A Mixed-Methods Investigation of Instructional Coaching for Teachers of Diverse Learners

Article *in* Urban Education · July 2014

DOI: 10.1177/0042085913481362

CITATIONS
17

READS
358

1 author:

Annela Teemant
Indiana University-Purdue University Indianapolis
22 PUBLICATIONS **133** CITATIONS

SEE PROFILE

Urban Education
2014, Vol. 49(5) 574–604

DOI: 10.1177/0042085913481362
uex.sagepub.com
SAGE

A Mixed-Methods Investigation of Instructional Coaching for Teachers of Diverse Learners

Annela Teemant[1]

Abstract

This longitudinal mixed-methods study investigates the efficacy and sustainability of instructional coaching outcomes among urban elementary teachers ($N = 36$) using focus group and quantitative pre-, post-, and 1 year-after intervention data. Coached teachers participated in a 30-hr workshop and then seven cycles of coaching (15 hr) targeting use of five research-based principles of learning—the Standards for Effective Pedagogy—for teaching diverse students. Findings demonstrate instructional coaching led to statistically significant (a) pedagogical transformation and (b) patterns of sustainability and attrition. Implications for theory, practice, and research are derived by interpreting quantitative and qualitative findings together.

Keywords

English as a Second Language (ESL), urban, culturally relevant pedagogy, teacher development, diversity, best practices, elementary school

Identifying value-added models and measures of professional development are increasingly important in an era of intense focus on improving teacher

[1]Indiana University Purdue University Indianapolis, IN, USA

Corresponding Author:
Annela Teemant, School of Education, Indiana University Purdue University Indianapolis, 902 West New York Street, ES 3121, Indianapolis, IN 46202, USA.
Email: ateemant@iupui.edu

quality (Desimone, 2009; Podgorksy & Springer, 2007). Nowhere is innovation more needed than for teachers who serve underachieving students of color, low-income students, urban students, or English Language Learners (ELLs; Hollins & Guzman, 2005). The U.S. student population is culturally, linguistically, and economically more diverse (e.g., NCES, 2010) while many teachers remain unprepared for diverse learners (Leland & Murtadha, 2011; Sleeter, 2008). Unfortunately, professional development for teachers of diverse learners is a neglected area of practice (Wei, Darling-Hammond, & Adamson, 2010) and research (Willis, 2009).

This article evaluates the efficacy of an instructional coaching intervention uniquely designed for urban teachers of low-income, multilingual, and multicultural students located in a large city with high economic disparities and in a large district with limited resources, high mobility rates, and teachers who commute into the schools where they teach. The instructional coaching intervention builds upon sociocultural theory (Vygotsky, 1978). As Tharp, Estrada, Dalton, and Yamauchi (2000, p. 44) observe, sociocultural theory values how "social, cultural, and historical processes" shape teaching and learning within the "social context of relationships." Through dialogue, teachers, and students—from very different backgrounds—can create a shared and valued context for learning (Sleeter, 2008). More knowledgeable others, such as teachers, parents, or even peers, can assist student learning through questioning, modeling, or providing feedback. "Teaching-as-assisted-performance" (Gallimore & Tharp, 1990, p. 197) takes place within the zone of proximal development (ZPD; Vygotsky, 1978). The ZPD represents the difference between what a person can do independently and what can be done with assistance. For Vygotsky (1997) the learning process requires the interactional space between the student and teacher to become active with dialogue and assistance to learn.

Sociocultural theory, therefore, provides an important foundation for professional development targeting urban educators. Tharp et al. (2000) identified research-based sociocultural principles of learning called the Standards for Effective Pedagogy (five standards). These five standards are: (a) Joint Productive Activity, a teacher and small group of students creating a shared product together; (b) Language and Literacy Development, employing sustained opportunities to read, write, or speak with assistance; (c) Contextualization, activating students' knowledge and skills from home, school, and community to learn new content; (d) Challenging Activities, providing students with performance standards, assistance, and feedback to learn cognitively challenging content; and (e) Instructional Conversation, engaging a small group of students in a sustained, student-dominated, goal-directed academic conversation by questioning for rationales and assisting learning.

The five standards instructional model builds on two essential dimensions: (a) use of small group activity centers and (b) targeted use of at least three of the five standards in learning activities (Tharp et al., 2000). Multiple studies have established the value of the five standards—individually and in combination—for improving achievement among diverse learners (e.g., Estrada, 2005; Saunders & Goldenberg, 1999; Tharp, 1982).

Doherty, Hilberg, Epaloose, and Tharp (2002) developed a five-point observation tool called the Standards Performance Continuum (SPC) to measure fidelity to the five standards instructional model (Figure 1). The "Not Observed" end of the continuum describes teacher-dominated, whole class, and behaviorist practices. The "Enacting" level operationalizes what sociocultural theory means by assisting student learning within the ZPD by being collaborative, language rich, conversational, cognitively demanding, and coconstructed.

Studies with the SPC have linked improved student achievement, especially for ELLs, to teacher fidelity to the five standards instructional model in settings where no professional development was provided (Doherty & Hilberg, 2007; Doherty, Hilberg, Pinal, & Tharp, 2003).

Using the SPC for evaluation, Teemant, Wink, and Tyra (2011) piloted use of the five standards as both an instructional model for students and an instructional coaching model with California teachers. Using a descriptive repeated-measures design, Teemant et al. (2011) found teachers ($N = 21$) successfully changed to small group configurations and an enacting (i.e., $M = 12.50 < 17.49$) rather than integrating ($M = 17.50$ to 20) level of fidelity to the model. Teachers implemented Joint Productive Activity and Language and Literacy Development with greater ease than Contextualization, Challenging Activities, or the Instructional Conversation. The intervention also closed the gap between teachers who were initially high and low implementers, demonstrating the value of instructional coaching for improving instructional practice. A subsequent study by Teemant and Hausman (2012) documented significant quasi-experimental effects and correlational evidence of achievement gains in language arts and English proficiency for students of coached teachers.

This study investigates teacher growth and responsiveness to five standards instructional coaching using a mixed-methods design. This study seeks to (a) validate the efficacy of the instructional coaching model using longitudinal quantitative data with a control group and (b) extend understanding of the sustainability of teacher change using quantitative and qualitative evidence. This study benefits coaches, policy makers, and researchers by evaluating an intervention specifically designed for urban teachers and students. It also contributes to understanding what teachers change, sustain, and value when coached using the five standards instructional model.

Figure 1. Standards performance continuum: A classroom observation rubric.

	NOT OBSERVED	EMERGING	DEVELOPING	ENACTING	INTEGRATING
General Definition:	*The standard is not observed.*	*One or more elements of the standard are enacted.*	*The teacher designs and enacts activities that demonstrate a partial enactment of the standard.*	*The teacher designs, enacts, and assists in activities that demonstrate a complete enactment of the standard.*	*The teacher designs, enacts, and assists in activities that demonstrate skillful integration of multiple standards simultaneously.*
Joint Productive Activity *Teacher and Students Producing Together*	Students work independently of one another.	Students are seated with a partner or group, AND (a) collaborate or assist one another, OR (b) are instructed in how to work in groups, OR (c) contribute individual work, not requiring collaboration, to a joint product.	The teacher and students collaborate on a joint product in a whole-class setting, OR students collaborate on a joint product in pairs or small groups.	The teacher and a small group of students collaborate on a joint product.	The teacher designs, enacts, and collaborates in joint productive activities that demonstrate skillful integration* of multiple standards simultaneously.

(continued)

Figure 1. (continued)

	NOT OBSERVED	EMERGING	DEVELOPING	ENACTING	INTEGRATING
Language & Literacy Development *Developing Language and Literacy Across the Curriculum*	Instruction is dominated by teacher talk.	(a) The teacher explicitly models appropriate language; OR (b) students engage in brief, repetitive, or drill-like reading, writing, or speaking activities; OR (c) students engage in social talk while working.	The teacher provides structured opportunities for academic language development in sustained reading, writing or speaking activities. (Sustained means at least 10 minutes. If it is a whole class arrangement, then more than 50% of the students are participating. No turn taking.)	The teacher designs and enacts instructional activities that *generate* language expression and development of 'content vocabulary,' AND *assists student language use or literacy development* through questioning, rephrasing, or modeling.	The teacher designs, enacts, and assists in language development activities that demonstrate skillful integration of multiple standards simultaneously.
Contextualization *Making Meaning – Connecting School to Students' Lives*	New information is presented in an abstract, disconnected manner.	The teacher (a) includes some aspect of students' everyday experience in instruction, OR (b) connects classroom activities by theme or builds	The teacher makes incidental connections between students' prior experience/ knowledge from home, school, or community and	The teacher integrates the new activity/academic concepts with students' prior knowledge from home, school, or community to	The teacher designs, enacts, and assists in contextualized activities that demonstrate skillful integration

(continued)

Figure 1. (continued)

	NOT OBSERVED	EMERGING	DEVELOPING	ENACTING	INTEGRATING
		on the current unit of instruction, OR (c) includes parents or community members in activities or instruction, OR (d) connects student comments to content concepts.	the new activity/ academic concepts.	connect everyday and schooled concepts.	of multiple standards simultaneously.
Challenging Activities *Teaching Complex Thinking*	Activities rely on repetition, recall, or duplication to produce factual or procedural information.	The teacher (a) accommodates students' varied ability levels, OR (b) sets and presents quality standards* for student performance, OR (c) provides students with feedback on their performance.	The teacher designs and enacts 'challenging activities' that connect instructional elements to academic content OR advance student understanding to more complex levels.	The teacher designs and enacts challenging activities with clear standards/ expectations and performance feedback, AND assists* the development of more complex thinking.	The teacher designs, enacts, and assists in challenging activities that demonstrate skillful integration of multiple standards simultaneously.

(continued)

Figure 1. (continued)

	NOT OBSERVED	EMERGING	DEVELOPING	ENACTING	INTEGRATING
Instructional Conversation *Teaching Through Conversation*	Lecture or whole-class instruction predominates.	With individuals or small groups of students, the teacher (a) responds in ways that are comfortable for students, OR (b) uses questioning, listening or rephrasing *to elicit student talk*, OR (c) converses on a nonacademic topic.	The teacher converses with a small group of students on an academic topic AND *elicits student talk* with questioning, listening, rephrasing, or modeling.	The teacher: designs and enacts an instructional conversation (IC) with a clear academic goal; listens carefully to assess and assist student understanding; AND questions students on their views, judgments, or rationales. Student talk occurs at higher rates than teacher talk.	The teacher designs, enacts, and assists in instructional conversations that demonstrate skillful integration of multiple standards simultaneously.

Instructional Coaching

Research on effective professional development (e.g., Garret, Porter, Desimone, Birman, & Yoon 2001) has established the importance of using sustained, collaborative, school-based approaches. Coaching (e.g., Joyce & Showers, 1995; Sailors & Shanklin, 2010), and instructional coaching as a particular type of coaching (Knight, 2009b), has emerged as an example of quality professional development. By definition, coaching provides teachers with individualized, continuous, and extended support from a more knowledgeable other. The collaborative coaching conversations promote growth by inviting ongoing cycles of reflection and action—or praxis (Freire, 1994)—about how to effectively implement new practices in the classroom (e.g., Knight, 2009a).

Cornett's and Knight's (2009) review of research showed coaching positively improves (a) teachers' attitudes, (b) skill transfer, (c) feelings of efficacy, and (d) student achievement. Literacy and mathematics coaching research is becoming more rigorous, moving beyond self-report data (e.g., Biancarosa, Bryk, & Dexter, 2010; Neuman & Wright, 2010; Vanderburg & Stephens, 2010). Sailors and Shanklin (2010, p. 5) assert, "While coaching may be new, it is no longer unproven." Nevertheless, coaching practices vary widely. Many questions remain about what coaching content, for what duration and intensity are the most effective for which teachers, in what setting, or for what purposes. The answers to such questions are pressing given today's politics around high-stakes accountability (Ravitch, 2010).

Instructional coaching focuses on improving classroom instruction. Instructional coaches (e.g., a peer, seasoned teacher, district leader, or external consultant) possess instructional expertise that they bring to bear in their work with teachers. Knight (2009b), for example, views instructional coaching as a partnership defined by equality, choice, empowerment and respect, authentic dialogue, reflection, praxis, and reciprocity. The coach–teacher dialogue is negotiated, evolving, and partnership specific, based in equity and shared expertise. Knight (2009b) asserts instructional coaching generally focuses on four elements: classroom management, content, instruction, or ongoing assessment.

The Five Standards Intervention

For this study, the five standards (Figure 1) defined the coaching intervention as both *process* and *pedagogical targets* for classroom teaching. The efficacy of instructional coaching was measured by the degree of fidelity on the SPC coached teachers achieved to the five standards instructional model. First,

teachers participated in a 30-hr university workshop, teaching about sociocultural theory, the five standards instructional model, the SPC observation rubric (Figure 1), and a 12-week phase in process for moving to use of small-group activity centers. The activity center organization included use of a teacher-led center (i.e., an instructional conversation) and multiple independent student-led centers. The norms, expectations, and procedures for promoting successful group work were shared (Hilberg, Chang, & Epaloose, 2003). Second, teachers participated in seven individual coaching sessions across the school year (i.e., approximately 15 contact hr). Only one instructional coach was hired as an external consultant with grant funding. She had 6 years of five standards instructional coaching experience. She was also an urban bilingual elementary teacher, National Board and Reading Recovery Certified, with a PhD in Literacy Education.

Similar to the teacher–student instructional targets, the coaching process relies on the five standards to operationalize the interaction between the coach and teacher. Following an intake interview to build rapport and establish shared expectations, each coaching cycle is a three-stage process: 30-min preconference, 45-min observation, and 30-min postconference. The preconference focuses on collaboratively planning a lesson for observation. The coach–teacher dialogue invites reflection on how to increase five standards use in the context of classroom instruction. During the 45-min observation, the coach captures evidence for SPC ratings and evidence of interactional patterns, assistance, questioning practices, and student thinking and talk. These data document how instructional choices shaped student learning, engagement, and community as well as progress incorporating the five standards. During the postconference, the coach and teacher use the observation data to assess implementation and student reaction against desired SPC targets. The coach–teacher dialogue is, therefore, defined as an Instructional Conversation: goal directed, dialogic, coconstructed, and full of meaningful assistance born out of shared classroom experience, expertise, and evaluation. The five standards guide the coaching conversation; however, teacher and student readiness determined goals for improvement on an ongoing basis. Teacher growth in coaching and student growth in learning are quickened through timely and contextualized assistance in the ZPD.

The five standards instructional coaching model has unique characteristics that support its use with urban educators. First, it is explicitly theoretical promoting use of a sociocultural pedagogy (Tharp et al., 2000) known to increase diverse learners' achievement (e.g., Doherty & Hilberg, 2007; Teemant & Hausman, 2012). Second, it is coherently theoretical in that the coach–teacher interactions mirror the assisted learning experiences idealized for teacher–student classroom interactions (Rogoff, 1990; Tharp & Gallimore,

1988). Third, the five standards instructional model is highly flexible. It works with any content area, grade level, or curriculum. The coach assists the teacher to use the sociocultural context and given resources to design small-group activities where the potential to learn is maximized by prioritizing collaboration, language use, connected learning, cognitive complexity, and teacher–student dialogue (Doherty et al., 2002). Fourth, it engenders organic teacher creativity and ownership of the curriculum rather than being prescriptive or based in mimicry. Fifth, the SPC focuses the intervention theoretically on fostering learning while practically providing quantitative accountability for teacher growth as a result of instructional coaching.

In summary, this study explores the efficacy of five standards instructional coaching as a means of promoting immediate and sustainable instructional innovation among urban teachers. The following quantitative, qualitative, and mixed-methods research questions (RQs) guide this study:

RQ1. Transformation: Do coached teachers significantly differ from control group teachers in use of the five standards as measured by individual standards and total score prior to, following, and 1 year after the intervention?

RQ2. Sustainability: Do coached teachers sustain the same level of fidelity to the five standards instructional model 1 year after the end of instructional coaching as measured by individual standards and total score?

RQ3. Teacher Response: What is the nature of teacher response to the instructional coaching professional development focused on the five standards instructional model?

RQ4. Interpretation: How does the qualitative focus group data inform understanding of quantitative instructional coaching outcomes?

Method

This embedded mixed-methods study (Creswell & Plano Clark, 2007) evaluates the efficacy of five standards instructional coaching to transform urban teacher pedagogy for diverse learners overtime. Qualitative data provides context for understanding findings of a primarily quantitative quasi-experimental study (i.e., use of a control group). Instructional coaching is expected to transform teacher pedagogy. The SPC measures fidelity to the five standards instructional model in this repeated measures design: pre- to postintervention observations (i.e., transformation) and 1-year-after-intervention observations (i.e., sustainability). Dependent variables include ratings for each

standard and total score. Focus group interviews with coached teachers post- and 1-year-after-intervention provide teacher perspectives on the intervention. The participants, data sources, and analyses are described below.

Participants

This longitudinal study, supported by U.S. Department of Education funding, involved an urban elementary school invited to become a demonstration site for ELL best practices. Following an orientation on expectations, benefits, funded activities, and informed consent, staff voted to participate (90% agreement) whether they personally planned to participate in the professional development or not. Coached teachers were paid US$2,000 for their participation.

Participants were 36 K-6 elementary teachers. Teachers were on average 45 years old with 16.5 years of experience. Teachers were predominately White (88%) and female (88%), representing every grade: *K* (14%), 1 (22%), 2 (14%), 3 (14%), 4 (14%), 5(8%), 6 (8%), and two mixed 5/6 classes (6%). During the intervention year (2008-2009), the students were 75% Hispanic and 16% African American, with small populations of White, Asian, and multiracial students. Ninety-five percent of students were on free/reduced lunch. Limited English Proficient students comprised 35% of the 422 students. During the language arts block, English as a Second Language (ESL) staff taught 90-min sheltered language arts classes for low-level ELLs, lowering the average class size from 23 to 13.3 students.

Three rounds of observations were conducted: Three 30-min preintervention (spring, 2008; $M = 85.42$ min; $SD = 7.85$), two 30-min postintervention (spring 2009; $M = 55.86$; $SD = 6.8$), and two 30-min 1-year-after-intervention observations (spring 2010; $M = 58.65$; $SD = 4.81$). Crossing three school years resulted in natural teacher variation in employment. Twenty-nine teachers were observed in spring 2008 for preintervention (11 control; 18 experimental) and 36 teachers for the spring 2009 postintervention (15 control; 21 experimental) periods. Twenty-six teachers were observed 1 year after the intervention in spring 2010 (6 control; 20 experimental). Twenty-one teachers (12 mainstream; 5 ESL; 4 Specialists) were observed for all three rounds (5 control; 16 experimental), and their data is used for the repeated measures analyses. No teachers dropped out of the study, but control group teachers who later participated in the professional development became experimental group members.

Data Sources and Analysis

Teacher fidelity to the five standards instructional model was quantitatively measured with the SPC (Figure 1). The "not observed" end of the continuum

anchors predominately whole-class, individual, rote, and abstract learning. The "enacting" end anchors sociocultural principles which value collaborative, dialogic, cognitively challenging, and assisted small-group learning. Each level of the rubric identifies teacher actions that represent increased use of sociocultural practices. The five levels of the rubric are: 0 = not observed; 1 = emerging (some element present); 2 = developing (partial enactment); 3 = enacting, meaning the sociocultural principle is fully enacted; and 4 = integrating, which is achieved when three of the five (3 × 3 rule) standards are fully enacted within a single activity. Four points are available per standard for a total score of 20 points. Hilberg (Personal communication, December 12, 2006) identified four value ranges for fidelity of implementation: (a) emerging < 7.50, (b) developing = 7.50-12.49, (c) enacting = 12.50-17.49, and (d) integrating = 17.50-20.00.

Three raters established SPC inter-rater reliability for this study. Each rater had 6 years of experience in using the instrument as instructional coaches and external evaluators. They received 5 days of SPC training from its developers (2002-2003), and later spent 4 months (2005) calibrating SPC use in coaching to achieve consensus. For this study, Case 2 Intraclass Correlation Coefficients (Shrout & Fleiss, 1979) were calculated using a two-way (Rater × Standard) mixed-effects ANOVA model (McGraw & Wong, 1996), where raters were identified as the random effect and each standard was considered a fixed effect. An average measure of reliability was used because each standard was rated eight times by three raters. The Intraclass Correlation Coefficients are considered high and appropriate for high stakes decisions (Walsh & Betz, 1990): Joint Productivity = 1.00, Language/Literacy = .84, Contextualization = .98, Challenging Activities = .97, Instructional Conversation = .96.

Five types of data analyses were conducted. First, frequencies, means, and standard deviations were calculated using SPC ratings for individual standards and total score. Second, ANOVAs were used to detect control and experimental group differences for pre-, post-, and 1-year-after-intervention observations. Third, using the 21 teachers observed for all three rounds, a within- and between-subjects repeated-measures ANOVA was used to detect group differences for individual standards and total score. The Wilks' λ value, F statistic, and partial η^2 value are reported. Effect sizes less than .20 are defined as small, less than .79 are medium, and greater than .80 are large (Cohen, 1988). As needed, Mauchly's test of sphericity resulted in corrections, where the conservative Lower-bound F statistic is then reported. Fourth, post hoc pairwise comparisons, using the least significant difference (LSD) method, identified significant differences ($p < .05$) in mean performance across observation rounds.

Fifth, recorded focus group discussions were transcribed and analyzed using content analysis techniques (Creswell, 2003). Focus group questions asked coached teachers to evaluate the instructional coaching experience, identifying strengths, surprises, necessary improvements, or positive outcomes of the program as well as points of difficulty, challenge, or resistance. Teachers were asked to identify changes in their thinking and teaching and share how students experienced the model. A professor not involved in the intervention conducted these discussions, allowing teachers the freedom to express their opinions, preferences, or concerns. Teacher responses were coded and categorized by questions and then reduced to present response patterns to the intervention. Authenticity of the qualitative data is established by (a) prolonged contact with teachers and the school, (b) triangulation of qualitative findings with classroom observation data in the interpretation phase, and (c) member-checking (Creswell, 2003). Validity for the mixed-methods section of the study was established by drawing meaningful implications from both the qualitative and quantitative datasets (Creswell & Plano Clark, 2007).

Results

Findings for research questions 1 to 3 are presented below. Research question 4, which focuses on how qualitative findings inform interpretation of quantitative data, will be addressed in the mixed-method discussion and implications section.

Teacher Transformation

RQ1 asked whether control and experimental group teachers differed significantly in use of the five standards instructional model at three points of time: pre-, post-, and 1-year-after-intervention. Table 1 presents the means and standard deviations for each standard and total score for each of the three rounds of observations. Overall, the control group, as expected, demonstrated very little change across rounds. The ranked use of individual standards varied by round for the control group with the exception of the Instructional Conversation, which was consistently the least employed strategy. Coached teachers' total score improvement (Gain Score = 10) showed them implementing the five standards instructional model at the highest or integrating level of fidelity (< 17.50) at the end of coaching, using four or more standards in activity design. A year after coaching, coached teachers sustained significant change, albeit at a lower or enacting level of fidelity (> 12.50 < 17.50), suggesting regular use of only three standards in activity design.

Table 1. Means and Standard Deviations for Five Standards by Observation Round and Group.

Five pedagogical standards		Preintervention C* (n = 11)	Preintervention E* (n = 18)	Postintervention C (n = 15)	Postintervention E (n =21)	1 year after C (n = 6)	1 year after E (n = 20)
Joint productive activity	*M*	2.33	1.76	1.53	3.55	1.58	3.00
	SD	0.49	0.57	0.93	0.92	1.24	0.97
Language/Literacy development	*M*	1.64	1.85	1.90	3.79	1.83	3.13
	SD	0.69	0.63	0.91	0.37	1.08	0.89
Contextualization	*M*	1.42	1.31	1.90	3.31	1.42	2.95
	SD	0.34	0.51	0.69	0.97	0.49	0.86
Challenging activities	*M*	1.48	1.72	1.73	3.50	1.75	2.78
	SD	0.90	0.85	0.86	0.77	1.13	1.15
Instructional conversation	*M*	0.76	0.74	1.23	3.52	1.25	2.63
	SD	0.58	0.52	1.04	0.87	1.37	1.26
Total score	*M*	7.64	7.40	8.03	17.67	7.92	14.45
	SD	2.20	1.78	2.92	3.38	5.00	4.85

Note. C = Control Group; E = Experimental or Coached Group

Preintervention. A one-way ANOVA of preintervention data revealed significant control and experimental groups differences in use of Joint Productivity, F (1, 27) = 7.64, p = .01. The control group had more specialists (e.g., Title I and Special Education), who worked with smaller groups of students (< 8) than mainstream teachers. This difference allowed the control group to more easily score at the Enacting Level for Joint Productivity. Preintervention observations revealed no other group differences. Both groups used the Instructional Conversation least. At baseline, these teachers were relatively homogeneous in their use of whole class, teacher-dominated, worksheet-driven, and decontextualized instructional practices.

Postintervention. A one-way ANOVA of postintervention data revealed statistically significant group differences for each standard and total score: (a) Joint Productivity, F (1, 34) = 41.36, p < .001; (b) Language/Literacy, F (1, 34) = 73.54, p < .001; (c) Contextualization, F (1, 34) = 23.33, p < .001; (d) Challenging Activities, F (1, 34) = 41.39, p < .001; (e) Instructional Conversation, F (1, 34) = 50.89, p < .001; (f) Total Score, F (1, 34) = 79.36, p < .001. This demonstrates instructional coaching effectively led to significant transformation of teacher pedagogy. Unlike control group teachers, coached teachers had students interacting and collaborating more with peers, receiving more differentiated assistance and feedback, and using more language

and literacy to learn cognitively challenging content at the end of instructional coaching.

One-Year-After-Intervention. A one-way ANOVA of after-intervention data revealed group differences for four of the five standards and total score: (a) Joint Productivity, $F(1, 24) = 8.65$, $p = .007$; (b) Language/Literacy, $F(1, 24) = 8.90$, $p = .006$; (c) Contextualization, $F(1, 24) = 17.18$, $p < .001$; (d) Instructional Conversation, $F(1, 24) = 5.33$, $p = .030$; and (e) Total Score, $F(1, 24) = 8.27$, $p = .008$. This is evidence that instructional coaching led to sustainable pedagogical differences between coached and noncoached groups even 1 year after coaching. Control group teachers consistently planned whole class, teacher-dominated instruction. Groups did not differ significantly in their use of Challenging Activities ($F(1, 24) = 3.68$, $p = .067$) 1 year following coaching, demonstrating coached teachers struggled to consistently use cognitively challenging activities.

Table 2. Means and Standard Deviations for Five Standards for Coached Teachers by Observation Round.

Five pedagogical standards		Preintervention	Postintervention	1 year after
Joint productive activity	*M*	1.81	3.44	2.94
	SD	0.53	1.03	1.00
Language/Literacy development	*M*	1.85	3.75	3.03
	SD	0.67	0.41	0.92
Contextualization	*M*	1.22	3.25	2.91
	SD	0.45	1.05	0.84
Challenging activities	*M*	1.67	3.44	2.69
	SD	0.84	0.81	1.20
Instructional conversation	*M*	0.77	3.44	2.56
	SD	0.51	0.96	1.25
Total score	*M*	7.34	17.31	14.13
	SD	1.88	3.66	4.94

Note. $N = 16$

Sustainability

RQ2 investigated the sustainability of coached teachers' fidelity to the five standards instructional model across time. Sixteen coached teachers participated in all three rounds of observations. Table 2 presents the means and

standard deviations for these coached teachers by round. Three patterns are noteworthy. First, Joint Productivity and Language/Literacy Development were the highest standards implemented across all three rounds, suggesting these standards were the easiest to implement. Second, standard deviations showed that coached teachers were more homogeneous during the preintervention than 1 year after the end of coaching. Teachers showed the greatest variation in their use of Challenging Activities and the Instructional Conversation 1 year after coaching, suggesting these standards are most challenging to sustain. Third, coached teachers on average dropped from the highest level of fidelity ($M = 17.31$; $SD = 3.66$; > 17.50) to the enacting level of fidelity ($M = 14.13$; $SD = 4.94$; > 12.50 < 17.49) 1 year after coaching. Teachers struggled to sustain use of high quality Instructional Conversations and Challenging Activities without ongoing coaching support.

The one-way repeated measure ANOVA revealed statistically significant teacher growth in use of the five standards individually and for total score by observation rounds with generally large effect sizes: (a) Joint Productivity Wilks' $\lambda = 0.29$, F (2, 14) = 17.24, $p < .001$, partial η^2 .71; (b) Language/Literacy Wilks' $\lambda = 0.12$, F (2, 14) = 58.94, $p < .001$, partial η^2 .89; (c) Contextualization Wilks' $\lambda = 0.17$, F (2, 14) = 34.06, $p < .001$, partial η^2 .83; (d) Challenging Activities Wilks' $\lambda = 0.30$, F (2,14) = 16.60, $p < .001$, partial η^2 .70; (e) Instructional Conversation Wilks' $\lambda = 0.15$, F (2, 14) = 41.19, $p < .001$, partial η^2 .86; and (f) Total Score Wilks' $\lambda = 0.15$, F (2, 14) = 39.13, $p < .001$, partial η^2 .85. These results provide quantitative evidence that instructional coaching transforms teachers' intentional use of sociocultural practices. The large effect sizes demonstrate that the instructional coaching intervention worked well (Coe, 2002).

LSD post hoc comparisons identified significant change by observation rounds. Appendix A contains Tables A1 to A6 (see appendix, available in the online version of this article at http://uex.sagepub.com/supplemental), with the mean differences, standard errors, and 95% confidence intervals for statistically significant ($p < .05$) pairwise comparisons in teacher performance by individual standard and Total Score. The significant mean gain scores (GS) from pre- to postintervention and pre- to 1-year-after intervention respectively are as follows: (a) Joint Productivity, 1.63 and 1.13; (b) Language/Literacy, 1.90 and 1.18; (c) Contextualization, 2.03 and 1.69; (d) Challenging Activities, 1.77 and 1.02; (e) Instructional Conversation, 2.67 and 1.79; and (f) Total Score, 9.97 and 6.78.

Coached teachers changed the most between pre- and postintervention in their use of the Instructional Conversation, followed by improved focus on Contextualization and Language/Literacy Development. Teachers changed the least in their use of Joint Productivity, which they were already using at a

higher level prior to coaching. Challenging Activities was the standard teachers grew the least in. These findings demonstrate that while teachers could easily change the organization of their classrooms—increasing use of small-group activity centers—they found it more difficult to consistently challenge and assist students in the learning process.

The mean decline scores between post- and 1-year-after-intervention scores were statistically significant for four of the standards and Total Score: (a) Joint Productivity, –.50; (b) Language/Literacy, –.72; (c) Challenging Activities, –.75; (e) Instructional Conversation, –.88; and (f) Total Score, –3.19. (See appendix A Tables A1-A6, available in the online version of this article at http://uex.sagepub.com/supplemental). The greatest decline occurred in coached teachers' use of the Instructional Conversation and Challenging Activities, suggesting these are the most difficult standards for teachers to sustain without support. Teachers did not significantly decrease use of Contextualization; they were already implementing this standard least at postintervention (Table A3). The total score decline (M = –3.19) revealed teachers implementing at the Enacting (> 12.5 but < 17.49) rather than the Integrating (> 17.49) level of fidelity 1 year after coaching. In practice, students were still interacting and collaborating, but on less cognitively challenging activities and with less assistance and feedback.

Qualitative Findings

RQ3 investigated the nature of teacher response to five standards instructional coaching. Four major themes emerged from the 2009 postintervention focus group: the importance of (a) individualized support, (b) student learning, (c) valued changes, and (d) shared challenges. The 2010 teacher response to the sustainability data is presented as a fifth theme.

1. **Individualized Support**. Teachers viewed five standards instructional coaching as valuable, personalized, and practical. One teacher (#1) said, "I gained more from this than anything that I have ever done." Many teachers identified the individualized assistance as particularly valuable. A teacher (#2) observed that the instructional coach "met me where I was, and she took me . . . to the next level." Although the instructional coach used the five standards and its rubric as growth targets to guide the coaching conversations, the teachers still felt the process was individualized. One teacher (#3) said: "It wasn't generic . . . Each person could go and get coached on what they need help with or what they wanted to work on. So it was your own independent goals." Another teacher (#5) observed, "It wasn't

like she had a set agenda. It was whatever you needed at that moment." She continued, "Whether we needed to reflect about the old lesson or on another one, we could just talk." Another teacher (#6) explained how coaching conversations evolved:

> At the beginning, we were specifically talking about, "Okay if I am going to do this, what are these children doing?" Just structure. Organization. By the end, we weren't planning my lessons anymore. We were talking about how to build on this.

Teachers valued instructional coaching for following them into the classroom. One teacher (#7) explained, "I think that this program—by having a coach in the building and in your classroom watching what you are doing—really gave you the tools to implement those things that you wanted to do." Another teacher (#8) summed up five standards instructional coaching this way: "It gave us 'the how' that is missing in all of the other ones."

2. **Student Learning.** Teachers valued how instructional coaching sharpened focus on student learning: "I think that the most important thing to me was that there was an extreme benefit to the students, and now I have a better way of teaching the students and getting them to use more language," (2nd-grade ESL teacher #9). Another ESL teacher (#8) explained:

> You had to think about some of these kids . . . and how can you make them grow too. . . Who are they best suited to be with? How were they going to improve on whatever you were working on? So that was interesting to think about . . . I thought that was a good thought process.

A 1st-grade teacher (#10) valued the focus on differentiated learning. She explained:

> It made differentiating instruction across ability levels so much easier in the classroom. It reached every single child. It made me feel like I was a more effective teacher. The best lessons I had were the days that I implemented those centers.

By implementing the five standards instructional model, several teachers felt students exceeded their expectations. For example, a kindergarten teacher (#4) shared:

> I was surprised with how much my kids would come up with. It was like I would plan something, and they would shoot it out of the water. . . I would be, "So, I guess I will re-plan my center because they already knew it all . . . I have to get deeper with this." I was just shocked with how much they come up with.

An ESL teacher (#9) explained her surprise, saying:

> I had my students . . . homogenously grouped according to their reading English level and the group that had the lowest English had the best projects when they were working in their independent center. They had the best conversations with me. I could really see that this structure was benefiting my lowest students.

A 2nd-grade teacher (#3) added, "The kids felt special. It [i.e., the instructional model] was very geared toward what they know, and the words that they say, and the projects that they made. So they were excited to do those things. They were very motivated." Another teacher (#11) concurred: "They wanted to do it . . . I learned. The kids learned, and it was great."

3. **Valued Changes**. While teachers mentioned growth in organization, classroom management, differentiation, and unit planning, teachers talked most about how instructional coaching made them more reflective. One teacher (#1) said, "The program was really good at forcing me to work around my own issues that keep me from being a better teacher." Another teacher (#9) credited the coaching process with "awakening me" and "making me more self aware and reflective in my teaching."

Teachers described four specific changes they valued. First, the five standards model specifically required teachers to move away from whole-class, teacher-dominated instruction to use of multiple small-group activity centers. This change was a challenge for many teachers. A 6th-grade teacher (#11) remembers thinking, "Centers with my 6th graders? Are you kidding me? I will have chaos happening if I give them even a little bit of freedom." She explained how the coach reassured her saying, "We can do this." She continued, "We talked through it. We tried it. We got better and better." The teacher recalled later overhearing a group of students working at an independent center: "They were so on it. I was amazed. . . The way that they had a conversation changed my whole outlook on this whole process."

Teachers collectively felt small-group configurations increased student engagement. For example, a 5th-grade teacher (#2) explained that when "I was teaching in front of the class, I was bored, and they were bored. When we did the centers . . . everybody was engaged." A 2nd-grade teacher (#6) shared that students were more aware of their learning:

> The centers all revolve around the same unit, and so they are more aware . . . it is all connected . . .They are making connections throughout the unit, and to their lives, and taking ownership of what they are learning.

An ESL teacher (#9) observed that when she, "saw how all the centers were more interconnected, and . . . when I started to add follow ups [i.e. follow up centers], I could see how beneficial it was to the students."

A second change teachers valued was taking greater ownership of the curriculum. In particular, teachers described using their textbooks differently as a result of coaching. A 1st-grade teacher (#3) realized, "It doesn't have to go from point A to point B . . . [that] was hard for me to grasp." A 2nd-grade teacher (#6) shared, "I am looking for the things that I want instead of going page by page." Another teacher (#4) explained, "I really had to stretch as a teacher because I really had to search for things that were past 'the cut and paste activities' in the teacher's manual. I had to make up my own activities."

Another 1st-grade teacher (#10) shared how the five standards coaching pushed her to make her teaching from the textbook more meaningful and relevant to her students' experiences:

> Scott Foresman actually comes with suggested centers that we do. So, we think, "We're doing centers." We are doing these little suggested things. Never mind that they are not meaningful. They are boring, and the kids hate them. She [i.e., the coach] would . . . help me . . . to see how to take the most mundane and uninteresting story and find something in it to get the kids. . . excited about, so that it became meaningful. . . I found that their comprehension went up on the assessment. . . We were looking at the big concept of what we were studying. That stuck with them, and so when they did the assessment, they did better. And that really surprised me. I didn't realize the impact would be that great."

A third outcome of five standards coaching was improved questioning skills for engaging students. One teacher (#13) shared that her coach, "taught me to ask 'why' a lot more. 'Why do you think that? Why did you say that? Why was your answer better?' It has been really neat to hear their reasons." She continued, "You realize, 'Okay, I get now how they connected the two.' I would never have [understood their thinking] if I didn't ask why." Another teacher (#3) suggested, "Our thinking kind of went from black and white—we ask a question and want this one answer—to being really just open ended . . .. It made them [i.e., the students] want to speak more." Expanding on these ideas, a teacher (#10) shared: "What this helped me to do is to ask them to think. It was hard for me. And it was hard for them. . . When they got the hang of it, they just had all kinds of ideas."

A fourth change multiple teachers discussed was how the instructional conversation changed their perspectives, expectations, and abilities to understand their students. Teachers learned to respect students' experiences and

not make assumptions about students' vocabulary, literacy, or world knowledge. A 1st-grade teacher (#3) shared: They have lived a different type of life than me. They know a lot of things. They come from different backgrounds . . . What you learn through the conversations gives you confidence in the kids." A kindergarten teacher (#4) observed:

> I didn't realize how horrific some of the experiences my children have had. For example, I had one little boy talking about walking across the border and seeing people dead. . . It is very sobering as their teacher. . . I know for sure we wouldn't have known about that unless we hadn't had that questioning [i.e. as part of the instructional conversations].

Although instructional coaching was guided by the pursuit of fidelity to the five standards instructional model, the quotations in this section provide evidence of how individualized coaching shaped, altered, and redefined teachers' thinking and beliefs about a variety of instructional considerations in addition to increasing fidelity to the five standards.

4. **Shared Challenges**. Teachers identified specific challenges in the 2009 focus group that foreshadow best, perhaps, the challenges faced in sustaining use of the five standards instructional model. For example, teachers acknowledged that changing habits was difficult. Several teachers identified with one teacher's challenge (#11): "My hardest part was giving up that control and letting the kids take control." Teachers also had to overcome self-doubt. A teacher (#10) shared that being "willing to let go and not be sure of everything and be willing to make mistakes and learn from them again [was] really hard."

Another challenge foreshadows the tension teachers had navigating between district mandates and best practices for children. Many teachers considered the mandated 9-week benchmark testing a barrier to quality teaching. A teacher (#14) shared:

> District testing takes so much time out of what you were doing. You get the kids rolling, and then you have benchmarks . . . You have to give them a story that takes an adult an hour to read . . . You want to use leveled readers, but you can't because they are not going to be tested on the leveled readers.

Another teacher (#10) observed that although the school was selected to participate in the professional development, "You still have to give the test . . . You still felt like you couldn't do it the way that you wanted."

Many teachers identified the district pacing guide as a challenge to fully implementing the five standards instructional model. One teacher (#2) said, "My hardest thing was feeling restrained by the pacing guide. . . We had permission to do the program, but 'Am I going to get in trouble for doing this?' That was always there. You can't go into the whole depth of what you wanted to do." Another teacher (#8) agreed:

> I'd be better off if they just gave us the book and said, "Figure it out on your own . . ." But no, they are throwing all this other stuff at us. All of these questions that they want us to ask. . . We are all intelligent people, and we lose that ability to think on our own if we have some books telling us how we should think.

Another teacher (#15) added, "It doesn't work when we have to do their set curriculum and a story a week." Teacher #10 summed up the exchanged this way: "It makes it really hard to be creative and embrace this kind of program."

5. **Response to Sustainability Data**. In fall 2010, coached teachers were ask for their perspectives on the decline in their use of the five standards instructional model between post- and 1-year-after-intervention observation rounds. Teachers identified three contributing factors. First, teachers explained redistricting resulted in large class sizes. One teacher (#11) shared, "When I was doing the coaching, I had 18 sixth graders in my room . . . The next year, I had 30. Class size makes a huge difference."

Second, as foreshadowed in their 2009 comments, teachers felt pressured by district mandates. A seasoned teacher (#5) shared, "Last year, I was really tied to a basal structure from the district. I had to have that test by Friday. I was really set to what they asked me to do." Another teacher (#2) observed, that without the excuse of the instructional coaching program, "Now I have to do what they want me to do. And I have to turn this test in. Before I had an excuse to be the dream teacher I wanted to be."

Teachers acknowledged that the instructional coach provided a needed level of accountability. One teacher (#11) said the coach, "held me accountable for growth. Then I was a year without her . . . I tried to push myself." Another teacher (#4) shared that without the coach, "I probably wouldn't have done it, because I was too nervous. . . She was a really big help at making the centers work."

In summary, teachers' qualitative responses provide evidence that they valued individualized assistance, increased student engagement, and becoming

more reflective. They also identified increasing class size, district mandates, and personal accountability as constraining factors in sustaining growth after the end of coaching. These qualitative responses reveal what and how teachers learned from five standards instructional coaching.

Mix-Method Discussion and Implications

This section responds to research question (RQ4): How does the qualitative data (RQ3) inform understanding of quantitative instructional coaching outcomes (RQ 1 & 2)? Taken together, these findings improve understanding of five standards instructional coaching as a sociocultural process and a pedagogical outcome for urban educators and suggest implications for theory, research, and professional development practices with urban teachers.

1. **Transformation**. Coached teachers demonstrated significant quantitative growth in use of the five standards instructional model in contrast to a control group. Large effect sizes confirmed the efficacy of the instructional coaching intervention. Coached teachers achieved the highest level of fidelity with seven coaching sessions. Instructional coaching led to significant transfer of new teaching skills from a workshop to the classroom, similar to other coaching studies (e.g., Cornett & Knight, 2009; Neuman & Wright, 2010; Teemant et al., 2011).

Coached teachers' qualitative responses credit their success with the five standards instructional model to individualized and responsive assistance. Although guided broadly by measurable performance targets, the coach–teacher conferences still encouraged genuine dialogue around a variety of immediate and individualized needs: self-doubt, use of textbooks, questioning, and student engagement. The coach, therefore, assisted teacher performance in a variety of ways that are best understood at the intersection of quantitative and qualitative datasets. Teachers' instructional practices became more sociocultural and their feelings of efficacy increased. Sosa and Gomez (2012) argue teacher efficacy increases Latino students' academic resilience, and teachers' comments suggested student learning and engagement increased as a result of new practices. Similar to Vanderburg and Stephens (2010), coached teachers became different teachers: Less controlling, more open-ended, and more welcoming of student participation.

Two theoretical implications emerge from these teacher transformation findings. First, using sociocultural theory explicitly and coherently to ground the five standards coaching process and its pedagogical targets moves

instructional coaching away from what Sailors and Shanklin (2010) describe as invented craft "with little or no support from research" (p. 2). A plethora of studies make viable and operational use of sociocultural perspectives in education (e.g., Moll, 1990; Rogoff, 1990; Tharp & Gallimore, 1988). The coach and teacher role becomes clearly defined as assisting learning within a learner's ZPD. In practice, coaching and teaching enact socially coconstructed conversations that improve learning in a particular classroom with real students, needs, and challenges. The impact of theoretical and practical coherence between coaching models and teaching targets designed for urban educators should be explored in greater depth.

Second, from a sociocultural perspective, instructional coaching is more than acquiring skills for management, content, instruction, or assessment. Instructional coaching—by being dialogic, relational, and culturally situated—is also about teacher identity, power, and agency (e.g., Lewis, Enciso, & Moje, 2007; Vanderburg & Stephens, 2010). Teachers highlighted how instructional coaching changed their teacher identities and led to questioning perceived mandates. Further study is warranted to understand how instructional innovation simultaneously shapes urban teacher identity, power, and agency when mediated by an instructional coach.

2. **Sustainability**. One year after the intervention, coached teachers sustained a significant level of pedagogical change in contrast to baseline observations and control group teachers. This study provides new quantitative evidence that five standards instructional coaching leads to both immediate and sustainable teaching improvements. Nevertheless, teachers also declined one level of fidelity a year after the end of coaching.

Qualitative findings foreshadowed that sustainability of new practices and ways of working with students was enhanced for teachers by perceived improvements in student learning, engagement, and excitement, which positively raised teacher expectations, enjoyment, and efficacy. Sustainability was diminished by challenges associated with changing old habits, giving up control, overcoming self-doubt, and remaining accountable for growth. The sustainability of instructional innovation, therefore, can be limited by engrained ways of being at the heart of teacher identity, power, and agency. Nevertheless, teachers collectively identified district mandates—embodied in frequent benchmark assessments and pacing guides—as the most palpable threat to immediate and sustainable teaching improvements. Several studies echo this tension (e.g., Damico & Rosaen, 2009; Stillman & Anderson, 2011; Valli & Chambliss, 2007).

Coached teachers rightly saw themselves as part of a larger district culture with its own set of test-based expectations. Wills and Sandholtz (2009, p. 106) argue that one consequence of test-based accountability is teachers' "decisions are significantly circumscribed by contextual pressures and time demands that devalue their professional experience, judgment, and expertise." Feedback from coached teachers supports Stillman's (2011) contention that although teachers acted as "agents of change" they did so within a "tightly monitored" system of "high-stakes accountability" (p. 136). This study provides evidence that institutional pressures constrain the sustainability of instructional coaching outcomes as well as teachers' perceptions of their own power and agency as teachers.

Two implications stand out. First, teachers identified district pacing guides and testing practices as "decelerators" of teaching innovation (Cornett & Knight, 2009, p. 211). As a new urban intervention, the current study was compartmentalized as an ESL initiative within a single school. Teacher feedback underscores the importance of achieving district-wide support when scaling up use of the intervention. As Wei et al. (2010) argue, professional development must be "seamlessly linked" (p. 2) to a shared vision of district-wide reform for greater impact (Horwitz et al., 2009).

Second, use of measurable performance targets, such as the five standards, allows growth to be captured overtime without compromising teacher professionalism. As Smagorinsky (2009) observes, "Teachers should make decisions based on their good judgment" informed by "careful and systematic observations of children." The instructional coaching conversation is elevated to a discussion of professional judgment. The sociocultural principles undergirding the five standards model allowed assisted learning to be situated (Lave & Wenger, 1990; Willis, 2009), responsive (Tharp & Gallimore, 1988), and socially constructed (Vygotsky, 1978). The findings suggest five standards coaching uniquely developed the autonomy, flexibility, creativity, and cognitive complexity of urban educators in ways that richly promoted a focus on student learning. While the sustainability of instructional coaching outcomes depends on teacher professionalism, funding depends on accountability. Therefore, the five standards instructional coaching model provides both clear starting points for coaching and a measure of accountability without sacrificing needed individualized support that builds teacher professionalism (Wills & Sandholtz, 2009).

3. **Growth Patterns**. The quantitative results for individual standards revealed which elements of instruction were more or less challenging to implement. Coached teachers quickly embraced increased use of collaboration—Joint Productive Activity—and reading, writing, and

speaking—Language and Literacy Development—in the design of centers. They valued small-group activities for increasing students' language use, engagement, and learning.

Similar to findings by Teemant et al. (2011), teachers struggled to initially grow and then sustain consistency with Contextualization and Challenging Activities. Teachers incidentally activate—rather than systematically integrate—students' previous knowledge from home, school, or community in the learning of new content. While activities may require higher order thinking, teachers are less likely to provide assistance and feedback to students in completing cognitively challenging tasks. In fact, coached teachers were not significantly different from control teachers in their attention to Challenging Activities a year after coaching.

Contextualization and Challenging Activities seem to present urban teachers with curriculum challenges. According to teachers, textbook topics and stories are insufficiently engaging and pacing guides limit the time available to support struggling learners. The five standards instructional model, when implemented, makes the textbook and ready-made materials secondary to and supportive of teachers' own understanding of their learners. These mixed-method findings indicate 1 year of instructional coaching effectively changes the organization of the classroom and teacher pedagogy. Nevertheless, teachers need additional time and support to confidently design activities that connect grade-level curriculum to students' lives and provide students the assistance they need to learn complex concepts.

Teachers found the Instructional Conversation, which values student–teacher dialogue to promote learning, the most difficult to enact. Teachers used the Instructional Conversation least at baseline, made the most growth in its use with coaching, and yet still used the Instructional Conversation least by the end of coaching. The Instructional Conversation was also the standard teachers experienced the greatest decline in a year after coaching. Nevertheless, coached teachers still used the Instructional Conversation more effectively than control group teachers. Qualitative feedback suggests teachers initially delay use of the Instructional Conversation until they trusted students could work productively without teacher monitoring. Once students' self-regulation skills were in place, the challenge of the Instructional Conversation rested in the teacher's development of open-ended questioning skills. Damico and Rosaen (2009) report similar challenges when teachers moved away from teacher-directed instruction toward more open-ended discussions. Neuman and Wright (2010, p. 84) found teacher–child interactions

the "most resistant to change." McIntyre, Kyle, and Moore (2006, p. 60), who studied implementation of the Instructional Conversation by one teacher, concluded that a teacher's role becomes "more complex" in the dialogic classroom.

This study provides practical evidence that instructional coaching accelerates urban teacher growth through goal-directed cycles of praxis (Freire, 1994). These mixed-method findings lend support to Cornett's and Knight's (2009) argument that coaching positively influences teacher attitudes, skill transfer, feelings of efficacy, and—in this case—perceptions of student achievement. Nevertheless, two implications emerge for improving five standards instructional coaching. First, the 30-hr workshop and 15-hr coaching process could be improved by aligning with teachers' demonstrated growth patterns. Standards that were quickly adopted should be targeted first. Other standards—Contextualization, Challenging Activities, and the Instructional Conversation—require more depth and intensity of support. Second, sustainability is a loftier goal than transformation. This study suggests that while seven cycles of instructional coaching leads to significant change and a degree of sustainability, the "dosage" (Neuman & Wright, 2010, p. 65) or "threshold" (Desimone, 2009, p.191) is not yet high enough to ensure sustained fidelity. Extending the duration of coaching into a 2nd year seems justified based on teachers' patterns of implementation for more difficult standards.

Future Research and Limitations

Three limitation and suggestions for future research stand out. First, this study relied on one instructional coach external to the school system. Although Teemant et al. (2011) found similar positive findings with three coaches, the real test of the model will be when it is taken to scale with district coaches. Second, the instructional coach and raters for this study achieved high inter-rater reliability for the SPC. More research is needed to determine if multiple instructional coaches in multiple settings can use the SPC with high reliability when taken to scale. Third, more rigorous qualitative methods should be considered. As Vanderburg and Stephens (2010) observe, little is known about what coaches do and how coaches assist teacher development. Careful analysis of video or audiotaped coaching conversations would contribute greatly to understanding the phenomenon of instructional coaching as a lived experience, revealing what happens in the interactional space between coach and teacher.

Conclusion

In this age of high-stakes accountability, effective professional development models are central to successful urban educational reform (Desimone, 2009; Wei et al., 2010). Nevertheless, professional development targeting urban teachers of diverse learners is relatively unexplored. This mixed-methods study contributes new insights into how instructional coaching transforms urban teacher practices when the performance targets are sociocultural. In addition, the study provides evidence that the five standards instructional coaching model leads to sustained teacher change 1 year after the end of coaching. Teachers were positive about the benefits of instructional coaching as a reflective process and pedagogical outcome. The Standards Performance Continuum, as a valid and reliable classroom observational tool, provides a needed measure of accountability for instructional coaching outcomes without sacrificing teacher professionalism. This evidence suggests the five standards represent a promising focus for professional development with urban teachers of diverse learners.

Declaration of Conflicting Interests

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

Funding

This research was supported in part by a National Professional Development Grant (T195N070233) from the U.S. Department of Education's Office of English Language Acquisition.

Reference

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

Coe, R. (2002, September). *It's the effect size, stupid: What effect size is and why it is important.* Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, England. Retrieved from http://www.leeds.ac.uk/educol/documents/00002182.htm.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillside, NJ: Lawrence Erlbaum.

Cornett, J., & Knight, J. (2009). Research on coaching. In J. Knight (Ed.), *Coaching: Approaches and perspectives* (pp. 192-216). Thousand Oaks, CA: Corwin Press.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: SAGE.

Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research.* Thousand Oaks, CA: SAGE.

Damico, J. S., & Rosaen, C. L. (2009). Creating epistemological pathways to a critical citizenry: Examination of a fifth-grade discussion of freedom. *Teachers College Record, 111*, 1163-1194.

Desimone, L. (2009). Improving impact studies of teacher professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181-199.

Doherty, R. W., & Hilberg, R. S. (2007). Standards for effective pedagogy, classroom organization, English proficiency, and student achievement. *Journal of Educational Research, 101*(1), 24-35.

Doherty, R. W., Hilberg, R. S., Epaloose, G., & Tharp, R. G. (2002). Standards performance continuum: Development and validation of a measure of effective pedagogy. *Journal of Educational Research, 96*(2), 78-89.

Doherty, R. W., Hilberg, R. S., Pinal, A., & Tharp, R. G. (2003). Five standards and student achievement. *NABE Journal of Research and Practice, 1*(1), 1-24.

Estrada, P. (2005). The courage to grow: A researcher and teacher linking professional development with small-group reading instructional and reading achievement. *Research in the Teaching of English, 39*(1), 320-364.

Freire, P. (1994). *Pedagogy of the oppressed*. New York, NY: Continuum.

Gallimore, R., & Tharp, R. (1990). Teaching mind in society: Teaching, schooling, and literate discourse. In L. C. Moll (Ed.), *Vygotsky and education: Instructional implications and applications of sociohistorical psychology* (pp. 175-205). New York, NY: Cambridge University Press.

Garret, 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*, 915-945.

Hilberg, R. S., Chang, J. M., & Epaloose, G. (2003). *Designing effective activity centers for diverse learners*. Santa Cruz: Center for Research on Education, Diversity & Excellence, University of California, Santa Cruz.

Hollins, E., & Guzman, M. T. (2005). Research on preparing teachers for diverse populations. In M. Cochran-Smith & K. Zeichner (Eds.), Studying teacher education: The report of the AERA Panel on Research and Teacher Education (pp. 477-548). Mahwah, NJ: Lawrence Erlbaum.

Horwitz, A. R., Uro, G., Price-Baugh, R., Simon, C., Uzzell, R., Lewis, S., & Casserly, M. (2009, October). *Succeeding with English language learners: Lessons learned from the Great City Schools.* Retrieved from the Council of Great City Schools website http://www.cgcs.org/publications/ELL_Report09.pdf

Joyce, B., & Showers, B. (1995). *Student achievement through staff development*. White Plains, NY: Longman.

Knight, J. (Ed.). (2009a). *Coaching: Approaches and perspectives*. Thousand Oakes, CA: Corwin Press.

Knight, J. (2009b). Instructional coaching. In J. Knight (Ed.), *Coaching approaches and perspectives* (pp. 29-55). Thousand Oakes, CA: Corwin Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Leland, C. H., & Murtadha, K. (2011). Cultural discourse on the frontline: Preparing and retaining urban teachers. *Urban Education, 46*, 895-912.

Lewis, C., Enciso, P., & Moje, E. B. (2007). *Reframing sociocultural research on literacy: Identify, agency, and power.* Mahwah, NJ: Lawrence Erlbaum Associates.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coeffients. *Psychological Methods, 1*(1), 30-46.

McIntyre, E., Kyle, D. W., & Moore, G. H. (2006). A primary-grade teacher's guidance toward small-group dialogue. *Reading Research Quarterly, 41*(1), 36-66.

Moll, L. C. (Ed.). (1990). *Vygotsky and education: Instructional implications and applications of sociohistorical psychology*. New York, NY: Cambridge University Press.

National Center for Education Statistics. (NCES, 2010). *The condition of education 2010*. Washington, DC: Institute of Educational Sciences, U.S. Department of Education. Retreived from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2010028.

Neuman, S. B., & Wright, T. S. (2010). Promoting language and literacy development for early childhood educators: A mixed-methods study of coursework and coaching. *Elementary School Journal, 111*(1), 63-86.

Podgorksy, M. J., & Springer, M. G. (2007). Teacher performance pay: A review. *Journal of Policy Analysis and Management, 26*, 909-949.

Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. New York, NY: Basic Books.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York, NY: Oxford University Press.

Sailors, M., & Shanklin, N. L. (2010). Introduction: Growing evidence to support coaching in literacy and mathematics. *Elementary School Journal, 111*(1), 1-6.

Saunders, W. M., & Goldenberg, C. (1999). *The effects of instructional conversations and literature logs on the story comprehension and thematic understanding of English proficiency and limited English proficient students.* Santa Cruz, CA: Center for Research on Education, Diversity & Excellence.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428. doi:10.1037/0033-2909.86.2.420

Sleeter, C. E. (2008). Preparing White teachers for diverse students. In M. Cochran-Smith, S. Feiman-Nemser, D. J. McIntyre, & K. E. Demers (Eds.), *Handbook of research on teacher education: Enduring questions in changing contexts* (3rd ed., pp. 559-582). New York, NY: Routledge, Taylor & Francis Group and the Association of Teacher Education.

Smagorinsky, P. (2009). The cultural practice of reading and the standardized assessment of reading instruction: When incommensurate worlds collide. *Educational Researcher, 38*, 522-527.

Stillman, J. (2011). Teacher learning in an era of high-stakes accountability: Productive tension and critical professional practice. *Teachers College Record, 113*, 133-180.

Stillman, J., & Anderson, L. (2011). To follow, reject, or flip the script. *Language Arts, 89*(1), 22-37.

Sosa, T., & Gomez, K. (2012). Connecting teacher efficacy beliefs in promoting resilience to support of Latino students. *Urban Education, 47*, 876-909.

Teemant, A., & Hausman, C. (2012). *Teacher pedagogy and student achievement in the context of critical sociocultural instructional coaching*. [Unpublished raw data].

Teemant, A. Wink, J., & Tyra, S. (2011). Effects of coaching on teacher use of sociocultural instructional practices. *Teaching and Teacher Education, 27*, 683-693.

Tharp, R. G. (1982). The effective instruction of comprehension: Results and description of the Kamehameha early education program. *Reading Research Quarterly, 17*, 503-527.

Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context*. New York, NY: Cambridge University Press.

Tharp, R. G., Estrada, P., Dalton, S. S., & Yamauchi, L. (2000). *Teaching transformed: Achieving excellence, fairness, inclusion, and harmony*. Boulder, CO: Westview Press.

Valli, L., & Chambliss, M. (2007). Creating classroom cultures: One teacher, two lessons and a high-stakes test. *Anthropology and Education Quarterly, 38*(1), 57-75.

Vanderburg, M., & Stephens, D. (2010). The impact of literacy coaches: What teachers value and how teachers change. *Elementary School Journal, 111*(1), 142-163.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. (M. Cole, V. John-Steiner, S. Scribner, & E Souberman, Eds.). Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1997). Educational psychology. Boca Raton, Florida: St. Lucie Press.

Walsh, W. B., & Betz, N. E. (1990). *Tests and assessment* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Wei, R. C., Darling-Hammond, L., & Adamson, F. (2010). *Professional development in the United States: Trends and challenges*. Retrieved from http://www.nsdc.org/news/NSDCstudytechnicalreport2010.pdf

Willis, A. I. (2009). Edupolitical research: Reading between the lines. *Educational Researcher, 38*, 528-536. doi:10.3102/0013189X09347584

Wills, J. S., & Sandholtz, J. H. (2009). Constrained professionalism: Dilemmas of teaching in the face of test-based accountability. *Teachers College Record, 111*, 106-114.

Author Biography

Annela Teemant, is an associate professor of Second Language Education at the Indiana University School of Education at Indiana University-Purdue University, Indianapolis (IUPUI). Her interests focus on developing, implementing, and researching quality teacher preparation for mainstream teachers of English Language Learners. She has been awarded four U.S. Department of Education grants focused on teacher quality.