Support for the Improvement of Practices through Intensive Coaching (SIPIC): A model of coaching for improving reading instruction and reading achievement

CrossMark

Misty Sailors [a, *], Larry Price [b]

[a] The University of Texas at San Antonio, USA
[b] Texas State University, USA

HIGHLIGHTS

- SIPIC was effective in improving teacher's instructional reading practices.
- There was increased student reading achievement in participating classes.
- There was increased reading achievement of readers who struggle with learning to read.
- Coached teachers offered more opportunities with cognitive reading strategies.
- Coach/teacher interactions were associated with improved instruction.

ARTICLE INFO

ABSTRACT

In this study, we examined one model of coaching, Support for the Improvement of Practices through Intensive Coaching (SIPIC), which draws from both direct and responsive models of coaching with classroom teachers. We found the model to be effective in improving the comprehension instruction of teachers and in raising the reading achievement of students, including students who struggle with learning to read. Additionally, we found that the interactions between coaches and teachers were statistical associated with the instructional practices of the teachers, demonstrating empirically that coaches' behaviors do influence the professional practices of teachers.

© 2014 Elsevier Ltd. All rights reserved.

1. Background on literacy coaching

One approach to improve reading achievement (the ability of students to comprehend text) is to improve the ability of teachers to effectively teach their students. As a result, there have been vast amounts of federal, state, and local monies spent on professional development each year (Borko, 2004) even though there are no clear directives for how these activities should look (Lipson, Mosenthal, Mekkelsen, & Russ, 2004). One such professional development that is showing great promise for improving literacy instruction is literacy coaching (Sailors & Shanklin, 2010). Although literacy coaching first appeared in the literature almost 80 years ago

(Cassidy, Garrett, Maxfield, & Patchett, 2010), research on it is relatively recent.

Findings from a recent literature review of coaching studies in the USA (Sailors, Minton, & Villarreal, 2013) reveal that coaching studies center on one of three themes. First, coaching studies demonstrated that coaches have many roles (Bean, Swan, & Knaub, 2003; Lynch & Ferguson, 2010; Mraz, Algozzine, & Watson, 2008; Walpole & Blamey, 2008). Second, studies have shown that teachers appreciate their coach (Bean, Draper, Hall, Vandermolen, & Zigmond, 2010; Diamond & Powell, 2011; Downer, Kraft-Sayre, & Pianta, 2009; Marsh, McCombs, & Martorell, 2012) and find their engagement with a coach to be generally positive (Ferguson, 2011). Third, teachers report they appreciate when their coaches share ideas and help and encourage them (Armstrong, Cusumano, Todd, & Cohen, 2008). In short, coaching receives high marks from the field.

* Corresponding author. The University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA. Tel.: +1 210 458 7288.
E-mail address: misty.sailors@utsa.edu (M. Sailors).

Research has also explored the role of literacy coaching on improving the perceptions and attitudes of classroom teachers. In Malawi, for example, a team of international colleagues (Sailors et al., 2014) examined the effectiveness of an innovative complementary reading program that included coaching. Two groups participated in the study: Treatment teachers received complementary teaching and learning materials, workshops, and directive coaching, and control teachers received no intervention. After this five-month intervention, treatment teachers were significantly more comfortable with their languages of instruction (Chichewa and English) and were more positive about their teaching ability, beliefs about the learning materials in their classroom, and beliefs about the culture of reading in their communities than control teachers. The authors suggested that the implementation of coaching was an important source of support in changing teacher's beliefs and attitudes.

Other research examined the effectiveness of coaching as it influenced teacher practice. In the USA, for example, Neuman and Wright (2010) studied differences in the effects of two models of professional development for pre-kindergarten language and literacy instruction—traditional university coursework only and on-site coaching. Teachers who received coaching outperformed teachers who received only coursework and those in the control group on environmental classroom measures. Similarly, Walpole and her colleagues (Walpole, McKenna, Uribe-Zerain, & Lamitina, 2010) examined coaching practices linked to teacher practice. Using structural equation modeling to explore the relationships between coaching factors and instructional factors, they identified coaching factors that were significant predictors of at least one instructional factor, accounting for differences by grade level.

But not all studies have shown such positive outcomes of the impact of coaching on practice. For example, in Marsh's investigation, less than half of coached teachers (47% of reading and 40% of social studies) reported that the reading coach had influenced them to make changes to their instruction to a moderate or great extent. Similarly, 24% of reading and 34% of social studies teachers reported that their coach had no influence on changes in their instruction (Marsh, McCombs, & Martorell, 2010). Likewise, Carlisle and Berebitsky (2010) reported that teachers who received coaching did not differ from teachers who were not coached on aspects of instruction relevant to the professional development program. Findings of Whitaker and colleagues (Whitaker, Kinzie, Kraft-Sayre, Mashburn, & Pianta, 2007) indicated no statistically significant differences existed between teachers who had a coach and those who did not on measures that captured practices related to that of the program. And, although the five-month Malawi intervention mentioned earlier was successful in changing the perceptions and beliefs of participating teachers (Sailors et al., 2014), the treatment was not effective in improving the instructional practices of those same teachers.

In addition, recent research has examined the impact of coaching on student reading achievement. Biancarosa and her colleagues (Biancarosa, Bryk, & Dexter, 2010) measured the effects of one literacy program, Literacy Collaborative, on long-term student learning. Using a value-added model, the research team found positive effects for the model on improvements in literacy learning. Similarly, Matsumura and her colleagues (Matsumura, Garnier, Correnti, Junker, & Bickel, 2010) investigated the effects of Content-Focused Coaching on new teachers recruited in a district that suffered from a high-turnover rate among its teaching staff. Their findings indicated that the program predicted significantly higher school-level gains by English language learners on the state standardized test.

Likewise, Sailors & Price (2010) explored the role of coaching as a means of professional development in improving comprehension instruction in elementary and middle school classrooms. They tested two models of professional development on the effects of improving the reading comprehension instruction of teachers and increasing the reading achievement of students in low-income schools. Using a random-effects, multilevel, pretest–posttest comparison group design to explore the effectiveness of the two models, the full intervention group (workshop plus coached) outperformed the partial intervention group (workshop only) across teacher observation and student achievement measures.

Other studies have reported mixed findings related to coaching and student achievement. In Belgium, for example, Van Keer & Verhaeghe (2005) compared year-round intensive coaching (35 contact hours) to a more "restrictive" model (15 professional development contact hours) for second and fifth grade teachers. They found that both treatments were equally effective in changing students' reading comprehension, fluency, strategy use, and self-efficacy. The authors hypothesized that the lack of differences could have been due to a small sample size, and that the teachers in the restricted group had already worked with the researchers in another professional development context, thus potentially convoluting the findings, and/or there was little data on the fidelity of implementation of the intervention by the teachers.

In a more recent study in the US, Lovett et al. (2008) implemented a coaching model to prepare high school teachers to remediate reading instruction for students with reading disabilities. In their longitudinal study, the authors worked with 23 teachers, teaching them to develop metacognitive models of literacy instruction, to become more reflective about their teaching practices, and to master effective multiple component approaches to reading interventions. Observations, feedback, modeling, and support were "integral" components in this intervention. Student outcome data indicated that there were greater gains in classrooms where teachers had an additional year of coaching support.

In summary, literacy coaching is a growing field and, although teachers appreciate and value their coaches, there are no definitive conclusions as to the effectiveness of coaching on teachers or student reading achievement (Sailors et al., 2013). To that end, we had two goals in this study: (a) to examine one model of coaching, Support for the Improvement of Practices through Intensive Coaching (SIPIC), testing its effectiveness on the instructional reading practices of elementary and middle school teachers and the reading achievement of their students and (b) to contribute to the general literature on coaching as a means of professional development for classroom teachers. Our research questions included, (a) What are the associated effects of the SIPIC model on the instructional reading practices of participating teachers? (b) What are the associated effects of the SIPIC model on the reading achievement within participating classrooms? (c) How often and in what ways do coaches support teachers when using the SIPIC model? (d) What aspects of the SIPIC model can be attributed to the improved instruction of participating teachers?

2. Directive and responsive coaching

The SIPIC model was designed to be classroom-based, embedded in the school day, and sustained over time with qualified coaches interacting with teachers and integrating promising practices into existing practices within a teacher's own classroom (Sailors & Shanklin, 2010). The model followed the belief that if teachers are to learn new practices and incorporate them effectively into their classroom, they must understand the theory, see the practices modeled, and have opportunities to discuss the practices with a knowledgeable other.

The SIPIC model is grounded in the "situative perspective" of knowledge, thinking, and learning (Greeno, 1997, 1998), which

states that human learning is guided by interactive systems of activity (Greeno, Smith, & Moore, 1993) in which individuals participate to achieve objectives that are meaningful in relation to their membership in communities of practice (Greeno, 1998). Learning, under this perspective, is guided by a set of activities that are somewhat multifaceted at the beginning so that the activities are more personally and socially meaningful to the learner and are dubbed "authentic" activities in some circles (Brown, Collins, & Duguid, 1989; Collins, Brown, & Newman, 1987). Learning occurs when people are engaged in quantitatively more and qualitatively complex activities. In order to become fully participatory in their community of practice, learners must have access to learning experiences that are social, situated, inclusive of tools of their practice, and built around discussions with knowledgeable others (Putnam & Borko, 2000).

Furthermore, SIPIC was based on theories of adult and higher education (Anderson, Reder, & Simon, 2000), specifically, teaching teachers what they needed to know at the time they needed and wanted to learn it, thereby assisting the teachers in thinking deeper about comprehension instruction. In fact, the model was grounded in teachers (a) volunteering to participate in the professional development, (b) selecting which reading strategy they would like their coach to focus on, and (c) selecting the type of exchange they would like to have with their coach. These are explained in a subsequent section.

We grounded the situated experiences of participating teachers within the frame of how literacy coaches position themselves with teachers (Ippolito, 2010). Although not as easily dichotomized as the labels may appear, responsive (Dozier, 2006) and directive (Deussen, Coskie, Robinson, & Autio, 2007) coaches play very different roles with teachers (Ippolito, 2010). Also known as development-reflective coaching (Kuijpers, Houtveen, & Wubbels, 2010), reflective coaching engages teachers in joint inquiry about teaching as a way of shifting teaching practices (Butler, Lauscher, Jarvis-Selinger, & Beckingham, 2004) through reflection (Costa & Garmston, 2005). Reflective coaches often borrow from the three-phase cycle described in Costa and Garmston's work, including pre-observation conferencing, observation, and post-observation conferencing.

Directive coaches, on the other hand, assume the role of "expert" and provide advice to teachers on how to implement a specific program using specific practices (Deussen et al., 2007), in some cases to ensure program fidelity (Bean et al., 2010). Labeled in many ways, including 'expert consultant' (Gersten, Morvant, & Brengelman, 1995), 'technical coach' (Denton & Hasbrouck, 2009), 'expert mentor' (Carlisle & Berebitsky, 2010), 'sounding board' and 'advisor' (Sperry, 2008), directive coaches often assume the role of knowledgeable specialists and "are assertive about what instructional practices teachers must implement" (Ippolito, 2010, p. 165). Often referred to as an "outside in" approach (Sheridan, Edwards, Marvin, & Knoche, 2009), studies on consultant coaches have demonstrated improvement in practices (Pianta, Mashburn, Downer, Hamre, & Justice, 2008). When effective, the process of directive coaching calls for frequent exchanges over a relatively short period in order to change practices, attitudes, and/or dispositions (Sheridan et al., 2009) and the nature of those exchanges must be "highly individualized" (Sheridan & Kratochwill, 2008).

While some teachers appreciate direct and explicit messages from the coach and report that they find it motivating to interact with coaching consultants (Shulman, 2004, p. 396), directive coaching is problematic with other teachers. Directive coaching sometimes creates conflict with teachers who resent being told what to do (Deussen et al., 2007) and report that they would rather have conversations with their peers than site administrators, district consultants, or university faculty (Lapp, Fisher, Flood, & Frey, 2003).

In short, the problem with directive coaching is "that its directness is as threatening as it is powerful" (McDonald, 1989, p. 211).

While research has demonstrated that coaches need to be able to negotiate and balance both reflective and directive coaching with teachers (Ippolito, 2010), the literature remains underdeveloped as to what a hybrid model would look like and how it would influence the instructional reading practices of teachers and the reading achievement of students. It was in that spirit that we designed the SIPIC model and tested it in this study. The model is a hybrid of directive and responsive coaching, as it was a combination of teaching the practice of intentional instruction (directive) through a reflective coaching cycle based on the need-to-know of participating teachers (reflective).

3. Methods

We employed a pretest–posttest control group design (Shadish, Cook, & Campbell, 2002) to test the effectiveness of the model under investigation, with a focus on improving the instructional comprehension practices of classroom teachers. That is, teachers who participated in this study learned how to be intentional in explaining the underlying sub-routines involved in cognitive reading strategies to their students. There were two groups of teachers in this study: one group received the full intervention (workshops plus the SIPIC model, the treatment group) and one group that received a partial intervention (workshop only, the control group). Our hypothesis was that even though we provided all teachers with initial information on intentional instruction, teachers in the treatment group would outperform those in the control group.

3.1. Participants

Participants included 120 teachers (50 control and 70 treatment) from 4 school districts located across two large metropolitan cities: one district was in central and three were in south Texas. Our participants taught across the grade levels: 16% taught second, 12% third, 13% fourth, 15% fifth, 17% sixth, 15% seventh, and 12% were eighth grade teachers. We had no first grade teachers in our database. The years of experience of participating teachers ranged from one to 36 years of experience (mean of 10.45; SD of 7.96). Only 5% of the teachers held a master's degree and of the remaining teachers, 39% reported having completed at least some graduate hours. Eighty-eight percent of the teachers earned their accreditation traditionally and 12% completed an alternative certification program. There were no statistical differences between the groups regarding years of experience, level of educational attainment, or type of certification.

This study focused on the professional development of regular education teachers; therefore, we did not work with special education, pre-kindergarten, or bilingual teachers. Because schools are social places and teachers closely interact with each other during the school day, a phenomenon known as experimental treatment diffusion presented a threat to the internal validity of the investigation (Shadish et al., 2002). In order to avoid this phenomenon, we randomly assigned schools to a treatment group. Table 1 provides the demographic information on the students, schools, and number of participating teachers from each district.

There were 1496 students participating in this study because of their teacher's participation and parental consent. The vast majority of these students (86%) were from low-income families as seen in Table 2. Because our intervention focused on the improved reading achievement of regular education students, our database does not contain assessment data of students labeled as special education.

Table 1
Demographic characteristics of participating districts, schools, teachers, and students.

District	Student count within district	School within district	Teacher count within school	
ID	n	ID	n	%
1	659	101	16	0.30
		102	6	0.11
		103	13	0.24
		104	5	0.09
		105	7	0.13
		106	2	0.04
		107	5	0.09
		n = 7	n = 54	
2	437	201	5	0.14
		202	9	0.24
		203	17	0.46
		204	4	0.10
		205	1	0.03
		206	1	0.03
		n = 6	n = 37	
3	131	301	9	0.75
		302	2	0.17
		303	1	0.08
		n = 3	n = 12	
4	269	401	4	0.24
		402	2	0.11
		403	1	0.06
		404	1	0.06
		405	1	0.06
		406	2	0.11
		407	1	0.06
		408	1	0.06
		409	1	0.06
		410	2	0.12
		411	1	0.06
		n = 11	n = 17	
Total N:	1496	27	120	

Table 2
Demographic characteristics of participating students.

Characteristic	n	%
Gender		
Male	668	45.00
Female	809	54.00
Missing	–	0.01
Ethnicity		
Hispanic	1230	82.00
White	124	8.00
African American	105	7.00
Asian/Pacific Islander	2	0.10
Other	2	0.10
Missing	–	2.20
Grade		
1	16	1.10
2	270	18.00
3	229	15.00
4	165	11.00
5	181	12.00
6	222	15.00
7	244	16.00
8	154	10.00
K	15	1.00
ELL		
Bilingual	126	8.00
English Language Learners	71	5.00
English only	1185	79.00
ESL	87	6.00
Missing	–	2.00

3.2. Intentional instruction

Reading is a complex, strategic process that requires both cognitive and metacognitive strategies (Block & Pressley, 2002). In fact, proficient readers use a set of cognitive reading strategies or "deliberate, goal-directed attempts to control and modify the reader's efforts to decode texts, understand words, and construct meaning of text" (Afflerbach, Pearson, & Paris, 2008, p. 368). These actions require that readers have metacognitive awareness, including declarative, procedural, and conditional knowledge (Paris, Lipson, & Wixson, 1983). This metacognitive knowledge is especially helpful when reading becomes difficult (Paris et al., 1983) and subsequently, when a reader must determine the appropriate reading strategies to use (Pressley, Borkowski, & Schneider, 1989).

We have dubbed the type of instruction that focuses on teaching metacognitive reading strategies "intentional instruction" (Sailors & Price, 2010) whereby teachers (a) provide opportunities for children to engage in reading strategies; (b) identify the reading strategies required during reading; and (c) explicate and discuss with students their own cognitive reading processes during reading. This last bullet includes telling children what strategy is being used, how they know to use it, why they are using it, and the steps they employ as they engage in the reading strategy.

3.3. Treatment groups and content of the professional development

Our participating school districts were adamant that the control teachers "got something, too;" subsequently, we designed a two-day summer workshop for all teachers to attend. Because there were too many cognitive reading strategies to learn in a 2-day period, we chose to engage the teachers with one cognitive reading strategy (drawing conclusions). The professional development team taught teachers how to: (a) identify places in text where readers are required to draw a conclusion in order to understand the text; (b) identify how readers know to draw a conclusion; (c) identify when it is appropriate to use "drawing conclusions"; (d) identify why readers draw conclusions; (e) explain the sub-routines involved in drawing conclusions; and (f) create spaces in which conversations about drawing conclusions (and the underlying sub-routines) could take place during reading.

We offered these two-day workshops four different times as a way of reducing the number of teachers that attended each workshop; the workshops were held before teachers were assigned to treatment groups and before data collection began (see section below). Although these workshops were held at the district headquarters and were devoid of students, they were based on structural and core features of effective professional development workshops for teachers (Birman, Desimone, Porter, & Garet, 2000; Desimone, 2002; Wayne, Yoon, Zhu, Cronen, & Garet, 2008). Specifically, the workshop involved active learning, promoted coherence, and aligned with district and state standards. Teachers practiced the content with each other, receiving feedback from their colleagues. There was no additional support outside these workshops for the teachers in the control group.

3.4. Coaching using the SIPIC model

The coaching model (SIPIC) extended the treatment for intervention teachers. Under the support they received through SIPIC, the teachers continued to learn to be intentional in their comprehension instruction through classroom-based support. The "external coaches" visited each classroom at least twice per month, and engaged the teachers in a variety of exchanges based on a movement toward full participation on the part of the teachers (Lave & Wenger, 1991). Before each visit, the coach would engage the teacher in an

email (or phone call), asking the teacher to identify the reading strategy that would be the focus of the lesson. The coach encouraged the teacher to choose a strategy based on the needs of her students or her own needs (one she was struggling to teach). The teacher also selected from one of the four possible types of exchanges: guided observations, co-teaching, guided reflections, or guided conversations. This served as a pre-conference and centered on discussions of the upcoming lesson, its design, and how it would align with other lessons in the classroom.

The first exchange, guided observations were demonstration lessons taught by the coach in the teacher's classroom. As the most pervasive type of exchange between coaches and teachers in schools today (Alvermann, Commeyras, Cramer, & Harnish, 2006; Deussen et al., 2007), we felt it necessary to structure the exchange carefully. While the coach was engaged in a demonstration lesson, she asked the teacher to engage in a set of guiding questions (see Fig. 1). The questions provided discussion points for the post-conference.

Co-teaching, the second type of exchange, involved the planning and teaching of a lesson together. Following the lesson, the coach and teacher engaged in a post-conference, discussing the strengths of the lesson as it pertained to intentional instruction. The post-conference ended with a conversation related to next-steps for future strategy lessons.

The third type of exchange, the "guided reflection," involved the coach documenting a lesson as the teacher taught it. During the pre-conference the coach would ask the teacher what observational data she would like collected, with a focus on becoming more intentional in her instruction. In some cases, teachers wanted coaches to watch for places where they allowed students to take the lead in the conversations around strategies, in other cases, teachers simply wanted coaches to document the way they described the focus strategy. Regardless, the coach documented the lesson, writing "like crazy" (Wolcott, 1995), creating an artifact of the lesson that they used for stimulated recall during the post-conference (Kucan, 2007).

The fourth exchange involved what we termed "guided conversations," those exchanges that involved the teacher and coach sitting down to discuss the teacher's struggles and challenges with intentional instruction. Drawing on the work of Costa and Garmston (2005), teachers and coaches engaged in thoughtful conversations intended to guide the thinking processes of the teacher related to explanations and stimulating discussion of reading strategies. These exchanges included brainstorming, giving and seeking of advice, and sharing of ideas and resources (Glazer & Hannafin, 2006, p. 181). These exchanges helped the teacher examine and deepen her inner thought processes related to reading strategies, what they were, when they are used and how to offer opportunities for students to explore them. These exchanges took place outside of the context of classroom instruction (e.g., during conference period, lunch, or after school).

3.5. Training of the SIPIC coaches and fidelity of implementation

The four coaches for this study were highly qualified based on professional standards (International Reading Association, 2004) and included two faculty members (including the first author), one retired reading specialist, and one district language specialist. The coaches attended a three day workshop that explored (a) theories of adult and higher education (Anderson et al., 2000); (b) types of knowledge held by teachers, including content (cognitive reading instruction), pedagogical (intentional instruction), and pedagogical content (how to incorporate this into their existing curriculum) (Shulman, 1986); (c) how to engage teachers in change using Guskey's (2002) model of teacher change; (d) concerns of teachers and levels of use of innovations as teachers engage in new innovations (Ward & McCotter, 2004); (e) how to focus on existing knowledge and practices of the teachers through observations and conversations (Cochran-Smith & Lytle, 1999; Lieberman, 1995); and (f) how to interact with teachers as learners while staying within their Zone of Proximal Development (Vygotsky, 1978). The coaches used the work of Duffy (2009) as a resource.

We collected data to ensure that the coaches were implementing the most critical components of the intervention in similar ways (Mowbray, Holter, Teague, & Bybee, 2003). Members of the data collection team used the teacher observation protocol (described below) on two lessons delivered by each of the coaches—one at the

---

- Classroom teacher's name: _____

- Visit number: _____

- Focus of the lesson and cognitive strategy: _____

Please think about these items during the demonstration lesson:

- Things I'm thinking about: (What my children are doing; Ways in which the coach

  is scaffolding learning; etc.)

- Ways in which the instructional strategy is similar to what I do in my classroom:

  (How are the texts similar? Instructional strategies? Language used by the coach?

  Response of my learners to the lesson? Etc…)

- Ways in which the lesson is different from when I teach strategy instruction: (How

  did the learners respond to this lesson? Where there "missed opportunities"? Etc…)

- Lingering questions I want to ask my coach.

Fig. 1. Guiding questions.

beginning and one midway through the study. University reading faculty used the data to determine if the coaches were using intentional instruction during the observed demonstration lessons. University reading faculty also monitored the coaching reports every month to ensure that the coaches were: (a) visiting the teachers; (b) allowing teachers to select the focus cognitive reading strategy of the exchange; (c) engaging the teachers in multiple ways (guided observations, reflection, conversations, and co-teaching).

3.6. Data collection

The research team collected a variety of data in order to examine the research questions. Three graduate students enrolled in the masters reading program and one adjunct reading faculty served as data collectors for this study. All held or were seeking advanced credentialing from the State as reading specialists and all were former classroom teachers. A project-employed research coordinator scheduled all data collection. Data collectors conducted nearly equal teacher observations, with two of them conducting a few observations more than others. Teachers were asked to schedule observations during a lesson in which reading in connected text was a significant part of the lesson. Data collectors also collected student data, including fall and spring reading achievement data and spring classification data (see below).

3.6.1. Student reading achievement data

We used the *Group Reading Assessment and Diagnostic Evaluation (GRADE)* to assess the reading achievement of students in this study. The GRADE is a group administered, norm-referenced, standardized reading assessment (American Guidance Services, 2001). We selected the GRADE because of the quality of its psychometric properties, its availability for each grade level requisite to this study, and because it measures reading comprehension. We used the GRADE's Standardized Growth Scale Values (GSVs) to measure change (i.e., increase or decrease at posttest) regarding students' achievement of reading comprehension. We used GSVs because these composite scores (i.e., a total test score) provide a measure of a student's reading achievement in reference to the entire range of achievement across all grades. Evidence of adequate score reliability (coefficient $\alpha = .89$) and validity evidence met recommended criteria (AERA/APA/NCME, 1999).

3.6.2. Data to classify students

We used a norm-referenced assessment, the Wide Range Achievement Test 3 (or WRAT-3) (Wilkinson, 1993), to identify participating students in our treatment and control classrooms as being above, on, and below grade level readers. The WRAT3 yields normative score values, can be used for ages 5 through 75 years, is individually administered, and provides absolute scale scores across the full range of development (Wilkinson, 1993). Evidence of adequate score reliability (coefficient $\alpha = .89$) and validity evidence met recommended criteria (AERA/APA/NCME, 1999). We did not use the WRAT3 to measure reading achievement. Rather, we used this classification scheme in our analysis of the reading achievement outputs.

3.6.3. Teacher observation data

We used the *Comprehension Instruction Observation Protocol System* (CIOPS) (Sailors, 2008), a category observation instrument that captured the teaching of reading strategies. The CIOPS is a combination of observational note taking and a quantitative coding process. A list of the cognitive strategies captured by the CIOPS is listed in Fig. 2. A list of scaffolding techniques is listed in Fig. 3. A full

description and the validity of the instrument is described elsewhere (Sailors & Price, 2010a).

3.6.4. Data coding and interrater reliability

During the two data collection time periods (early fall and late spring), observers took continuous and descriptive field notes (following the work of Wolcott, 1995) during the two, 45-min observations in each classroom. These field notes focused on the context of the classroom, materials used, text-based comprehension and/or comprehension strategy instruction, and the instructional strategies employed by the teachers. After the classroom visit, but within a 24-h period, the observer transferred the narrative notes into the CIOPS system and coded them.

In order to reduce the inference required of the observers during the coding process (Herbert & Attridge, 1975), the narrative data were reduced to the smallest possible reasonable and recognizable units as they were coded (Martin, 1977). In designing the instrument, we selected units to be coded based on the cognitive reading strategies that we identified in our review of the literature. The data collectors identified and marked those instances in the lesson in which the teacher provided an opportunity to engage the students in discussion around a cognitive reading strategy. We defined opportunity as "a set of circumstances that makes it possible to do something," such as the engagement in a reading strategy. If there was no evidence of such opportunities, the collector marked "no strategy." In the case of an opportunity to engage, the data collector then marked the instructional interactions between the teacher and the students in the class. As the cognitive reading strategy under discussion changed, so did the coding. For example, the text that appears in Fig. 4 is a portion of data from an observation that has been coded and transferred into the CIOPS system (see Fig. 5).

Intensive training and monitoring in the use of the teacher observation data were necessary to obtain reliable results. Observers were trained for two full days in a university-based classroom environment in which they were introduced to the theoretical frame of the study, how to "write like crazy" (Wolcott, 1995), and how to code the data. Training took place using an electronic training module, including a manual, a practice video, and various examples of video clips of teachers engaged in teaching cognitive reading strategies. On the third day, observers were placed in non-study schools with the trainer. To evaluate interrater reliability, data were compared between each pair of points of agreement between the observer and trainer (i.e., labeling of the cognitive reading strategy and type of exchange between teacher and students). In the morning, half of the observers were placed in a classroom with the trainer, marking data independent of each other. Point-by-point agreement using Cohen's Kappa statistic revealed an interrater reliability of .80. Later that same morning, the second half of the observers were placed in a different classroom with the trainer and the pair collected and marked data independent of each other. Point-by-point agreement using Cohen's Kappa statistic revealed an interrater reliability of .81. Results of the interrater reliability analyses using the Kappa coefficient were acceptable in both instances with an estimate of $r \geq .80$ (Fleiss, 1981). Throughout data collection, 10% of the lessons were checked for ongoing interrater reliability with results remaining above a Kappa of .80. Data collection and coaching visits were never conducted during the same week.

3.6.5. Coaching logs

We used coach logs to help coaches document their exchanges with teachers and to answer our research questions. After each exchange with a teacher, coaches completed a report that included information on the date of the exchange, the amount of time spent

Activating and Accessing Prior Knowledge*	Asking, Answering, and Reformulating Clarification Questions*
Classifying and Categorizing*	Connecting*
Critical Literacy: Authenticity; Generalizability; Perspective; and Purpose of Author*	Determining Main Idea / Theme*
Finding, Extracting, and Paraphrasing Important Details*	Drawing Conclusions*
Integrating External Text Features*	Monitoring*
Predicting / Confirming / Re-predicting*	Previewing*
Recalling*	Responding Aesthetically*
Setting a Purpose*	Skimming and Scanning*
Summarizing*	Synthesizing*
Test Taking Strategies	Using Fix-Up strategies
Visualizing*	Word Identification and Word Recognition

Fig. 2. Reading strategies captured by the CIOPS (*denotes foci strategy).

Assessed the strategy
Engaged in defining the strategy
Engaged in demonstrating the strategy
Engaged in explaining the process
Engaged in modeling when to use the strategy
Engaged in modeling why to use the strategy
Engaged in naming the strategy
Engaged in providing other examples
Engaged in practice with the strategy
Engaged in questioning techniques
Engaged in reminding of use of strategy
Engaged in telling when to use the strategy
Engaged in telling why to use the strategy

Fig. 3. Scaffolding techniques captured by the CIOPS.

with the teacher, the nature of the exchange, the cognitive strategy that was the focus of the exchange, and who initiated the exchange (coach or teacher). The coaches met as a team biweekly to discuss their interactions with the teachers. In attendance was with a member of the professional development team (faculty members in literacy education); these meetings allowed us to feel confident that the integrity of the intervention was being maintained throughout the study (O'Donnell, 2008).

3.7. Data analyses

Prior to answering our research questions, we screened the data for missing and/or extreme values, explored distributional patterns of frequency counts of observational data, and examined the tenability of the general linear model assumptions of linearity, normality, homoscedasticity, and independence of observations within each classroom and school and for the total sample. To increase the sensitivity of the analyses, we examined these data characteristics and model assumptions at every level of our analysis. Based on the results of the data screening, we proceeded by using a hierarchical modeling analytic strategy, allowing us to examine the fixed and random variation in student change as measured by the GRADE within between treatment and control classrooms, with students nested within classrooms and classrooms nested within schools. The respective group and the classroom in which individual students were naturally nested or clustered constituted the fixed effect portion of our model within the multilevel design.

We allowed the student-level data on the GRADE to randomly vary across time using an autoregressive level-1 covariance structure allowing us to model the dependence (i.e., the correlated

Sample from Teacher 01123 (post-observation)
Overview: Teacher and students are moving toward the carpet with an issue of Weekly Reader. The teacher hands out article and then tells stds to sit on reading rug in back of room. Teacher sets the easel up at the front of the carpet; it has a chart of it that was partially filled in last week.

T- I still have the volcanoes that we studied last time on the chart. Who can raise their hand and tell me something we learned last week? S- It burns. T- Right. It'll burn your rubber boots. S- It rumbles. T- You can hear the rumble from the plates shifting. T- What else do you remember from last week? <Stds read various facts that appear on the instructional aid. >

T- Who can raise their hand and remind me what an inference is? S- What you know and what it tells you. T- Good. He said that wonderfully; what you already know and what the text tells you. Did it tell you word for word that it will melt your boots? No we made an inference. You also need to be thinking about what you know and what you want to know. That's the first thing we are going to do. (Erases chart)

T- Does anyone know how to say this? Vesuvius. Is this a real picture? S- No T- No, it's a painting. What does depicting mean? (Teacher reads definition from text) T- Look at the picture on the top. Now let's look at the map. What does Italy look like? S- Boot T- Yeah, that's right. T- Capital city? What is really close to Mt. Vesuvius? S- The capital. T- Right, Naples.

<Lesson continues>

Fig. 4. Sample text from an observation.

structure) among student-level scores nested within individual classrooms across time. We allowed intercepts and slopes to randomly vary for student achievement data as a way of modeling individual change. In our analysis, we used the Hierarchical Linear Modeling computer program, version 6.0, (Raudenbush, Bryk, Cheong, Congdon, & DuToit, 2004). For analysis of teacher observational data, we proceeded by conducting between groups chi-square analyses of change scores based on frequency counts of

observational data within classrooms (Von Eye & Niedermeier, 1999).

We examined the effect of the intervention on teachers' use of intentional instruction expressed as the mean number of times we observed teachers implementing comprehension strategies. There were no differences in teachers' use of strategies at pretest. Therefore, we conducted a between study groups chi-square test to examine the effect between teacher groups. Our goal was to see if
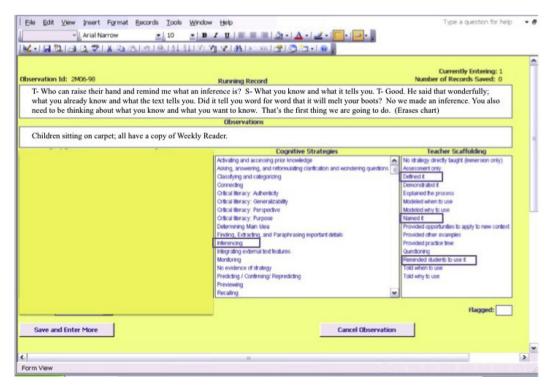

Fig. 5. Screen shot with sample text coded inside the CIOPS.

the net change (expressed as a change score) was statistically significant between the groups. We used a chi-square analysis because the change or difference scores were not normally distributed.

4. Findings

The purpose of this study was to demonstrate the effectiveness of the SIPIC model on the instructional comprehension practices of classroom teachers as exhibited by the reading achievement of their students. We compared the opportunities teachers provided for students to engage in cognitive reading strategies and the instructional interactions between teachers and students, specifically exploring the way in which these teachers explained the strategies to their students, following intentional instructional. We also examined the extent to which the coaching model may have contributed to changes in teacher practices and the outcomes of students.

4.1. Impact on teaching practices

In order to answer the question, "What are the associated effects of the SIPIC model on the instructional reading practices of participating teachers?" we began by using "micro" individual-level (teacher) observational data points by recoding these individual units into larger composite-level variables thereby allowing us to examine the teaching practices of participating teachers. We used two composite variables, "opportunities to engage in comprehension strategies" and "constructed explanations" both drew from raw frequency counts of the CIOPS data (observed events as nominal yes/no for each of the strategies and interactional explanations). The first, "opportunities to engage in comprehension strategies" was designed to capture the occurrence of cognitive reading strategies and has been shown in previous research to be positively associated with increases in measures of comprehension (Sailors & Price, 2010). We used Almasi's (2013, p. 12) organizational structure and summed all raw variables together except for a few: the word identification/word recognition, fix-up, and test taking strategies. Reliability analysis (i.e., internal consistency) provided adequacy for the internal consistency of the opportunity to engage variable ($\alpha = .78$).

The second composite variable, "constructed explanations," allowed us to examine the constructed explanations around cognitive reading strategies. This composite variable consisted of all of the raw variables, except questioning, assessing, and reminding. We based the linear composite variable on the frequencies of teachers either exhibiting these practices or not. Reliability analysis using non-parametric correlational statistics provided marginal adequacy ($\alpha = .79$) for the internal consistency structure of this composite variable. A more detailed description of the composite variables can be found elsewhere (Sailors & Price, 2010).

There were no statistically significant differences between groups at pretest for either of these composite variables; therefore we conducted a between study groups posttest analysis of the means. When comparing the differences in the opportunities to engage in cognitive strategies ("opportunity" variable), we found a statistically significant difference in the direction of the treatment group regarding the frequency of use of the outcome variable "opportunity" ($X^2 (22) = 36.17, p < .05$). The magnitude of the effect size was medium (Cramer's $V = .50$, 95% CI = .43–.56). Interested readers are referred to Cohen (1988) for further information regarding the range and recommended interpretation of effect sizes (small = .10–.34; med = .35–.6; large = >.60). When comparing the differences in frequency of use of intentional instructional practices ("constructed explanations" variable) we found a statistically significant difference ($X^2 (17) = 27.17, p < .05$)

between the groups. The observational trend for this difference favored the treatment group and the practical effect for this analysis was observed as being large (Cramer's $V = .62$, 95% CI = .41–.77).

4.2. Impact on reading achievement

In order to answer the question, "What were the associated effects of the SIPIC model on reading achievement within participating classrooms?" we focused on the achievement level of students as measured by the GRADE. A second more specific question of interest was, whether or not there were any differential gains made by students of differing levels of achievement; we used the WRAT3 for this question. To examine these research questions, we used pretest and posttest student-level scores on the GRADE hierarchically nested within classrooms, and classrooms nested within schools, for the treatment and control groups to measure the change in student achievement across the school year.

The results of the hierarchical random coefficient model (i.e., allowing the intercepts and slopes to vary) analysis revealed statistically significant fixed and random effects. The fixed effects portion of the model yielded a 13.8-point increase from pretest to posttest ($\overline{X} = 428.2 - \overline{X} = 442.06$) for the full intervention group and an 8.69-point increase from pretest to posttest ($\overline{X} = 428.36 - \overline{X} = 437.0$) for the control group. Table 3 below provides the results of the hierarchical analysis. The reliability of the teacher (classroom) parameter estimates across time was very high ($r = .94$), providing support for a high level of discrimination between groups. Further, the full intervention group exhibited a large practical growth effect ($d = .45$) whereas the control group exhibited a small growth effect ($d = .24$). The overall summary of the total outcome variability explained by our multilevel random coefficients growth model exhibited excellent explanatory power with a pseudo R^2 of .89 (Singer & Willett, 2003). This indicates a strong relationship between the impacts of the treatment on student reading achievement.

We also examined the impact on student learner level (on, below, and above grade level readers as measured by the WRAT3). We again used pretest and posttest student-level scores of the GRADE hierarchically nested within classrooms, nested within schools, for the treatment and control by learner grade level across the school year. The fixed effects portion of the overall model main effect was observed as being statistically significant ($t = 369.37$, df = 24.69, $p < .001$). A particularly important practical finding from

Table 3
Multilevel effects of professional development model on student reading achievement.

Fixed effect	Coefficient	S.E.	df	t
Intercept (average)	414.01***	6.98	126.67	59.23
Group (treatment)	7.75	10.74	125.77	0.72
Time (average change in time slope)	14.19***	3.38	111.46	4.19
Group × Time (interaction of study group by time component)	7.59	5.17	110.55	1.46
Random effect	Parameter	S.E.		
Level-one (i.e., student-level) random effects:				
Intercept residual variance	499.11***	137.03		
Student level (slope variance)	281.28***	0.00		

***Indicates a significant effect at $p < .001$. On average, students mean score on GRADE was 414.01. Students in the control group scored an average of 7.75 points lower than those in the treatment group at posttest. Residual variance of the intercept in the random portion of the model was significant at $p < .001$; Random residual slope/change variance for individual students over time was significant at $p < .001$.

this analysis is that "below grade level" students in the treatment group exhibited the largest effect size ($d = .73$) from pre to posttest, an improvement of 28.73 score points. Also, the "above grade level" students in the treatment group exhibited an effect size of ($d = .70$) from pre to posttest, an improvement of 21.68 score points. The overall summary of the total outcome variability explained by our factorial multilevel random coefficients growth model exhibited excellent explanatory power with a pseudo R^2 of .81 (Singer & Willett, 2003).

We were also interested in the effects of the intervention overall and by learner grade level (interaction effect) on the achievement scores of students on the GRADE. We observed a statistically significant 3-way interaction effect, that is, a group X focus child level X time ($F = 3.54$, df = 2, 24; $p < .05$) for student learner level as assessed by the WRAT3. Specifically, we detected a statistical interaction effect in GRADE change scores over time between the children categorized as being below average. Strikingly, the mean score for control group children at the "below" category decreased from 424.5 to 413.5 (11 points) over time, while the mean scores for children in the "below" category group increased from 382.4 to 406.2 (23.8 points). The general linear model equation yielded a group X focus child level X time effect size of .23 (medium to large range). The magnitude of the effect size for the general linear model regression/ANOVA (η^2) assumes the following ranges (small = .02, medium = .15, and large = .35).

4.3. Exchanges between coaches and teachers

In order to answer our first question, "What types of exchanges took place between coaches and teachers when using the SIPIC model?" we analyzed the coaching logs. Our analysis indicated that the coaches visited each teacher an average of 7.41 times during the year, with a range from 1 to 16 times (SD 4.01). Likewise, the mean average time of the total coaching visits for each teacher was 355.49 min, with a range from 45 to 830 min (SD 201.75). Thirty-eight percent of the exchanges involved guided conversations and the remainder (62%) was classroom-based (involving the coach, teacher, and at least some of the students in the classroom). These exchanges included guided observations (50%), co-teaching (25%), and guided reflections with the coach (25%). Using Almasi's (2003) scheme for thinking about reading strategies, our analysis indicated a vast majority of the exchanges focused on cognitive reading strategies (98%), while a very small number (2%) focused on fix-up strategies. No exchanges focused on word identification strategies.

4.4. Aspects of SIPIC and improved instruction

In order to answer the question, "What aspects of the SIPIC model can be attributed to the improved instruction of participating teachers?" we explored the quantity of the interventions (measured by the number of minutes coaches spent engaged with the participating teachers) and the two composite variables. We observed no statistically significant differences in the direction of the treatment group regarding the quantity of the intervention and the opportunities ("opportunity" variable) teachers provided their students to engage in cognitive strategies. However, we did observe a statistically significant difference in the use of intentional instruction ("constructed explanations" variable) and the quantity of the intervention between groups (X^2 (70) = 104.95; $p < .01$, effect size = .62. 95% CI = .46–.79).

The next sets of analyses explored the various types of possible exchanges within the intervention and the improvement of instructional practices of the teachers. Our findings indicated that there were statistically significant associations between the types of exchanges and the instructional practices of the teachers. Our

results were as follows: Demonstration lessons X^2 (112) = 176.56 $p < .001$) between the groups with an effect size of Cramer's $V = .67$, 95% CI = .56–.79. For "co-teaching" X^2 (42) = 100.55 $p < .001$) between the groups with an effect size of Cramer's $V = .67$, 95% CI = .56–.79. For "feedback" X^2 (112) = 161.01 $p < .001$) between the groups with an effect size of Cramer's $V = .64$, 95% CI = .50–.78. For "teacher initiated exchange" X^2 (140) = 211.28, $p < .001$) between the groups with an effect size of Cramer's $V = .74$, 95% CI = .58–.90. Finally, for "coach initiated" change X^2 (84) = 109.89, $p < .01$) between the groups with an effect size of Cramer's $V = .66$, 95% CI = .50–.82.

In summary, we found the SIPIC model of coaching to be effective in improving the practices of reading teachers and increasing the achievement scores of students, especially those who struggle with learning to read. In the next section, we discuss these findings, relating them to existing research on coaching, changes in teaching practices, and student achievement.

5. Discussion

In this study, we tested the effectiveness of the SIPIC model, namely to see if a hybrid model of coaching could improve the instructional reading practices of elementary and middle school teachers and improve the reading achievement of their students. As a hybrid model, SIPIC combined aspects of both directive coaching (with a focus on comprehension instruction) and responsive coaching (working with teachers who *self-selected* to learn, with a focus on *what* teachers wanted to learn, in ways that met *how* the teacher wanted to learn). Our findings are noteworthy for several reasons. First, our findings indicate that coaching continues to serve as a viable means of improving the instructional practices of reading teachers. Namely, the treatment teachers not only provided more opportunities to engage in cognitive reading instruction than control teachers, they also offered better explanations/co-constructions of those reading strategies, with large medium to large effect sizes, respectively. These findings align with years of research that shows teachers can learn to be more explicit in their comprehension instruction (see the works of Gerald Duffy and Michael Pressley) and more recent research that shows coaching can provide the type of support teachers need to improve their reading strategy instruction (Hsieh, Hemmeter, McCollum, & Ostrosky, 2009).

Second, our findings indicate that coaching may have had an impact on the reading achievement of students in participating classrooms. Namely, students in classrooms that received coaching had greater gains on the standardized reading measures than students in control classrooms and there were greater practical effects when comparing the two. These findings align with research that demonstrated that teacher explicitness in cognitive reading instruction leads to an increase in reading achievement on the part of students. For example, in his seminal research, Duffy and his colleagues (Duffy et al., 1986, 1987) found that students of teachers who had been taught to explain reading strategies had significantly greater awareness of strategies, why those strategies were important, and how to use the strategies than students of the comparison teachers. Others have demonstrated similar findings—children can be taught to be strategic in their thinking (Pressley, 2000). Previous research suggests that it takes time for students to become strategic readers (Duffy, 2002); our work suggests that an intensive model of professional development for classroom teachers (such as SIPIC) provides the context for this learning to take place.

Third, and perhaps more interesting, were our findings related to students who were labeled as "below grade level" (based on their WRAT scores) in coached classrooms as the largest effects sizes in improvement on the GRADE was seen in these students.

Furthermore, while control students categorized as "below grade level" decreased in their mean score, those same students in treatment classrooms increased nearly 24 points in their mean scores, with a medium practical effect size. These findings confirm and extend previous research that examined the impact of coaching on student reading achievement. For example, in their study, Biancarosa and her team (Biancarosa et al., 2010) found that children who entered the school year with lower levels of literacy learned at a faster rate than those who entered with higher levels of literacy in classrooms of teachers who were supported by coaches. Similarly, Lockwood and colleagues (Lockwood, McCombs, & Marsh, 2010) found the most noticeable (and positive effects) of their coaching program among students with lower levels of language comprehension at the beginning of the school year. Finally, Porche, Pallante, and Snow (2012) found that "at-risk" students showed greater literacy gains when their teachers received coaching as an intervention.

Collectively, these findings may indicate that students really do need their teachers to be more explicit in their reading instruction. These findings may also suggest that our treatment teachers were learning to become adaptive in their instruction, making decisions for their students based on the selection and adaptation of ideas from a variety of ideologies, methods, materials, and programs (Hoffman & Duffy, 1999). These findings have direct impact on the way in which we view cognitive strategy instruction with exceptional students—those who read below level and those who read above level and the way we think about the professional development of teachers who are committed to meeting the instructional needs of all students. If we want teachers to become reflective and thoughtful, we must provide them with intensive and appropriate professional development to do so.

Fourth, we found a statistical association between the time teachers spent with their coach and the opportunities they offered to engage their students in cognitive strategy instruction. This finding aligns with one other study that examined time as a variable in explaining student outcomes. Specifically, Elish-Piper & L'Allier (2010)'s regression analysis demonstrated a positive relationship between the amount of time literacy coaches spent observing in first grade classrooms and student gains on reading assessments. These are important findings as many studies have shown that coaches' time is divided between time in classrooms and in completing other responsibilities (Bean, Cassidy, Grumet, Shelton, & Wallis, 2002; Bean et al., 2003). Our findings may indicate that a concerted focus of a coach's time in classrooms (working with teachers) may be an effective use of that coach's time.

Finally, we found that the exchanges between coaches and teachers we tested were statistically associated with the instructional practices of the teachers. These findings make a unique contribution to the field of research on coaching. While studies have reported that teachers appreciate their coach (Bean et al., 2010; Lovett et al., 2008), report they learn from their coaches (Nielsen, Barry, & Stabb, 2007), and appreciate feedback from their coach (Nielsen, Barry, & Addison, 2007; Nielsen, Barry, & Stabb, 2008; Scott, Corina, & Carlisle, 2012), our study empirically demonstrates that coaches' behaviors may influence the professional practices of teachers with whom they work.

6. Limitations

Although our findings suggest that the SIPIC model of coaching may have influenced teaching and learning outcomes, there are limitations to this study. First, although we have provided evidence pointing to the importance of each of the components of the SIPIC model, we believe that future investigations may demonstrate the uniqueness of each of the components and their associations to teacher learning. Each of the teachers in this study received some of each of the components; future studies, in which treatment teachers receive one and only one of the exchanges as the intervention, may be more informative.

Second, our coaches acted as outside consultants in this study. We do not know of the impact of the SIPIC model when implemented with school-based coaches. While the contributions of our coaches to the literacy instruction and learning of our participants were valuable, we cannot say from our findings what the impact would be if the intervention mirrored "business as usual" practices. An efficacy trial in which school-based coaches act as the intervention might inform the field. And, the role of peer coaches using the model may inform the field as many teachers already engage in peer coaching practices, albeit informally (Zwart, Wubbels, Bolhuis, & Bergen, 2008). Finally, future research must investigate the support that various types of coaches need in order to successfully implement a model like SIPIC, as there are many different qualities of coaches engaged with teachers across the country (Roller, 2006).

Third, our study looked at the treatment on one group of teachers—those who volunteered to be in the study. Others have documented the positive impact that volunteerism in professional development activities has on the practices of teachers (Linek, Fleener, Fazio, Raine, & Klakamp, 2003; Yamagata-Lynch, 2003). Future studies might look at the implementation of the SIPIC model at a school level and the impact the model has on non-voluntary teachers, both quantitatively and qualitatively.

Fourth, the type of data we collected may have limited our study. We were unable to collect longitudinal data on our participating teachers and subsequently, we do not know what long lasting effects the model had on their instruction. Additionally, as other studies have demonstrated, there is variability within teachers' reading lessons (Croninger & Valli, 2009). More than two observations in each classroom might have yielded a deeper understanding of the implementation of the practices we measured. Similarly, the model was only tested in the support of improving instructional reading practices. It should be tested on the support it may offer (or not) to math and science teachers and teachers at various grade levels. Finally, we were not able to collect data that would allow us to explore the ways in which coaches relied on responsive or directive practices, or how they wove their way through the two in mediated ways. This may be a topic for future research.

Finally, we did not compare the SIPIC model to other existing (and effective) models of professional development situated in the same paradigm. For example, Putnam and Borko (2000) describe two models of professional development. The first, "break set" (p. 6) is where teachers learn to re-think their instruction in summer workshops housed in sites other than school buildings. The second, case-based learning is where teachers are provided with vicarious encounters through cases. These forms of professional development might be compared to the SIPIC model for further testing.

7. Conclusions

The field of reading education has embraced coaching for a variety of reasons. Some may be because coaching is "hot" (Cassidy & Ortlieb, 2013), while other reasons might have to do with policy mandates. We believe that our findings offer compelling evidence that SIPIC is a viable model of professional development for classroom teachers across grade levels and content areas when it is focused on reading instruction.

The findings of this study remind us that teachers need time and intensive support in their classrooms by highly qualified coaches in order to improve their practices. One decade ago in her presidential address at the annual meeting of the American Educational

Research Association, Borko (2004) called for researchers to engage in different types of inquiries using a vast array of research tools to generate sources of knowledge that help the field make informed decisions about the types of professional development we offer to teachers. This message was the impetus for and our continued line of research, seeking to understand the role of coaches, teachers as learners, and the context of the professional development as a way of improving professional development for classroom reading teachers. Further, Borko encouraged researchers to consider the balances and trade-offs between fidelity of implementation and adaptation of the program to its context and to consider which elements of the program must be preserved to ensure the integrity of its underlying goals and principles. Future documentation of "mutual adaptation" of the SIPIC model may yield contributions to our understanding of the types and levels of support teachers need to become better reading teachers and the types and levels of support coaches need in order to help teachers do so.

Acknowledgments

References

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher, 61*, 364—373.

Almasi, J. F. (2003). *Teaching strategic processes in reading.* New York: Guildford Press.

Alvermann, D., Commeyras, M., Cramer, S., & Harnish, D. (2006). *Evaluation of Georgia Reading First implementation, progress, and impact 2004—2005.* College of Education, University of Georgia.

American Educational Research Association/American Psychological Association/National Council on Measurement in Education. [AERA/APA/NCME]. (1999). *Standards for educational and psychological testing.* Washington, DC: Author.

American Guidance Services. (2001). *Group Reading Assessment and Diagnostic Evaluation (GRADE).* Circle Pines, MN: Author.

Anderson, J. R., Reder, L. M., & Simon, H. A. (2000, Summer). *Applications and Misapplications of Cognitive Psychology to Mathematics Education. Texas Educational Review.* Available at http://act-r.psy.cmu.edu/publications/pubinfo.php?id=146.

Armstrong, K., Cusumano, D. L., Todd, M., & Cohen, R. (2008). Beliefs about implementing literacy skills scale. *Journal of Early Childhood Teacher Education, 29*(4), 297—308.

Bean, R. M., Cassidy, J., Grumet, J. V., Shelton, D., & Wallis, S. R. (2002). What do reading specialists do? Results from a national survey. *Reading Teacher, 55*, 2—10.

Bean, R. M., Draper, J. A., Hall, V., Vandermolen, J., & Zigmond, N. (2010). Coaches and coaching in Reading First schools: a reality check. *Elementary School Journal, 111*, 87—114.

Bean, R. M., Swan, A. L., & Knaub, R. (2003). Reading specialists in schools with exemplary reading programs: functional, versatile, and prepared. *Reading Teacher, 56*, 446—455.

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*, 7—34.

Birman, B. F., Desimone, L., Porter, A. C., & Garet, M. S. (2000). Designing professional development that works. *Educational Leadership, 57*, 28—33.

Block, C. C., & Pressley, M. (2002). *Comprehension instruction: Research-based best practices.* New York: Guilford Press.

Borko, H. (2004). Professional development and teacher learning: mapping the terrain. *Educational Researcher, 33*, 3—15.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32—42.

Butler, D. L., Lauscher, H. N., Jarvis-Selinger, S., & Beckingham, B. (2004). Collaboration and self-regulation in teachers' professional development. *Teaching and Teacher Education, 20*, 435—455.

Carlisle, J., & Berebitsky, D. (2010). Literacy coaching as a component of professional development. *Reading and Writing, 24*(7), 773—800.

Cassidy, J., Garrett, S. D., Maxfield, P., & Patchett, C. (2010). Literacy coaching: yesterday, today and tomorrow. In J. Cassidy, S. D. Garrett, & M. Sailors (Eds.), *Literacy Coaching: Research & Practice.* Corpus Christi, TX: Center for the Development of Educational Research.

Cassidy, J., & Ortlieb, E. (2013). What was hot (and not) in literacy. *Journal of Adolescent & Adult Literacy, 57*, 21—29.

Cochran-Smith, M., & Lytle, S. (1999). Relationship of knowledge and practice: teacher learning in communities. In A. Iran-Nejad, & C. D. Pearson (Eds.), *Review of research in education* (Vol. 24, pp. 249—306). Washington, DC: American Educational Research Association.

Cohen, J. (1988). *Statistical power for the behavioral sciences.* New York: Laurence Earlbaum.

Collins, A., Brown, J. S., & Newman, S. E. (1987). *Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics.* Technical Report No. 403. (No. 403). Urbana, Illinois: Illinois University.

Costa, A. L., & Garmston, R. J. (2005). *Cognitive coaching: A foundation for renaissance schools.* Christopher-Gordon Publishers.

Croninger, R. G., & Valli, L. (2009). "Where is the action?" Challenges to studying the teaching of reading in elementary classrooms. *Educational Researcher, 38*, 100—108.

Denton, C. A., & Hasbrouck, J. (2009). A description of instructional coaching and its relationship to consultation. *Journal of Educational & Psychological Consultation, 19*(2), 150—190.

Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research, 72*, 433—450.

Deussen, T., Coskie, T., Robinson, L., & Autio, E. (2007). *"Coach" can mean many things: Five categories of literacy coaches in Reading First* (Issues & Answers Report, REL 2007-No. 005). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. Available at http://ies.ed.gov/ncee/edlabs/regions/northwest/pdf/REL_2007005.pdf.

Diamond, K. E., & Powell, D. R. (2011). An iterative approach to the development of a professional development intervention for head start teachers. *Journal of Early Intervention, 33*, 75—93.

Downer, J., Kraft-Sayre, M. E., & Pianta, R. C. (2009). Ongoing, web-mediated professional development focused on teacher—child interactions: early childhood educators' usage rates and self-reported satisfaction. *Early Education and Development, 20*, 321—345.

Dozier, C. (2006). *Responsive literacy coaching: Tools for creating and sustaining purposeful change.* Portland, ME: Stenhouse.

Duffy, G. G., Roehler, L. R., Meloth, M. S., Vavrus, L. G., Book, C., Putnam, J., et al. (1986). The relationship between explicit verbal explanations during reading skill instruction and student awareness and achievement: a study of reading teacher effects. *Reading Research Quarterly, 21*, 237—252.

Duffy, G. G., Roehler, L. R., Sivan, E., Rackliffe, G., Book, C., Meloth, M. S., et al. (1987). Effects of explaining the reasoning associated with using reading strategies. *Reading Research Quarterly, 23*, 347—368.

Duffy, G. G. (2002). The case for direct explanation of strategies. In C. C. Block, & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 28—41). New York: Guilford Press.

Duffy, G. G. (2009). *Explaining reading. A resource for teaching concepts, skills, and strategies.* Guilford Press.

Elish-Piper, L., & L'Allier, S. K. (2010). Exploring the relationship between literacy coaching and student reading achievement in grades K-1. *Literacy Research and Instruction, 49*, 162—174.

Ferguson, K. (2011). Examining the relationship of power and resistance in literacy coaching in three school contexts. In P. J. Dunston, L. B. Gambrell, K. Headley, S. K. Fullerton, P. M. Stecker, V. R. Gillis, et al. (Eds.), *60th yearbook of the Literacy Research Association* (pp. 165—178). Oak Creek, WI: Literacy Research Association.

Fleiss, J. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: John Wiley & Sons.

Gersten, R., Morvant, M., & Brengelman, S. (1995). Close to the classroom is close to the bone: coaching as a means to translate research into classroom practice. *Exceptional Children, 62*, 52—66.

Glazer, E. M., & Hannafin, M. J. (2006). The collaborative apprenticeship model: situated professional development within school settings. *Teaching and Teacher Education, 22*, 179—193.

Greeno, J. G. (1997). Response: on claims that answer the wrong questions. *Educational Researcher, 26*(1), 5—17.

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist, 53*(1), 5—26.

Greeno, J. G., Smith, D. R., & Moore, J. L. (1993). Transfer of situated learning. In D. K. Detterman, & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 99—167). Norwood, NJ: Ablex.

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching: Theory and Practice, 8*, 381—391.

Herbert, J., & Attridge, C. (1975). A guide for developers and users of observation systems and manuals. *American Educational Research Journal, 12*, 1—20.

Hoffman, J. V., & Duffy, G. G. (1999). In pursuit of an illusion: the flawed search for a perfect method. *Reading Teacher, 53*, 10—16.

Hsieh, W. Y., Hemmeter, M. L., McCollum, J. A., & Ostrosky, M. M. (2009). Using coaching to increase preschool teachers' use of emergent literacy teaching strategies. *Early Childhood Research Quarterly, 24*(3), 229—247.

International Reading Association. (2004). *The role and qualifications of the reading coach in the United States.* Newark, DE: Author.

Ippolito, J. (2010). Three ways that literacy coaches balance responsive and directive relationships with teachers. *The Elementary School Journal, 111*(1), 164—190.

Kucan, L. (2007). Insights from teachers who analyzed transcripts of their own classroom discussions. *The Reading Teacher, 61*, 228—236.

Kuijpers, J. M., Houtveen, A. A. M., & Wubbels, T. (2010). An integrated professional development model for effective teaching. *Teaching and Teacher Education, 26*, 1687—1694.

Lapp, D., Fisher, D., Flood, J., & Frey, N. (2003). Dual role of the urban reading specialist. *Journal of Staff Development, 24*(20), 33–36.

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

Lieberman, A. (1995). Practices that support teacher development. *Phi Delta Kappan, 76*, 591–596.

Linek, W. M., Fleener, C., Fazio, M., Raine, I. L., & Klakamp, K. (2003). The impact of shifting from "How teachers teach" to "How children learn". *Journal of Educational Research, 97*, 78–89.

Lipson, M. L., Mosenthal, J. H., Mekkelsen, J., & Russ, B. (2004). Building knowledge and fashioning success one school at a time. *The Reading Teacher, 57*, 534–542.

Lockwood, J. R., McCombs, J. S., & Marsh, J. (2010). Linking reading coaches and student achievement: evidence from Florida middle schools. *Educational Evaluation and Policy Analysis, 32*, 372–388.

Lovett, M. W., Lacerenza, L., De Palma, M., Benson, J. J., Steinbach, K. A., & Frijters, J. C. (2008). Preparing teachers to remediate reading disabilities in high school: what is needed for effective professional development? *Teaching and Teacher Education, 24*, 1083–1097.

Lynch, J., & Ferguson, K. (2010). Reflections of elementary school literacy coaches on practice: roles and perspectives. *Canadian Journal of Education, 33*, 199–227.

Marsh, J. A., McCombs, J. S., & Martorell, F. (2010). How instructional coaches support data-driven decision making: policy implementation and effects in Florida middle schools. *Educational Policy, 24*, 872–907.

Marsh, J. A., McCombs, J. S., & Martorell, F. (2012). Reading coach quality: findings from Florida middle schools. *Literacy Research and Instruction, 51*, 1–26.

Martin, J. (1977). The development and use of classroom observation instruments. *Canadian Journal of Education, 2*, 43–54.

Matsumara, L. C., Garnier, H. E., Correnti, R., Junker, B., & Bickel, D. D. (2010). Investigating the effectiveness of a comprehensive literacy coaching program in schools with high teacher mobility. *Elementary School Journal, 111*, 35–62.

McDonald, J. P. (1989). When outsiders try to change schools from the inside. *Phi Delta Kappan, 71*, 206–212.

Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: development, measurement, and validation. *American Journal of Evaluation, 24*, 315–340.

Mraz, M., Algozzine, B., & Watson, P. (2008). Perceptions and expectations of roles and responsibilities of literacy coaching. *Literacy Research and Instruction, 47*, 141–157.

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

Nielsen, D. C., Barry, A. L., & Addison, A. B. (2007). A model of a new-teacher induction program and teacher perceptions of beneficial components. *Action in Teacher Education, 28*, 14–24.

Nielsen, D. C., Barry, A. L., & Stabb, P. T. (2008). Teachers' reflections of professional change during a literacy-reform initiative. *Teaching and Teacher Education, 24*, 1288–1303.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research, 78*, 33–84.

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8*, 293–316.

Pianta, R., Mashburn, A., Downer, J., Hamre, B., & Justice, L. (2008). Effects of web-mediated professional development resources on teacher–child interactions in pre-kindergarten classes. *Early Childhood Research Quarterly, 23*(4), 431–451.

Porche, M., Pallante, D. H., & Snow, C. E. (2012). Professional development for reading achievement results from the Collaborative Language and Literacy Instruction Project (CLLIP). *Elementary School Journal, 112*, 649–671.

Pressley, M. (2000). What should comprehension instruction be the instruction of?. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 545–562) Mahwah, NJ: Lawrence Erlbaum.

Pressley, M., Borkowski, J. G., & Schneider, W. (1989). Good information processing: what it is and how education can promote it. *International Journal of Educational Research, 13*, 857–867.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*, 4–15.

Raudenbush, S., Bryk, A., Cheong, F., Congdon, R., & DuToit, M. (2004). *Optimal HLM 6.0 Software [Computer software]*. Chicago: Scientific Software.

Roller, C. (2006). *Reading and literacy coaches report on hiring requirements and duties survey*. Newark, DE: International Reading Association.

Sailors, M. (2008). *Comprehension Instruction Observation Protocol: Capturing the instructional comprehension practices of classroom teachers, grades 2–8 and the metacognition of their students*. Data Collection Instrument. San Antonio, TX: The University of Texas at San Antonio.

Sailors, M., & Price, L. (2010). Professional Development that supports the teaching of cognitive reading strategy instruction. *Elementary School Journal, 110*, 301–323.

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

Sailors, M., Minton, S., & Villarreal, L. (2013). *A systematic review of literacy coaching, Annual Meeting of the Literacy Research Association*. Dallas, TX: Literacy Research Association.

Sailors, M., Hoffman, J. V., Pearson, P. D., McClung, N., Shin, J., Phiri, L., & Saka, T. (2014). Supporting change in literacy instruction in Malawi. *Reading Research Quarterly, 49*, 209–231.

Scott, S. E., Cortina, K. S., & Carlisle, J. F. (2012). Understanding coach-based professional development in Reading First: how do coaches spend their time and how do teachers perceive coaches' work? *Literacy Research and Instruction, 51*, 68–85.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin Co.

Sheridan, S. M., Edwards, C. P., Marvin, C. A., & Knoche, L. L. (2009). Professional development in early childhood programs: process issues and research needs. *Early Education and Development, 20*(3), 377–401.

Sheridan, S. M., & Kratochwill, T. R. (2008). *Conjoint behavioral consultation: Promoting family–school connections and interventions* (2nd ed.). New York, NY: Springer Publishing.

Shulman, L. S. (1986). Those who understand: knowledge growth in teaching. *Educational Researcher, 15*, 4–14.

Shulman, J. H. (2004). From inspired vision to impossible dream: the dangers of imbalanced mentoring. *Journal of Teacher Education, 55*, 393–406.

Singer, J., & Willett, J. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Sperry, L. (2008). Executive coaching: an intervention, role function, or profession? *Consulting Psychology Journal: Practice and Research, 60*, 33–37.

Van Keer, H., & Verhaeghe, J. P. (2005). Effects of explicit reading strategies instruction and peer tutoring on second and fifth graders' reading comprehension and self-efficacy perception. *Journal of Experimental Education, 73*, 291–329.

Von Eye, A., & Niedermeier, K. E. (1999). *Statistical analysis of longitudinal categorical data in the social and behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum.

Vygotsky, L. S. (1978). *Mind in society: Development of higher psychological processes*. Cambridge, MA: Harvard Educational Press.

Walpole, S., & Blamey, K. (2008). Elementary literacy coaches: the reality of dual roles. *Reading Teacher, 62*(3), 222–231.

Walpole, S., McKenna, M. C., Uribe-Zarain, X., & Lamitina, D. (2010). The relationship between coaching and instruction in the primary grades: evidence from high-poverty schools. *Elementary School Journal, 111*, 115–140.

Ward, J. R., & McCotter, S. S. (2004). Reflection as a visible outcome for preservice teachers. *Teaching and Teacher Education, 20*, 243–257.

Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting with teacher professional development: motives and methods. *Educational Researcher, 37*, 469–479.

Whitaker, S., Kinzie, M., Kraft-Sayre, M. E., Mashburn, A., & Pianta, R. C. (2007). Use and evaluation of web-based professional development services across participant levels of support. *Early Childhood Education Journal, 34*, 379–386.

Wilkinson, G. S. (1993). *WRAT-3: Wide range achievement test administration manual/Gary S. Wilkinson*. Wilmington, Del: Wide Range, Inc.

Wolcott, H. F. (1995). *The art of fieldwork*. Walnut Creek, CA: AltaMira Press.

Yamagata-Lynch, L. C. (2003). How a technology professional development program fits into teachers' work life. *Teaching and Teacher Education, 19*, 591–607.

Zwart, R. C., Wubbels, T., Bolhuis, S., & Bergen, T. C. M. (2008). Teacher learning through reciprocal peer coaching: an analysis of activity sequences. *Teaching and Teacher Education, 24*, 982–1002.