 2014	17

Notes. The dates shown represent the first meeting between the teacher and coach for each iteration. Not applicable (n/a) implies that the coach did not collaborate with that teacher during that particular iteration.

[a]Teacher 25 at school 4 was replaced half-way through Iteration 2 because the original teacher took an administrative leave.

Learning Through Iterative Experimentation

For each iteration, the coaches in our study implemented the coaching model as designed, and then met for two to three days after implementing the iteration to discuss strengths and weaknesses of the model. During focus group sessions, coaches shared video of themselves coaching and reviewed teacher implementation data. While reviewing data, the group identified challenges associated with implementing the instructional coaching model. When possible, the researchers and coaches collaborated to identify new strategies for moving through these challenges to improve the coaching process. For some issues, however, the group was unable to identify refinements during the sessions, and then researchers consulted the literature and/or interviewed experts to find new strategies for refining the coaching process. When improvements were identified outside of the focus groups, research team members shared the new refinements through onsite training before coaches implemented the next iteration.

For the first iteration, coaches implemented a coaching model that starts with coaches video recording teachers' instructional practices, sharing that video with teachers, and using a framework for goal setting called SMART goals (Doran, 1981). A SMART goal is one that is specific, measurable, attainable, relevant, and time-based. This goal-setting framework was used to help identify new instructional practices that teachers were interested in implementing. At the end of Iteration 1, coaches and researchers (the "study team") identified challenges with goal setting. The study team determined that teacher-focused goals, centered around how teachers implement teaching strategies, were not leading to sustained change. Study team members recognized that the goal setting process was important as it helped teachers

and coaches establish boundaries and maintain focus on a shared vision. The study team thus determined that a new framework for goal setting was needed.

Researchers reviewed literature on goal setting (e.g., Locke & Latham, 2006; McEwan et al., 2016; McKeown, 2014; Van den Bergh et al., 2014; Van Kuijk et al., 2016) and proposed a new model called PEERS goals. PEERS goals are powerful (make a big difference in students' lives); easy (simple, clear, and easy to understand); emotionally compelling (matter to teachers); reachable (have a clearly identified outcome and strategy); and student-focused (address a student achievement, behavior, or attitude outcome). For Iteration 2, the coaches implemented goal setting built around the PEERS framework.

At the end of Iteration 2, after reviewing data, and in particular as a result of reviewing video of coaching conversations recorded during Iteration 2, during focus group discussions, coaches and researchers identified the need for coaches to improve their questioning skills. Specifically, the study team agreed that the instructional coaches needed better questions and questioning skills to engage teachers in critical self-reflection. To address this issue, a researcher on the project interviewed leading coaching experts for advice on questioning skills, including Joellen Killion, Cathy Toll, Bruce Wellman, and Lucy West. After the interviews with the coaching experts were completed, we created a list of coaching questions and shared it with the instructional coaches. The coaches used the questions during coaching conversations, and in particular to set goals, monitor progress, and make adaptations to interventions so as to assist teachers to move closer to and hit their goals during Iteration 3.

At the conclusion of the third iteration, the study team met to discuss implementation of the coaching model. Our main focus during this final meeting was on how the use of new types of questions impacted coaching conversations and what next steps coaches planned on taking to improve their practice. Coaches reported that the targeted training in effective questioning helped them identify areas of improvement with teachers. Together with the use of video, coaches found that improved questioning techniques—such as focusing on open-ended questions with no specific answer in mind—improved teachers self-reflection.

Data and Methods

In this section, we describe our site selection and data collection processes and then review the analytic approach for assessing the costs and effects of the coaching model.

Site Selection

We partnered with a school district in a small city of approximately 10,000 residents in the Pacific Northwest. The district employed instructional coaches for several years but was interested in adopting a new framework to support coaches' work. At the beginning

of the 2012–2013 school year, district leaders contacted the research center with which authors of this study were affiliated and we worked with district leaders to select instructional coaches to implement the new coaching framework. At the time of the study, the district operated four elementary schools, one middle school, one comprehensive high school, and one alternative high school and served approximately 4,000 students (similar to the national average). One coach at each elementary school and one at the middle school was selected based on recommendation from the associate superintendent for instruction. Most schools staffed two full-time coaches, and selected coaches typically had more experience in their role than the other coach at the same school. Collaborating teachers were identified through a combination of teacher volunteering, coach solicitation, and the principal supporting a particular coach–teacher collaboration. We conducted monthly half-day professional development sessions for coaches during the year prior to beginning the project and throughout the study. Sessions covered the guiding principles and theory of action of the coaching model. Most sessions were conducted virtually. The district's student population is primarily low-income (about 80% of students are eligible for free and reduced-price lunch, FRL) and about 87% identify as Latinx, 12% as White, and 1% as another race/ethnicity. In our findings section, we provide additional information about contextual factors that likely shaped how and for whom the coaching cycles influenced teachers' instruction and student outcomes.

Data Collection

Implementation of the coaching cycles and data collection for this research study began at the start of the 2013–2014 school year and took place over two years. All instructional coaches in our study filled out time logs throughout each iteration that documented the total time coaches, teachers, principals, and other school staff allocated to instructional coaching practices. Coaches also tracked goals, data collected to monitor progress toward reaching goals, and the status of whether goals were reached. Coaches submitted time logs and data about goals to research assistants at the end of each school week. Additionally, a research assistant worked with teachers to log the total time allocated to instructional coaching and confirm the status of goals. Research assistants followed up with teachers and coaches to confirm time reports and goal status during monthly interviews. Finally, a research assistant made note of all physical resources (e.g., video cameras, tablets, curricular materials) that coaches used or provided to teachers.

Analytic Approach

Estimating costs. We used a methodological approach called the ingredients method to assess costs (Levin et al., 2017), which consists of two basic steps. First, we collected data on the quality and quantity of all resources required of the intervention, which included physical resources such as video cameras and instructional materials, as well as

salaried work time of teachers, coaches, principals, and coach trainers. For teachers, we included only salaried work time that could otherwise be allocated for alternate purposes or interventions such as grading student work, participating in professional learning communities, or attending other professional development. Thus, teacher time included all interactions with instructional coaches or time spent in preparation for coaching meetings that took place during planning periods, before and after school, and during non-student days. For coaches, we included all salaried work time in which the coach was actively engaged in the coaching intervention, which could include meeting directly with teachers, preparing for meetings, or completing other coaching-related activities.

To calculate the total amount of personnel time allocated to the coaching program during each iteration, we sum the total number of hours devoted to coaching for each individual teacher and coach, during each iteration, and divide by the total hours worked per year (1,440). This approach provides an estimate of the full-time equivalent (FTE) proportion of annual salaried work time devoted to the coaching program for each individual involved, for each iteration. Focus group meetings with coaches and teacher interviews were excluded from these calculations because they were conducted for research purposes and not generally part of the actual coaching model. We held these meetings outside the regular school day so that research related data collection would not disrupt implementation of the coaching model.

We tracked costs of the initial training of coaches, the time that principals allocated to the coaching intervention, and physical resources, but exclude them from the presentation of results because they represent less than 5% of total costs and do not alter our overall results. Our estimates of total annual cost, including all resources, range from approximately $4,800 to $9,500 per teacher, or about 2%–4 of district expenditures. These costs are moderate given prior studies showing that districts spend approximately 6%–9% of total expenditures on all teacher professional development activities (Miles et al., 2004).

The second step of the ingredients method involves placing a dollar value on each resource. We used the Database of Educational Resource Prices published by the Center for Benefit Cost Analysis in Education to acquire the costs of personnel. This database provides national average salaries and fringe benefits for teachers, instructional coaches, coach trainers, and principals. For example, the salary for teachers with a BA and six to nine years of experience, which is roughly the education/experience profile of teachers in our study, is $41,800 (2007–2008 dollars) plus an additional 29.5% in fringe benefits. We adjusted all salaries to 2015 dollars using the Bureau of Labor Statistics Consumer Price Index for Urban Wage earners. When we include fringe benefits and adjust salaries to present value, the annual compensation for teachers and coaches is $61,068 and $88,534, respectively. To calculate the total cost of the coaching model during each iteration, we multiply the yearly salary of teachers and coaches by the FTE proportion of time allocated to the intervention.

Estimating effectiveness. We measure outcomes of the instructional coaching model by assessing whether the teacher and coach met the goal they agreed upon during the coaching process. The idea behind this approach is that the coaching model

is more likely to positively impact instruction and student achievement when teachers are reaching a greater number of goals. Based on literature cited earlier, we assume that the greater number of goals being met at each school, and in each coaching cycle, the larger the impact of the coaching intervention.

As part of the coaching model, all goals were student outcome-based and measurable. One concern is that the quality or ambition of goals may decrease over time, which would upwardly bias our estimates of the effectiveness of coaching. We therefore monitored goals closely through each iteration. We found that the level of ambition was relatively consistent across iterations. Although the foci of goals were teacher-generated, many were similar across schools and iterations in part because coaches all received the same training on goal-setting and instructional strategies and the same coaches worked through each iteration. For example, a teacher might say they want to improve student engagement. Based on their training for this study, coaches often suggested that teachers work toward getting 80% of students engaged at three different points during the class period, using a "time-on-task" observation instrument (Reinke et al., 2008).[3] Other teachers wanted to foster deeper learning from daily lessons, and coaches often suggested that teachers work towards having 90% of students correctly answer their "exit ticket problem" at the end of class, or their "bell ringer" assignment at the beginning of class. A complete listing of the goals that teachers set during each iteration is available from the authors upon request (but excluded here due to space limitations).

We remind the reader here that our analytic approach differs from the typical experimental design, where the researchers' concerns center on causal inference, controlling for contextual factors, and removing sources of bias. Instead, our purpose is to implement an intervention, track data that provides some insights into whether the intervention might be working (without knowing with absolute certainty the cause–effect relationship), making changes to the intervention based on lessons learned, and then implementing the intervention again with informed changes (Barab & Squire, 2004; Cohen-Vogel et al., 2015; Means & Penuel, 2005). As such, we are less concerned with potential sources of bias such as the need to counterbalance our research design (e.g., Best & Kahn, 2005).

The final step of the analytic approach is to combine measures of costs and effects for each iteration. We do this by dividing the total cost of each iteration (which consists of teacher and coach salaried work time) by the number of goals reached in each iteration. Our final outcome measure is a cost-effectiveness ratio showing the cost per goal. Although this outcome measure is difficult to interpret in isolation, we use it to make comparisons across iterations so that we can track how the cost per goal changes from the first iteration through the third iteration.

Findings

We discuss findings in three sections, describing in turn the costs, measures of effectiveness, and estimates of cost-effectiveness ratios.

Costs of Instructional Coach and Teacher Time

As expected, the salaried work time of instructional coaches represented the largest share of costs of the instructional coaching program across all iterations. Coach time across each iteration is displayed in Table 2. In columns 2 and 3, we report the approximate start and end date for each school, for each iteration (dates are approximate because dates varied across teachers within schools). Column 4 shows the proportion of the school year that the iteration took place. Using the same example as before, in School 1, Iteration 1 lasted an average of 23 weeks across teachers or about 62% of the school year, given a 37-week school year. This figure is shown in the first row of column 4.

Table 2. Average Full-Time Equivalent (FTE) Coach Time Allocated to Coaching Intervention, by Iteration.

School	Start Date	End Date	Proportion of School Year	Coach FTE During Iteration	Total FTE of Iteration
			Iteration 1		
1	9/6/13	1/24/14	0.622	0.518	0.322
2	9/6/13	1/10/14	0.604	0.854	0.515
3	9/6/13	1/31/14	0.622	0.664	0.413
4	9/6/13	2/14/14	0.581	0.729	0.424
5	9/6/13	1/10/14	0.584	0.805	0.470
			Iteration 2		
1	1/31/14	5/30/14	0.396	0.387	0.153
2	1/17/14	5/30/14	0.468	0.705	0.330
3	2/7/14	5/30/14	0.412	0.476	0.196
4	2/21/14	5/30/14	0.378	0.411	0.156
5	1/17/14	5/30/14	0.459	0.643	0.295
			Iteration 3		
1	8/29/14	1/16/15	0.459	0.451	0.207
2	8/29/14	1/16/15	0.459	0.668	0.307
3	8/29/14	1/16/15	0.459	0.443	0.203
4	8/29/14	1/16/15	0.459	0.479	0.220
5	8/29/14	1/16/15	0.459	0.697	0.320

Notes. Start and end dates at each school varied across teachers within each school, so we report the approximate date each iteration began in each school. The fourth column, labeled Proportion of the school year, reports the proportion of the school year that each iteration was implemented. These figures represent averages across teachers within each school because iterations took place at slightly different time periods across teachers in the same school. For example, in School 1, Iteration 1 represented an average of 62.2% of the school year, as is reported in row 1, column 4 of this table. The next column, labeled Coach FTE during iteration, reports the proportion of salaried work time coaches devoted to coaching teachers in our study during each iteration. The product of columns 4 and 5 is the total FTE coach time allocated during each iteration, which we report in column 6.

We found substantial variation in the amount of time coaches allocated to working with teachers, which we report in Table 2. Across all three iterations, coaches spent, on average, between 39% and 85% of their salaried work time coaching the teachers who participated in our study. For example, as shown in column 5 of Table 2, during the first iteration the coach in School 1 allocated 52% of salaried work time to coaching, while the coach in School 5 allocated 81%. Information on how the proportion of coach time allocated each week changed over the course of the study is available from the authors upon request. The final column of Table 2 shows the total coach time allocated in each school, across the three iterations (this column is simply the product of columns 4 and 5). We use the figures shown in column 5 to estimate the total cost of coach time in each iteration.

In Table 3, we report the cost of coach time. The second column is the total FTE coach time in each school for Iteration 1 (the same numbers shown in the final column of Table 2). The next column shows the total cost of coach time at each school and the total cost of coach time for each iteration is reported in the final row. The total cost of coach time during Iteration 1 was far greater than that of Iteration 2 and 3. This is due in part to the longer time span of Iteration 1 as well as the amount of time coaches were allocating to coaching activities.

The other major cost ingredient of the instructional coaching model that changes across iterations is teacher time. We report the cost of teacher time in Table 4. As with coach time, the first iteration had the largest personnel costs, the second iteration had the lowest, and the third iteration was in the middle of these two. As with coach time allocations, there is at least as much variation in teacher time costs across schools within each iteration as there is between iterations. This suggests that coaching is being implemented in slightly different ways even during the same iteration. That said, when we altered the coaching model after Iteration 1 the cost decreased in all schools during the second iteration. Finally, our adjustments for the third iteration generally increased the costs of teacher time at all schools with the exception of School 3.

Table 3. Total Cost of Coach Time Allocated to Coaching Intervention, by School and Iteration.

	Iteration 1		Iteration 2		Iteration 3	
School	FTE	Total Cost	FTE	Total Cost	FTE	Total Cost
1	0.322	$28,515	0.153	$13,565	0.207	$18,359
2	0.515	$45,635	0.330	$29,240	0.307	$27,173
3	0.413	$36,564	0.196	$17,382	0.203	$18,007
4	0.424	$37,523	0.156	$13,776	0.220	$19,471
5	0.470	$41,628	0.295	$26,135	0.320	$28,339
Total	2.145	$189,864	1.131	$100,098	1.258	$111,348

Notes. The estimated total compensation (wages and fringe benefits) for instructional coaches in 2014–2015 dollars is $88,534, based on the Database of Educational Resource Prices published by the Center for Benefit Cost Analysis in Education, which reports national average wages for various educational personnel

Table 4. Total Cost of Teacher Time Allocated to Coaching Intervention, by School and Iteration.

	Iteration 1		Iteration 2		Iteration 3	
School	FTE	Total cost	Hours	Total cost	Hours	Total cost
1	0.011	$693	0.004	$237	0.012	$732
2	0.009	$565	0.004	$226	0.008	$504
3	0.018	$1,133	0.015	$932	0.013	$777
4	0.004	$264	0.002	$127	0.009	$530
5	0.010	$630	0.005	$297	0.007	$435
Total	0.054	$3,286	0.030	$1,819	0.049	$2,979

Note. The estimated total compensation (wages and fringe benefits) for teachers with a BA and 6–9 years of experience in 2014–2015 dollars is $61,068 (Levin et al., 2012).

Outcomes

Teachers successfully met many of the goals they established during the coaching cycles. During the first iteration, a total of 26 goals were set and teachers met these goals in 16 cases. Only 14 goals were attempted during Iteration 2 and of those, 10 were achieved. The lower number of goals set was partly a factor of a smaller number of teachers being coached from the first iteration to the second (23 teachers as opposed to 14). In addition, several teacher-coach dyads completed two coaching cycles and set two goals during the first iteration, whereas each teacher and coach only made it through one coaching cycle and only set one goal during Iteration 2 (recall that each iteration took roughly one semester). Coaches collaborated with only two teachers each during the third iteration, which resulted in a total of 10 teacher-coach dyads. However, a total of 15 goals were set. Teachers met 10 of those 15 goals during Iteration 3.

Cost-Effectiveness Ratios

We report the final cost-effectiveness ratios in Table 5. The second and third columns repeat the cost of coach and teacher time displayed in the prior two tables, respectively. Column 4 sums the total of cost of personnel time and the fifth column shows the results for our outcome measure, the number of goals reached. Finally, in Column 6 we report the cost-effectiveness ratios, which can be interpreted as the cost per goal achieved. During the first iteration, the total cost for each goal was $12,072. Because the coach in School 1 met 5 of the 6 goals attempted, while the cost was relatively similar to other schools, the cost per goal in School 1 was substantially lower than other schools. Iteration 2 was more cost-effective according to our measures, with a cost per goal of $10,192. Again, we see substantial variation across schools within Iteration 2. For example, because School 2 only met 1 goal, the cost per goal was far greater than other schools during the second iteration. Finally, the third iteration was slightly less cost-effective compared to the second iteration, but still more cost-effective than Iteration 1.

Table 5. Total Cost of Personnel Time Allocated to Coaching Intervention, by School and Iteration.

School	Total Cost of Coach Time	Total Cost of Teacher Time	Total Cost	Number of Goals Reached	Cost Per Goal
			Iteration 1		
1	$28,515	$693	$29,207	5	$5,841
2	$45,635	$565	$46,200	3	$15,400
3	$36,564	$1,133	$37,697	3	$12,566
4	$37,523	$264	$37,787	2	$18,894
5	$41,628	$630	$42,258	3	$14,086
Total	$189,864	$3,286	$193,150	16	$12,072
			Iteration 2		
1	$13,565	$237	$13,801	2	$6,901
2	$29,240	$226	$29,466	1	$29,466
3	$17,382	$932	$18,314	3	$6,105
4	$13,776	$127	$13,903	2	$6,952
5	$26,135	$297	$26,432	2	$13,216
Total	$100,098	$1,819	$101,917	10	$10,192
			Iteration 3		
1	$18,359	$732	$19,091	2	$9,546
2	$27,173	$504	$27,677	2	$13,838
3	$18,007	$777	$18,784	3	$6,261
4	$19,471	$530	$20,001	2	$10,001
5	$28,339	$435	$28,774	2	$14,387
Total	$111,348	$2,979	$114,328	11	$10,393

Context and Implications

Description of the Context

As part of our research, we collected qualitative and quantitative data about the local context, emphasizing features that likely shaped the outcomes of our work. In the interest of space, we omit description of context-specific factors that, while distinctive, are not likely to have influenced how and for whom the coaching cycles altered teachers' instruction and student outcomes. First, teachers in our sample reported having access to an adequate level of resources and generally felt supported in their work. Principals at all five schools arranged shared planning time among teacher grade-level teams. That educators generally felt supported likely enhanced their

willingness to work with instructional coaches and external researchers and also may have increased their willingness to implement new pedagogical approaches.

Second, the district was experiencing enrollment growth, which required hiring new teachers and other staff and opening a new elementary school in 2012–2013 (that school was in its second and third years of operation during data collection). Principals described some difficulties recruiting new teachers and significant time commitment related to new teacher on-boarding. However, the expansion of the teaching workforce created new opportunities. Newly hired teachers were eager to adopt and share pedagogical techniques that they had not learned in their preparation programs. Our findings may thus be especially relevant to contexts with larger shares of new and inexperienced teachers or growing school districts.

Third, the district experienced demographic change over the past 20 years. Veteran teachers described growth in the percent of Latinx students and data show that from 1994–1995 to 2014–2015, the number of Latinx students almost doubled (from about 1,800 to 3,500), while the number of White students declined by half (900 to 450). Most teachers identified as white and some school leaders suggested that a cultural and racial/ethnic mismatch between students and faculty could have stymied teachers' efforts to connect with students, but the "small-town" context of the district may have fostered closer social ties. Coaches were able to recruit teachers with relative ease and cited changing student demographics as a motivating factor for veteran teachers to seek out coaching opportunities. Our prior work shows that additional collaborating teachers reduce the average cost of coaching programs, but teachers are sometimes reluctant to work with instructional coaches. In short, the changing student demographics, coupled with strong ties between the district and local community, may have helped coaches identify a greater number of collaborating teachers, which both reduced the average cost of the coaching program, and likely helped promote greater impacts on instruction.

Implications

Little prior research assesses the costs and effects of instructional coaching and this study represents a preliminary effort to address this gap. Extant studies find coaching is substantially more costly than traditional professional development (Knight, 2012) but impacts on student outcomes may warrant this investment (Kane & Rosenquist, 2018). While our study does not provide a causal interpretation, results suggest that an iterative model of coaching can decrease the costs of coaching without diminishing outcomes. By using a cost model, and calculating a cost per unit of outcome, practitioners who support coaches can monitor the costs of coaching programs and track results. A simple measure such as the cost per goal reached may help school leaders focus improvement efforts that lead to better outcomes. Our results also suggest that specific kinds of changes may have more impact on coaching efficiency than others. In our case, coaches

saw the greatest increase in cost-effectiveness when they refined their approach to goal setting during the second iteration. Similar results did not occur, however, when coaches refined their questioning skills to foster more critical self-reflection. What remains unclear is how much the increased cost-effectiveness was a result of timing, coaches just improving their practice from one iteration to the next, changes in the coaching model, or changes in teachers. About half of the Iteration 2 teachers participated in Iteration 1, while Iteration 3 involved all new teachers. Future continuous improvement research that examines these factors will provide deeper understanding of how to improve instructional coaching in local contexts.

More research is also needed to identify aspects of the iterative process that are most helpful for coaches. Specifically, did coaches learn more from watching video of their own practice, seeing video of other coaches coaching, discussing video of the practices, or receiving direct instruction in specific aspects of coaching such as goal-setting or effective questioning? Relatedly, future work might explore teacher factors that moderate the impact or cost of coaching. For example, are teachers with more experience or particular dispositions, more likely to benefit from coaching? Is coaching more cost-effective in particular settings? Indeed, much more could be learned by applying the design and methods of this study to coaching models in other contexts. Finally, our outcome measure is a potential limitation of the study in that we are unable to show with certainty what kind of impact reaching a goal has on teachers and students. Future research may for example, collect additional data on teacher and student outcomes.

Conclusion

This study is the first to apply design-based improvement science methods to track the cost-effectiveness of instructional coaching. We show that school leaders can work to improve the cost-effectiveness of coaching models over time. Importantly, many of the benefits of instructional coaching, such as building positive school climate, are not easily measured. To that end, our study also demonstrates how scholars can actively engage with school practitioners through design-based continuous improvement research. We urge researchers and educational practitioners to continue to collaborate on efforts to ensure that educational investments are allocated efficiently and that all students have access to high-quality learning environments.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

David S. Knight https://orcid.org/0000-0002-4067-155X

Notes

1. Readers may be familiar with other paradigmatic perspectives of value, such as "use value" and "exchange value" (e.g., Noyes, 2016; Peters, 2002). Our definition of cost-effectiveness is drawn from the cost-effectiveness literature (see Levin et al., 2017 for an overview).
2. Estimates of the cost per teacher and expenditure as a percent of total expenditures are provided only for additional context and are not part of our cost-effectiveness analysis. Estimates are based on the approximate annual per-student expenditures of school districts in the state and year in which we worked, $11,000, and assuming average class sizes of 25 students. Our sample includes 37 teachers, but coaches reported collaborating with additional teachers outside the study.
3. In classes where a greater proportion of students were already on task (based on coaches' and teachers' assessments), teachers set more ambitious goals, such as having, for example, 85 percent of students engaged. In other words, two goals may differ in the percent of students reaching certain targets, but they may be equally ambitious, since targets are based on current levels. The 49 goals created over three iterations fit into 14 unique goals, with variations on the timing and percent of students reaching various targets. While the process for creating goals changed from Iteration 1 to 2, Iteration 2 teachers did not have substantially different needs for their students and generally created similar goals to those of Iteration 1 teachers. By Iteration 3, variations of all 14 of the unique goals had been used by other teachers in the prior two iterations.

References

Aos, S., & Pennucci, A. (2013). *K-12 class size reductions and student outcomes: A review of the evidence and benefit-cost analysis* (Document No. 13-01-2201). Washington State Institute for Public Policy.

Atteberry, A., & Bryk, A. S. (2011). Analyzing teacher participation in literacy coaching activities. *The Elementary School Journal*, *112*(2), 356–382.

Barab, S., & Squire, K. (2004). Research-based design: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1–14.

Barnett, W. S., & Masse, L. N. (2007). Comparative benefit–cost analysis of the Abecedarian program and its policy implications. *Economics of Education Review*, *26*(1), 113–125.

Belfield, C. R., Nores, M., Barnett, S., & Schweinhart, L. (2006). The high/scope perry preschool program cost–benefit analysis using data from the age-40 followup. *Journal of Human Resources*, *41*(1), 162–190.

Belfield, C. R., & Levin, H. M. (Eds.). (2007). *The price we pay: Economic and social consequences of inadequate education*. Brookings Institution Press.

Bell, P. (2004). On the theoretical breadth of design-based research in education. *Educational Psychologist*, *39*(4), 243–253.

Bernstein, R. (1983). Beyond objectivism and relativism: Science, hermeneutics, and praxis. University of Pennsylvania Press.

Best, J., & Kahn, J. (2005). *Research in education* (10th ed.). Pearson.

Biancarosa, G., Bryk, A. S., & Dexter, E. R. (2010). Assessing the value-added effects of literacy collaborative professional development on student learning. *The Elementary School Journal*, *111*(1), 7–34.

Block, P. (1993). Stewardship: Choosing Service over self-interest. *Berrett-Koehler Publishers*.

Borman, G. D., & Hewes, G. M. (2002). The long-term effects and cost-effectiveness of success for all. *Educational Evaluation and Policy Analysis*, *24*(4), 243–266.

Bohm, S. (1996) On dialogue. Routledge.

Bruce, C. D., Esmonde, I., Ross, J., Dookie, L., & Beatty, R. (2010). The effects of sustained classroom-embedded teacher professional learning on teacher efficacy and related student achievement. *Teaching and Teacher Education*, *26*(8), 1598–1608.

Bryk, A. S. (2009). Support a science of performance improvement. *Phi Delta Kappan*, *90*(8), 597–600.

Bryk, A. S., Gomez, L. M., & Grunow, A. (2010). *Getting ideas into action: Building networked improvement communities in education*. Carnegie Foundation for the Advancement of Teaching.

Buckingham, M., & Coffman, C. (1999). *First, break all the rules: What the world's. greatest managers do differently*. Simon and Schuster.

Cheatham, G. A., & Ostrosky, M. M. (2013). Goal setting during early childhood parent-teacher conferences: A comparison of three groups of parents. *Journal of Research in Childhood Education*, *27*(2), 166–189.

City, E. A. (2011). Learning from instructional rounds. *Educational Leadership*, *69*(2), 36–41.

City, E. A., Elmore, R. F., Fiarman, S. E., & Teitel, L. (2009). *Instructional rounds in education: A network approach to improving teaching and learning*. Harvard Education Press.

Cobb, P., Jackson, K., Smith, T., Sorum, M., & Henrick, E. (2013). Design research with educational systems: Investigating and supporting improvements in the quality of mathematics teaching and learning at scale. *National Society for the Study of Education Yearbook*, *112*(2), 320–349.

Cohen-Vogel, L., Cannata, M., Rutledge, S. A., & Socol, A. R. (2016). A model of continuous improvement in high schools: A process for research, innovation design, implementation, and scale. *Teachers College Record*, *118*(13), 1–26.

Cohen-Vogel, L., Tichnor-Wagner, A., Allen, D., Harrison, C., Kainz, K., Socol, A. R., & Wang, Q. (2015). Implementing educational innovations at scale: Transforming researchers into continuous improvement scientists. *Educational Policy*, *29*(1), 257–277.

Croft, A., Coggshall, J. G., Dolan, M., & Powers, E. (2010). *Job-embedded professional development: What it is, who is responsible, and how to get it done well* (Issue Brief). National Comprehensive Center for Teacher Quality. https://eric.ed.gov/?id=ED520830

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad*. School Redesign Network at Stanford University. National Staff Development Council.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, *24*(2), 81–112.

DuFour, R. (2014). Harnessing the power of PLCs. *Educational Leadership*, *71*(8), 30–35.

Educational Sciences Reform Act (2002). P.L. 107–279. https://www2.ed.gov/policy/rschstat/leg/PL107-279.pdf

Freire, P. (1970). The adult literacy process as cultural action for freedom. *Harvard Educational Review*, *40*(2), 205–225.

Fishman, B. J., Penuel, W. R., Allen, A. R., Cheng, B. H., & Sabelli, N. (2013). Design-based implementation research: An emerging model for transforming the relationship of research and practice. *National Society for the Study of Education*, *112*(2), 136–156.

Foster, J. M., Toma, E. F., & Troske, S. P. (2013). Does teacher professional development improve math and science outcomes and is it cost effective?. *Journal of Education Finance*, *38*(3), 255–275

Garet, M. S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., Uekawa, K., Falk, A., Bloom, H. S., Doolittle, F., Zhu, P., & Sztejnberg, L. (2008). *The impact of two professional development interventions on early reading instruction and achievement*. U.S. Department of Education, NCEE.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915–945.

Garet, M., Wayne, A., Stancavage, F., Taylor, J., Walters, K., Song, M., Brown, S., Hurlburt, S., Zhu, P., Sepanik, S., & Doolittle, F. (2011). *Middle school mathematics professional development impact study: Findings after the first year of implementation*. Department of Education, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics*, *95*(7), 798–812.

Hill, H. (2015). *NEPC review: The mirage: Confronting the hard truth about our quest for teacher development*. National Educational Policy Center.

Hill, K. L., & Desimone, L. M. (2018). Job embedded learning: How school leaders can use job embedded learning as a mechanism for school improvement. In S. J. Zepeda & J. A. Ponticell (Eds.), *The Wiley handbook of educational supervision* (pp. 101–129). John Wiley & Sons.

Isaacs, W. (1999). *Dialogue and the art of thinking together: A pioneering approach to communicating in business and in life*. Broadway Business.

Kane, B. D., & Rosenquist, B. (2018). Making the most of instructional coaches. *Phi Delta Kappan*, *99*(7), 21–25.

Kelly, A. E. (2006). Quality criteria for design research: Evidence and commitments. In J. Van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 166–184). Routledge.

Killion, J. P., & Todnem, G. R. (1991). A process for personal theory building. *Educational Leadership*, *48*(6), 14–16.

Knight, J. (2007). *Instructional coaching: A partnership approach to improving instruction*. Corwin Press.

Knight, J. (2011). What good coaches do. *Educational Leadership*, *69*(2), 18–22.

Knight, D. S. (2012). Assessing the cost of instructional coaching. *Journal of Education Finance*, *38*(1), 52–80.

Knight, J., Bradley, B. A., Hock, M., Skrtic, T. M., Knight, D. S., Brasseur-Hock, I., Clark, J., Ruggles, M., & Hatton, C. (2012). Record, replay, reflect: Videotaped lessons accelerate learning for teachers and coaches. *Journal of Staff Development*, *33*(2), 18–23.

Knight, D. S., Landry, S. H., Zucker, T. A., Merz, E. C., Guttentag, C. L. & Taylor, H. B. (2019). Cost-effectiveness of early childhood interventions to enhance preschool: Evidence from a

randomized experiment in Head Start centers enrolling historically underserved populations. *Journal of Policy Analysis and Management*, *38*(4), 891–917. https://doi.org/10.1002/pam.22145

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, *88*(4), 547–588.

Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science*,*15*(5), 265–268.

Levin, H. M. (2001). Waiting for Godot: Cost-effectiveness analysis in education. In R. Light (Ed.), *Evaluations that surprise*. Jossey-Bass.

Levin, H. M., Belfield, C., Hollands, F., Bowden, B. A., Cheng, H., Shand, R., Pan, Y., & Hanisch-Cerda, B. (2012). *Cost-effectiveness analysis of interventions that improve high school completion*. Center for Benefit-Cost Studies of Education, Teachers College, Columbia University. Retrieved, January 4, 2014, from http://www.ny.frb.org/research/education_seminar_series/IESHighSchoolCompletion.pdf

Levin, H. M., & Belfield, C. (2015). Guiding the development and use of cost-effectiveness analysis in education. *Journal of Research on Educational Effectiveness*, *8*(3), 400–418.

Levin, H. M., Glass, G. V., & Meister, G. (1987). A cost-effectiveness analysis of computer-assisted instruction. *Evaluation Review*, *11*(1), 50–72.

Levin, H. M., & McEwan, P. J. (2000). *Cost-effectiveness analysis: Methods and applications*. SAGE Publications.

Levin, H. M., McEwan, P. J., Belfield, C., Bowden, A. B., & Shand, R. (2017). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis*. SAGE Publications, Inc.

Mansfield, C. F., Beltman, S., Broadley, T., & Weatherby-Fell, N. (2016). Building resilience in teacher education: An evidenced informed framework. *Teaching and Teacher Education*, *54*, 77–87.

Marsh, J. A., McCombs, J., & Martorell, F. (2009). How instructional coaches support data-driven decision making: Policy implementation and effects in Florida middle schools. *Educational Policy*, *24*(6), 872–907. https://doi.org/10.1177/0895904809341467

McEwan, D., Harden, S. M., Zumbo, B. D., Sylvester, B. D., Kaulius, M., Ruissen, G. R., Dowd, A. J., & Beauchamp, M. R. (2016). The effectiveness of multi-component goal setting interventions for changing physical activity behaviour: A systematic review and meta-analysis. *Health Psychology Review*, *10*(1), 67–88.

McKeown, M. (2014). *The strategy book: How to think and act strategically to deliver outstanding results* (2nd ed.). Pearson.

Means, B., & Penuel, W. R. (2005). Research to support scaling up technology-based innovations. In C. Dede, J. Honan, & L. Peters (Eds.), *Scaling up success: Lessons from technology-based educational improvement* (pp. 176–197). Jossey-Bass.

Miles, K. H., Odden, A., Fermanich, M., & Archibald, S. (2004). Inside the black box of school district spending on professional development: Lessons from five urban districts. *Journal of Education Finance*, *30*(1), 1–26.

Noyes, A. (2016). Bringing Bourdieu to mathematics education: A response to Williams and Choudry. *Research in Mathematics Education*, *18*(1), 22–26.

Parrish, T. B. (1994). A cost analysis of alternative instructional models for limited English proficient students in California. *Journal of Education Finance*, *19*(3), 256–278.

Peters, M. (2002). Education policy research and the global knowledge economy. *Educational Philosophy and Theory*, *34*(1), 91–102.
Reinke, W. M., Lewis-Palmer, T., & Merrell, K. (2008). The classroom checkup: A classwide teacher consultation model for increasing praise and decreasing disruptive behavior. *School Psychology Review*, *37*(3), 315–332.
Rosenholtz, S. J., Bassler, O., & Hoover-Dempsey, K. (1986). Organizational conditions of teacher learning. *Teaching and Teacher Education*, *2*(2), 91–104.
Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.
Schein, E. H. (2009). The corporate culture survival guide. John Wiley & Sons.
Schön, D. A. (Ed.). (1991). The reflective turn: Case studies in and on educational practice. Teachers College Press
Shogren, K. A., Palmer, S. B., Wehmeyer, M. L., Williams-Diehm, K., & Little, T. D. (2012). Effect of intervention with the self-determined learning model of instruction on access and goal attainment. *Remedial and Special Education*, *33*(5), 320–330.
Stufflebeam, D. L., & Zhang, G. (2017). *The CIPP evaluation model: How to evaluate for improvement and accountability*. Guilford Publications.
Stufflebeam, D. L. (1968). *Evaluation as enlightenment for decision-making*. Evaluation Center, Ohio State University.
Texas Education Agency. (2016). *Teacher goal-setting and professional development (GSPD) plan sample document*. https://teachfortexas.org/Resource_Files/Evaluation_Process/GSPD_Sample_Document.pdf
The New Teacher Project. (2015). *The mirage: Confronting the hard truth about our quest for teacher development*. https://tntp.org/assets/documents/TNTP-Mirage_2015.pdf
Van den Bergh, L., Ros, A., & Beijaard, D. (2014). Improving teacher feedback during active learning: Effects of a professional development program. *American Educational Research Journal*, *51*(4), 772–809.
Van Kuijk, M. F., Deunk, M. I., Bosker, R. J., & Ritzema, E. S. (2016). Goals, data use, and instruction: The effect of a teacher professional development program on reading achievement. *School Effectiveness and School Improvement*, *27*(2), 135–156.
Washington Office of Superintendent of Public Instruction. (2016). Teacher/principal evaluation program, training modules, goal setting in educator evaluation. Author.

Author Biographies

David S. Knight, Ph.D., is an assistant professor of education finance and policy at the University of Washington College of Education. His research focuses on equity in the allocation of school resources, educator labor markets, and cost-effectiveness analysis. David received his PhD in urban education policy and master's degree in economics from the University of Southern California and bachelor's degrees in economics and anthropology from the University of Kansas.

Thomas M. Skrtic, Ph.D., is Williamson Family Distinguished Professor of Special Education at the University of Kansas. His research interests include disability, race, and class inequalities in school and society, education and special education policy and politics, and educational policy implementation.