Effects of ESL Instructional Coaching on Secondary Teacher Use of Sociocultural Instructional Practices

ANNELA TEEMANT,
Indiana University-Purdue University Indianapolis

YUHAO CEN,
Shanghai Jiao Tong University, China

AMY WILSON,
Indiana University-Purdue University Indianapolis

This longitudinal and descriptive quantitative study investigates the efficacy of an English as a Second Language (ESL) instructional coaching intervention with urban secondary teachers (N = 22). Coached teachers participated in a 30-hour workshop and then six cycles of coaching targeting use of five research-based sociocultural principles of (language) learning called the Standards for Effective Pedagogy. Findings demonstrate instructional coaching led to unique and statistically significant (a) pedagogical transformation and (b) patterns of development for STEM and non-STEM secondary teachers. Implications for improving the professional development model for STEM teachers are discussed.

Key Words: Instructional coaching; Professional development; Sociocultural pedagogy; Urban Secondary Teachers; STEM.

Instructional Practices

Promoting language learning in secondary core academic classrooms is a pressing need given the steady increase of English language learners (ELLs) in American classrooms (Aud, et al., 2013; National Center for Education Statistics, 2010). Teacher preparation, unfortunately, has not kept pace with shifting demands, leaving predominantly White, middle-class teachers unprepared to

meet the educational needs of multilingual, multicultural, low-income, and urban student populations (Hollin & Guzman, 2005; Sleeter, 2008).

Secondary educators are especially challenged to serve diverse student populations equitably. Horn (2012) observed, "historically marginalized groups of students are often severely underserved in mathematics classrooms" (p. 6). Crisp and Nora (2012) credit the persistent under-representation of minorities in science and mathematics to ongoing "disparities in teacher quality, school funding, and monies spent on instructional resources" (p.4). Both new standards in mathematics (Common Core State Standards Initiative, 2012) and science (NGSS Lead States, 2013) call for the integration of language and literacy into instruction (Lee, Quinn, & Valdéz, 2013). New standards envision all students understanding, communicating, and applying learning in real world contexts (NGSS Lead States, 2013). Such new student performance expectations are pushing STEM educators "to create *learning environments* that support student participation" (Horn, 2012, p. 3).

These evolving teacher expectations are radical—not minor—shifts for secondary educators, who are used to traditional lecture-based pedagogy. As teachers of culturally and linguistically diverse students, secondary teachers need more than knowledge of language to promote language learning. They need more than a checklist of strategies, accommodations, or better scripting of lessons to increase achievement. In this paper, we argue secondary teachers need to first experience a fundamental shift in their theory and pedagogy of learning to create the kind of learning environment that would result in ELLs' understanding, communicating, and applying knowledge in real world contexts.

This paper examines the effectiveness of an English as a Second Language (ESL) instructional coaching intervention (Teemant, 2014; Teemant & Reveles, 2012; Teemant, Wink,

& Tyra, 2011) targeting secondary mathematics/science (STEM) and non-STEM teachers. Building upon sociocultural theory (Vygotsky, 1978), the intervention focuses on use of five sociocultural practices called the Standards for Effective Pedagogy or Five Standards (Tharp, Estrada, Dalton, & Yamauchi, 2000). Tharp et al. theorize that ELL student achievement increases to the degree students are asked to collaborate, use rich language and literacy, build on prior knowledge and experience, use higher order thinking skills, and discuss their learning with ongoing assistance and feedback. The instructional model combines teacher use of the Five Standards with small group configurations.

This descriptive longitudinal paper (a) describes the instructional coaching intervention and (b) presents quantitative evidence of teacher change, patterns of development, and STEM and non-STEM teacher differences in use of the Five Standards instructional model. This study contributes to the professional development knowledge base by documenting the successes and challenges of pursuing radical pedagogical change with urban secondary teachers for the benefit of ELLs.

Instructional Coaching in Theory and Practice

Teacher professional development is one of the most productive ways to improve student achievement (e.g., Guskey, 2000), and instructional coaching, as a particular type of coaching (Knight, 2009), is considered a highly effective component of a professional development process (e.g., Cornett & Knight, 2009; Joyce & Showers, 1995; Sailors & Shanklin, 2010). Instructional coaching provides teachers support in the classroom on an ongoing, collaborative, and extended basis. These characteristics create opportunities for high quality professional development (Desimone, 2009; Wei et al., 2010).

The theoretical foundations, focus, and content of coaching models vary greatly. Sociocultural theory (Vygotsky, 1978) is considered essential in the preparation of teachers of culturally and linguistically diverse learners because it explicitly recognizes that learning is social and knowledge is cultural (Rogoff, 1990; Tharp & Gallimore, 1988). For Vygotsky (1997), teachers are situated as more knowledgeable others who play an active role in student learning. It is the interaction between teacher and students that allows students to receive assistance in learning concepts. Assistance is considered most valuable when it is situated within the zone of proximal development (ZPD), slightly beyond current abilities.

Although several qualitative (e.g., Arnau, Kahrs, & Kruskamp, 2004; Brown, Reumann-Moore, Hugh, Christman, & Riffer, 2009; Carrera, 2010), quantitative (e.g., Hearn, 2010; Vogt & Rogalla, 2009), and mixed methods (e.g., Marsh, Sloan McCombs, & Martorell, 2010; Murray, Ma, & Mazur, 2009) studies of coaching with secondary educators have been conducted, none have explicitly focused on sociocultural pedagogical practices. According to Tharp et al., (2000), at the highest level of sociocultural practice, teachers engage in: (a) *Joint Productive Activity*—a teacher and small group of students produce 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*—defining expectations, and then providing assistance and feedback to students; and (e) *Instructional Conversation*—engaging a small group of students in a sustained, student-dominated, goal-directed academic conversation that questions rationales and assists learning.

For professional development and evaluation purposes, Doherty, Hilberg, Epaloose, and Tharp (2002) developed a classroom observation tool called the Standards Performance

Continuum (SPC). Figure 1 defines incremental and observable steps a teacher can take to move from predominately behavioristic and teacher-dominated instruction at the "Not Observed" level toward increasingly sociocultural practices at the "Enacting" level. Teachers promote the most learning when they use at least three of the Five Standards in a single activity. A growing number of studies connect teacher fidelity to the Five Standards instructional model with increased English attainment and academic achievement for both native and non-native speakers of English (e.g., Doherty & Hilberg, 2007; Doherty, Hilberg, Pinal, & Tharp, 2003; Estrada, 2005; Saunders & Goldenberg, 1999; Author et al., 2013; Tharp, 1982).

In the instructional coaching context, the SPC defines the performance targets for teachers and allows coaches to document the evolving quality of teacher implementation of the Five Standards over time. Several recent studies document the efficacy of this instructional coaching model (a) with elementary teachers ((Teemant, 2014; Teemant & Reveles, 2012; Teemant, Wink, & Tyra, 2011), and (b) in producing significant student achievement gains (Teemant & Hausmann, 2013). What has yet to be explored is whether urban secondary teachers benefit to the same degree.

The Five Standards instructional coaching process has been described in detail in Hilberg, Doherty, and Reveles (2004). Similar to previous instructional coaching studies (Teemant, 2014; Teemant & Reveles, 2012), secondary teachers participated in a 30-hour university workshop, where sociocultural theory, the Five Standards instructional model, and the SPC observation tool were the focus on learning using small group activity centers, video clips, and readings. Teachers were also introduced to a 12-week phase in process, which detailed norms, procedures, and logistics for moving from predominately whole-class instruction to multiple, simultaneous, and differentiated small group activity centers (see Hilberg, Chang, and Epaloose, 2003).

Second, teachers participated in six individual coaching sessions across the school year (i.e., approximately 12 contact hours). With federal grant funding[1], one instructional coach was hired as an expert external consultant with six years of Five Standards instructional coaching experience. The coach has experience as urban bilingual elementary teacher, is National Board and Reading Recovery Certified, and has a Ph.D. in Literacy Education.

The Five Standards define the growth targets for teachers and guide the instructional coaching process itself. After an introductory interview establishing rapport, norms, and expectations, each coaching cycle is a three-stage process: 30-minute pre-conference, 45-minute observation, and 30-minute post-conference. The coach-teacher pre- and post-conferences are Instructional Conversations: that is, they are goal-directed, conversational, negotiated, and driven by teacher readiness. The pre-conference focuses on collaborative lesson planning. During the classroom observation, the coach gathers evidence of student interaction and thinking, questioning patterns, types of assistance, and SPC ratings. The post-conference conversation elicits reflection on teacher and student learning.

The Five Standards instructional model is intended for use in any content area or grade level and is not prescriptive. The teacher maintains control of the curriculum and instruction although the teacher is encouraged with the growth targets to design activities that promote collaboration, language use, connected learning, cognitive complexity, and teacher-student dialogue (Doherty et al., 2002).

In summary, this study of instructional coaching, with STEM and non-STEM regular classroom teachers, explores the efficacy of the Five Standards instructional coaching model as a

[1] 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.

means of promoting visible instructional innovation among urban secondary teachers. The following quantitative research questions (RQs) guide this study:

RQ1. Transformation: Is there an increase in teacher use of the Five Standards as measured by individual standards and total score?

RQ2. Pattern of Development: What patterns of implementation for the Five Standards Instructional Model emerge across coaching cycles for all teachers?

RQ 3. STEM: Are there significant differences between STEM and NON-STEM teacher groups in transformation or patterns of development?

Methods

This study is principally a descriptive and longitudinal quantitative study that evaluates the effectiveness of a one-year ESL instructional coaching intervention with secondary urban teachers. Based on previous studies, teacher use of sociocultural teaching practices is expected to increase as a result of coaching. Teacher fidelity to the Five Standards instructional model is measured with the SPC, using a repeated measures design across six cycles of instructional coaching. Quantitative data provides evidence of teacher transformation, patterns of development, and STEM/non-STEM comparisons. The context and participants, instrument, and analyses are described below.

Context and Participants

With U.S. Department of Education funding, teachers from two urban secondary schools participated in this one-year study. During the 2009-2010 school year, both participating secondary schools had high student diversity and poverty (School 1 = 42% White; 23% Black, 30% Hispanic, and 5% multi-racial with 76% on free/reduced lunch; School 2 = 7% White; 64% Black, 25% Hispanic, and 4% multi-racial with 71% free/reduced lunch), poor graduation rates

(School 1 = 47%; School 2 = 50%), and were in a fifth year of academic probation. After a faculty meeting where the expectations, benefits, funding, and informed consent process were described, more than 90% of the teaching staff supported school participation whether or not they planned to personally participate.

Participants were 22 urban teachers (4 male) from two high schools (School 1=12; School 2=10 teachers) who completed a 30-hour summer workshop and six cycles of Five Standards instructional coaching. Coached teachers were paid $2,000. Five teachers were STEM and 17 were non-STEM teachers. Teachers were 73% White, 18% Black, and 5% Hispanic, averaging 13.9 years of experience (12 < 13 years). Class sizes were small (School 1 = 13.2; School 2 = 15.8), with grades 7 (18%), 9 (18%), 10-11 (18%), and 9-12 (46%) represented.

Instrument

As described in a previous study (Teemant, 2014), fidelity to the Five Standards Instructional Model was quantitatively measured using the Standards Performance Continuum (SPC). As Figure 1 shows, 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 x 3 rule) standards are fully enacted within a single activity. Four points are available per standards 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=

	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.
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 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 teacher makes incidental connections between students' prior experience/knowledge from home, school, or community and the new activity/academic concepts.	The teacher integrates the new activity/academic concepts with students' prior knowledge from home, school, or community to connect everyday and schooled concepts.	The teacher designs, enacts, and assists in contextualized activities that demonstrate skillful integration 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.
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.
Critical Stance *Teaching to Transform Inequities*	Instruction reflects appropriate content area standards.	The teacher designs instruction using variety, which includes (a) multiple sources of information; OR (b) values and respects multiple perspectives; OR (c) supports learning through multiple modalities.	Using variety, the teacher designs instruction that positions students to generate new knowledge resulting in (a) original contributions, products, or expertise; OR (b) students' questioning and reflecting on issues from multiple perspectives.	The teacher designs or facilitates instruction that consciously engages learners in (a) interrogating conventional wisdom and practices; AND (b) reflection upon ramifications of such practices; AND (c) actively seeks to transform inequities within their scope of influence within the classroom and larger community.	The teacher designs, enacts, and assists in critical stance activities that demonstrate skillful integration of multiple standards simultaneously.

12.50 – 17.49; and (d) integrating= 17.50 – 20.00. For more description of the SPC see Doherty et al. (2003).

Three raters established SPC inter-rater reliability for this study. Each rater had six years of experience in using the instrument as instructional coaches and external evaluators. They received five days of SPC training from its developers (2002-2003), and later spent four 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 x 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.

Data analysis

Dependent variables include ratings for each standard and total score. Data analysis occurred in two stages. First, frequencies, means, and standard deviations were calculated. Second, one-way repeated measure ANOVAs were conducted to reveal patterns of development for each of the Five Standards. In order to understand differences in STEM and non-STEM teachers, a *t*-test was conducted for group based on their baseline observation. This analysis revealed if patterns of ultimate development varied by STEM/nonSTEM. Mauchly's test of sphericity resulted in corrections as needed. In general, the Wilks' Lambda value, *F* statistics, and a partial eta-squared value are reported. Effect sizes are defined by Cohen (1988) as small (<.20), medium

(>.20 and <.79) and large (>.80). Tests of within-subjects contrasts identify significant linear and quadratic trends in the data across cycles. Marked line graphs of development are provided.

Results

The results of this study are reported by research question (RQ). Teacher use of the Five Standards, patterns of development, and differences between STEM and non-STEM teachers follow.

1. Teacher Use of the Five Standard. Table 1 presents the means and standard deviations for each standard and total score by coaching cycle. In response to RQ1, secondary teachers' use of each standard consistently increased from coaching cycle one to three. Language and Literacy Development, Challenging Activities, and the Instructional Conversation each showed a plateau-effect during cycle four, experienced a drop during cycle five, and increased during cycle six again. Both Joint Productive Activity and Contextualization plateaued during cycles four and five, with Contextualization increasing again during the sixth coaching cycle. Joint Productive Activity showed the greatest variation, as measured by standard deviations, among teachers across all coaching cycles, with the exception of Contextualization in cycle two.

The one-way repeated measure ANOVAs revealed teacher growth was statistically significant in use of Language/Literacy Development, Challenging Activities, and Total Score with medium effect sizes: (a) Language/Literacy, Wilks' Lambda =0.29 , $F(5,17) = 8.35$, $p < .000$, partial eta-squared = .71; (b) Challenging Activities, Wilks' Lambda = 0.46, $F(5,17) = 3.95$, $p = .015$, partial eta-squared = .54; (c) Total Score, Wilks' Lambda = 0.50, $F(5,17) = 3.46$, $p = .024$, partial eta-squared = .50.

The LSD comparisons revealed significantly greater mean use of each standard at coaching cycle six than at coaching cycle one. On a four-point scale, the gain scores (GS) in

Five Pedagogical Standards		Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6
Joint Productive Activity (JPA)	*M*	1.23	2.14	2.18	2.09	2.18	2.09
	SD	0.81	1.25	1.37	1.54	1.47	1.51
Language/Literacy Development (LLD)	*M*	1.41	2.32	2.45	2.41	2.32	2.64
	SD	0.50	1.13	1.14	1.26	1.25	1.09
Contextualization (CTX)	*M*	1.23	2.00	2.05	1.95	2.05	2.27
	SD	0.61	1.27	1.36	1.33	1.36	1.28
Challenging Activities (CA)	*M*	1.32	1.82	2.18	2.27	1.86	2.55
	SD	0.57	0.80	1.22	1.20	1.04	1.18
Instructional Conversation (IC)	*M*	0.77	1.09	1.59	1.59	1.45	1.68
	SD	0.43	0.92	1.30	1.30	1.26	1.39
Total Score	*M*	5.95	9.36	10.50	10.32	9.86	11.23
	SD	2.04	4.65	6.01	6.17	5.89	6.06

Total N=22

Table 1: *Means and Standard Deviations for Five Standards and Total Score by Coaching Cycle*

order from most to least growth follow: Language/ Literacy (GS = 1.23), Challenging Activities from (GS = 1.23), Contextualization from (GS = 1.04), Instructional Conversation (GS = 0.91), and Joint Productive Activity (GS = 0.86). Despite this growth (Total Score gain = 5.28), teachers on average only achieved a developing level of fidelity (Total Score *M* between 7.50 and 12.49) by coaching cycle six (*M* = 11.23, *SD* = 6.06), falling two levels short of the highest

or integrating level (M >17.50). In practice, this means teachers moved away from behaviorist-oriented and teacher-dominated instructional practices but were not, on average, able to realize a sociocultural classroom, rich with challenging, contextualized, and sustained teacher- and peer-assisted small group collaboration. Students were interacting only briefly with peers on cognitively challenging activities without teacher assistance, feedback, or clear expectations for performance.

2. General Patterns of Development. Teachers' cycle-to-cycle pattern of development was investigated for each standard using repeated measures contrasts and graph line plots of means (RQ2). Analysis revealed a significant linear trend for four individual standard and Total Score with medium effect sizes: Language/Literacy, $F(1,21) = 15.22$, $p = .001$, partial eta-squared = .42; Contextualization, $F(1,21) = 6.94$, $p = .015$, partial eta-squared = .25; Challenging Activities, $F(1,21) = 12.41$, $p = .002$, partial eta-squared = .37; Instructional Conversation, $F(1,21) = 7.52$, $p = .012$, partial eta-squared = .26; Total Score, $F(1,21) = 9.68$, $p = .005$, partial eta-squared = .32. This upward linear trend demonstrates teachers generally improved from cycles one to six on all standards except consistent use of Joint Productive Activity.

The data also revealed a single bend trend—or a significant quadratic trend—for two standards and the total score with medium effect sizes: Joint Productivity, $F(1,21) = 5.81$, $p = .025$, partial eta-squared = .22; Language/Literacy, $F(1,21) =10.22$, $p = .004$, partial eta-squared = .33; and Total Score, $F(1,21) = 5.65$, $p = .027$, partial eta-squared = .21. The quadratic trend demonstrates that teachers experienced slight declines in implementing Joint Productive Activity and Language/Literacy during coaching cycles four and five, which is also reflected in the Total Score.

A cubic trend with a medium effect size was revealed in the development pattern of Challenging Activities, $F(1,21) = 2.00$, $p = .017$, partial eta-squared = .24. After consistent improvement from cycle one to cycle four, teachers experienced a drop in cycle five and a rise again during cycle six.

Figure 2 visually presents these patterns of fluctuations comparing each standard to each other. First, from cycle one to six, teachers used Language and Literacy Development at higher levels than the other standards, which suggests this standard was the easiest for teachers to implement. Second, the Instructional Conversation was the least observed standard in use across the cycles, suggesting it is the most difficult for secondary teachers to implement. Lastly, the development patterns for Joint Productive Activity, Contextualization, and Challenging Activities reveal the standards teachers struggle to implement with consistency.

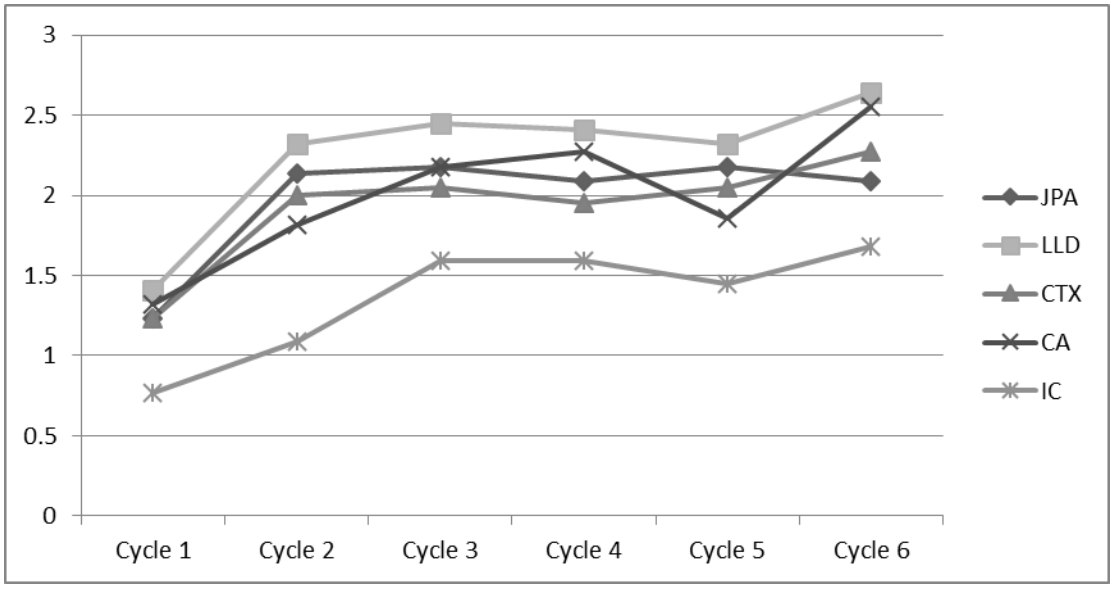

Figure 2. Marked line graph of each standard in contrast to each other.

In practice, these findings suggest that secondary teachers require more than six coaching sessions to achieve the highest level of fidelity to the model. Secondary teachers even with instructional coaching support were unable to consistently create the classroom conditions for language learning. Specifically, secondary urban teachers struggle with effective and consistent use of small group configurations that would create improved conditions and the likelihood for peer and teacher assistance, feedback, language use, and learning.

3. STEM/Non-STEM Groups. Table 2 presents the means and standard deviations for use of the Five Standards by coaching cycle for STEM and non-STEM teachers (RQ3). Four patterns stand out: (a) STEM teachers consistently enacted the Five Standards Instructional Model at a lower level than non-STEM teachers across coaching cycles two to six; (b) STEM teachers increased use of Joint Productivity from cycle one to two, but then they experienced a consistent decline thereafter; (c) STEM teachers struggled in implementing Challenging Activities in their classrooms, as reflected by the fluctuation in the mean scores across the cycles; and (d) the STEM group enacted the Instructional Conversation least among the Five Standards. These patterns demonstrate that STEM and non-STEM teachers have different patterns of development and responsiveness to instructional coaching when the Five Standards are the growth targets.

Figure 3 presents a graph comparing STEM and non-STEM groups by Total Score across coaching cycles. The *t*-test on Total Score shows no significant group difference at cycle one, $p = .502$. By cycle six, the Total Score group differences are significant, $p = .049$ with a large effect size ($d = 1.26$). For STEM teachers, the Total Score did not increase from coaching cycle one to cycle six; for non-STEM teachers, the Total Score increase from cycle one to six was significant, $p = .001$, $d = 1.4$. Therefore, instructional coaching widened the instructional ga

Standards		Cycle 1		Cycle 2		Cycle 3		Cycle 4		Cycle 5		Cycle 6	
		S	NS	S	NS	S	NS	S	NS	S	NS	S	NS
Joint Productive Activity	*M*	1.40	1.18	1.80	2.24	1.60	2.35	1.60	2.24	1.40	2.41	1.00	2.41
	SD	0.89	0.81	0.45	1.39	0.89	1.46	0.55	1.72	0.89	1.54	1.00	1.52
Language/ Literacy Dev.	*M*	1.00	1.53	1.60	2.53	1.80	2.65	1.80	2.59	1.40	2.59	1.60	2.94
	SD	0.00	0.51	0.55	1.18	0.45	1.22	0.45	1.37	0.55	1.28	0.55	1.03
Contextualization	*M*	1.00	1.29	1.20	2.24	1.00	2.35	1.00	2.24	1.20	2.29	1.40	2.53
	SD	0.00	0.69	0.45	1.35	0.71	1.37	0.00	1.39	0.45	1.45	0.55	1.33
Challenging Activities	*M*	1.40	1.29	1.80	1.82	1.40	2.41	1.80	2.41	1.20	2.06	1.80	2.76
	SD	0.55	0.59	0.45	0.88	0.55	1.28	0.45	1.33	0.45	1.09	0.45	1.25
Instructional Conversation	*M*	0.60	0.82	0.80	1.18	0.80	1.82	1.00	1.76	0.80	1.65	0.80	1.94
	SD	0.55	0.39	0.45	1.02	0.45	1.38	0.00	1.44	0.45	1.37	0.45	1.48
Total Score	*M*	5.40	6.12	7.20	10.00	6.60	11.65	7.20	11.24	6.00	11.00	6.60	12.59
	SD	1.52	2.18	1.64	5.07	2.61	6.28	0.84	6.77	2.00	6.21	2.61	6.15

Note: N=22: STEM = S (*n* = 5); Non-STEM = NS (*n* =17);

Table 2: Means and Standard Deviations for Five Standards and Total Score by Cycle and STEM and Non-STEM Groups

between STEM and non-STEM teachers, with STEM teachers' growth remaining relatively stagnant in comparison. STEM educators seem less able to innovate in their practices in ways that benefit ELLs.

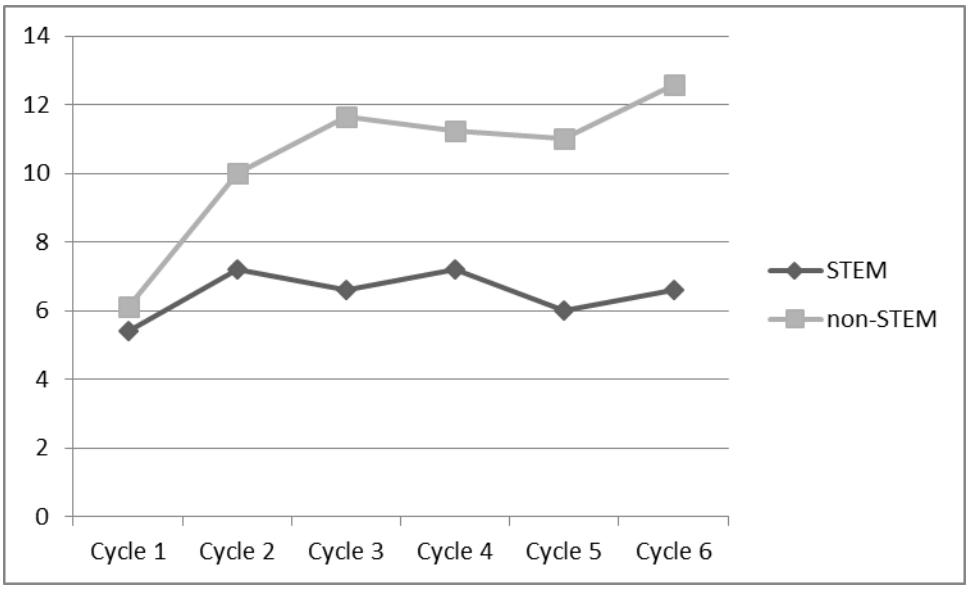

Figure 3. Marked line graph of total score by STEM and non-STEM groups

Discussion and Implications

This section discusses the efficacy of Five Standards instructional coaching with secondary teachers, in particular for STEM and non-STEM urban educators. Implications for theory, research, and professional development practices for secondary urban teachers are provided.

1. Transformation. Similar to elementary studies (Author, 2013; Author et al., 2011), these secondary urban teachers demonstrate statistically significant growth in their use of the Five Standards instructional model from the first to sixth cycle of instructional coaching, with medium effect sizes. However, they collectively grew only one in four possible levels of fidelity from the emerging to developing level on the SPC (Figure 1). When STEM and non-STEM teacher growth is compared, the difference is significant with a large effect size: STEM teachers

show no significant growth in use of the sociocultural practices valued in promoting language learning over six coaching cycles. Non-STEM teachers appear more able to transfer workshop learning to the classroom with the assistance of an instructional coach (e.g., Cornett & Knight, 2009; Neuman & Wright, 2010; Teemant et al., 2011).

In contrast, STEM teachers remain largely in teacher-dominated whole class configurations, with very brief opportunities for student talk, collaboration, or assistance of any type. In a three-year study, Johnson (2010) observed similar patterns of science teachers "slowly" becoming "more comfortable with allowing students to become more active learners" (p. 244).

In terms of practice, this study provides evidence that six cycles of instructional coaching across a single school year are not enough to support secondary urban teachers in reaching the highest levels of fidelity to sociocultural instructional practices. Extending the length of coaching as well as the frequency of coaching seems appropriate given limited teacher outcomes (Johnson, 2010). As Desimone (2009) and Neuman & Wright (2010) argue, more research is needed to establish the amount of coaching needed for sustainable instructional innovation, especially for STEM teachers.

Theoretically, two implications from the STEM/non-STEM differences stand out. First, while instructional coaching has broadly been defined by alternatively focusing on issues of management, content, instruction, or assessment (Knight, 2009), it is possible that more research will reveal a prioritizing order to coaching issues based on discipline-specific needs. The starting points for learning for STEM and non-STEM teachers are significantly different. Non-STEM teachers are ready to experiment with small group configurations while classroom management limits instructional innovation for STEM teachers. A study by Brown et al., (2007) also found English teachers more able than math teachers to implement new instructional

practices as a result of coaching. If, to what degree, and how instructional innovation is shaped by discipline-specific traditions and needs ought to be investigated further.

Second, instructional coaches must mediate such perceived discipline-specific constructions of student and/or teacher identities, agency, and use of power in order to elicit instructional innovation. As previous studies demonstrate (Author et al., 2011; Author, 2013; Author et al., 2013; Author et al., 2012), urban elementary teachers are able to fully enact sociocultural practices while developing greater autonomy, flexibility, and cognitive complexity around instruction. More research is needed to establish how the secondary culture of (a) the discipline, (b) the urban setting, and (c) the larger school context shape, intersect, or constrain teachers' use of sociocultural instructional practices.

2. Pattern of Development. The quantitative findings clearly demonstrate which of the Five Standards are more or less difficult to enact for secondary teachers. First, Language/Literacy Development is the easiest to implement. They improve from teacher-dominated talk to brief (STEM) and more sustained (non-STEM, at least 10 minutes) opportunities for student reading, writing, or speaking. Non-STEM teachers are more likely to (a) design activities that generate student talk and (b) to provide assistance through questioning, rephrasing, and modeling. Non-STEM teachers value the increased student talk and engagement resulting from their instructional changes.

Similar to other studies (Author et al. 2011; Author 2013; Yamauchi, Taum, & Wyatt, 2006), the Instructional Conversation is the standard teachers use the least. This is not entirely surprising. The Instructional Conversation requires teachers to be a full participant with a single small group. Although the Instructional Conversation is most commonly considered a literacy strategy, there is also evidence that the Instructional Conversation is being used in the context of

science (Stoddart, Solis, Tolbert, & Bravo, 2010) and mathematics to significantly increase achievement (e.g., Dalton & Sison, 1995; Fisher & Kopenski, 2008; Hilberg, Tharp, DeGeest, 2000). Johnson (2010) argues effective science instruction rests on "discussing observations, solving problems, and communicating findings" (p. 236). Therefore, new content-area demands for building students' abilities to talk mathematics (CCSSI, 2012) and talk science (NGSS Lead States, 2013) elevate the importance of developing STEM teachers' skills in crafting successful small group discussions, even Instructional Conversations. Horn (2012, p. 4) observes, "The focus on student thinking requires a genuine curiosity," on the part of teachers "about young people and their ideas" and that should "include opportunities for student sense making or questions that move understanding forward" and emphasize "deep conceptual understanding" (p. 5). Despite such calls for more dialogic interactions between teachers and students, a very real gap exists between idealized and realized STEM instruction in urban settings.

For the standards of Joint Productive Activity, Contextualization, and Challenging Activities, STEM teachers struggle to move beyond independent seatwork, abstract presentations, or basic recall and repetition of content. Horn (2012) reminds us that "According to the Third International Mathematics and Science Study (TIMSS), the dominant pattern of classroom instruction in the United States is learning terms and practicing procedures" (p. 4). The findings for non-STEM teachers, however, indicate greater willingness to ask students to collaborate on cognitively challenging and contextualized tasks. From a sociocultural perspective, neither group of secondary teachers provides assistance and feedback to students in the process of learning.

Theoretically, this high variability among STEM and non-STEM teachers in use of Joint Productive Activity, Contextualization, and Challenging Activities suggests teachers did not

have adequate support in their zone of proximal development by the instructional coach. Research using recorded coaching conversations could shed further light on the true nature of challenges encountered in teacher learning related to the Five Standards.

In practical terms, this high variability among STEM and non-STEM teachers also suggests these groups may benefit from being part of separate workshop events, allowing more differentiated and discipline-specific video and case study examples during the 30-hour workshop to support teacher implementation. STEM teachers would also benefit from more intense focus on classroom management skills. It is also possible that the phase-in process needs to be developed collaboratively, tailored, and made more manageable for STEM educators.

Findings indicate that a workshop plus one year of instructional coaching is not enough to close the gap between (a) STEM and non-STEM teachers outcomes nor (b) secondary and elementary teacher outcomes (e.g., Teemant, 2014; Teemant et al., 2011). Future research, therefore, should continue to investigate what combination and duration of assisted learning events lead to greater use of sociocultural instructional practices among secondary teachers, especially STEM educators.

Conclusion

The Five Standards instructional model represents sociocultural theory as observable pedagogical practices that promote English language learning. Nevertheless, the use of the Five Standards with urban secondary teachers as a way to prepare all teachers for ELLs is a relatively new avenue of professional development research. The findings of this study demonstrate secondary teachers' pedagogically practices became more sociocultural as a result of instructional coaching. However, this study also contributes new evidence that STEM and non-

STEM teachers significantly differ in their responsiveness to instructional coaching, with discipline-specific traditions and needs shaping coaching outcomes.

ABOUT THE AUTHORS

Annela Teemant is an Associate Professor of Second Language Education (Ph.D., Ohio State University, 1997) at IUPUI. She has been awarded four U.S. Department of Education ESL teacher quality grants to study the preparation of teachers working with English Learners.

Inquiries should be directed to ateemant@iupui.edu

Yuhao Cen is an Assistant Professor (Ph.D., Indiana University Bloomington, 2012) at the Graduate School of Education, Shanghai Jiao Tong University. Her research interests include teaching and learning, student development, and college experiences

Inquiries should be directed to ycen@sjtu.edu.cu

Amy Wilson (M.S. ESL, 2013) is a research assistant and adjunct instructor for the IUPUI English as a Second Language Program.

Inquiries should be directed to amymwils@umail.iu.edu

REFERENCES

Arnau, L., Kahrs, J., & Kruskamp, B. (2004). Peer Coaching: Veteran High School Teachers Take the Lead on Learning. *NASSP Bulletin, 88*(639), 26-41.

Aud, S., Wilkinson-Flicker, S., Kristapovich, P., Rathbun, A., Wang, X., and Zhang, J. (2013). The Condition of Education 2013 (NCES 2013-037). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved from http://nces.ed.gov/pubsearch

Brown, D., Reumann-Moore, R., Hugh, R., Christman, J. B., & Riffer, M. (2009).

Links to learning and sustainability: Year three report of the Pennsylvania High School Coaching Initiative. Philadelphia: Research for Action. Retrieved from http://www.researchforaction.org/publication-listing/?id=20

Brown, D., Reumann-Moore, R., Hugh, R., Christman, J. B., Riffer, M., du Plessis, P., & Maluk, H. P. (2007). Making a difference: Year two report of the Pennsylvania High School Coaching Initiative. Philadelphia: Research for Action.

Carrera, H. C. (2010). *Enhancing Teacher Practice through Coaching: A Case Study in an English Language Learner Environment.* (Unpublished doctoral dissertation). Temple University, Philadelphia, Pennsylvania.

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

Common Core State Standards Initiative. (2012). *Common Core State Standards for Mathematics.* Retrieved from http://www.corestandards.org/Math

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

Crisp, G., & Nora, A. (2012). Overview of Hispanics in science, mathematics, engineering, and technology (STEM): K-16 representation, preparation, and participation. Retrieved from http://www.hacu.net/images/hacu/OPAI/H3ERC/2012_papers/Crisp%20nora%20-hispanics%20in%20stem%20-%20updated%202012.pdf

Dalton, S., & Sison, J. (1995). *Enacting instructional conversation with Spanish-*

speaking students in middle school mathematics. Research Report 12. Santa Cruz, CA: National Center for Research on Cultural Diversity and Second Language Learning.

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.

Fisher, D., & Kopenski, D. (2007). Using item analyses and instructional conversations to improve mathematics achievement. *Teaching Children Mathematics 14*(5), 278-282.

Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press.

Hearn, R. M. (2010). *An Evaluation of Instructional Coaching at Selected High*

Schools in North Louisiana and Its Effect on Student Achievement, Organizational Climate, and Teacher Efficacy. (Unpublished doctoral dissertation). Louisiana Tech University, Louisiana

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

Hilberg, R.S., Doherty, R.W., & Reveles, C. (2004). Training manual and resources for five standards coaching. Unpublished manuscript. University of California, Santa Cruz.

Hilberg, R. S., Tharp, R. G., & DeGeest, L. (2000). Efficacy of CREDE's standards-based instruction in American Indian mathematics classes. *Equity and Excellence in Education, 33*(2), 32-40.

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.

Horn, I.S. (2012). *Strength in numbers: Collaborative learning in secondary mathematics*. Reston, VA: The National Council of Teachers of Mathematics, Inc.

Johnson, C.C. (2010). Transformative professional development for in-service teachers: Enabling change in science teaching to meet the needs of Hispanic English language learner students. In D. W. Sunal, C.S., Sunal, & E. L. Wright (Eds.), *Teaching science with Hispanic ELLs in K-16 classrooms* (pp. 233-252). Charlotte, NC: Information Age Publishing, INC.

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

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

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

Lee, O., Quinn, H., & Valdés, G. (2013). Science and language learning for English language learners in relation to Next Generation Science Standards and with implications for Common Core State Standards for English language arts and mathematics. *Educational Researcher, 42*(4), 223-233.

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

Marsh, J. A., Sloan McCombs, J., & Martorell, F. (2010). How Instructional Coaches Support Data-Driven Decision Making. *Educational Policy, 24*(6), 872-907.

Murray, S., Ma, X., & Mazur, J. (2009). Effects of Peer Coaching on Teachers' Collaborative Interactions and Students' Mathematics Achievement. *Journal of Educational Research, 102*(3), 203-212.

National Center for Education Statistics. (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.

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington DC: The National Academies Press.

Neuman, S.B., & Wright, T.S. (2010). Promoting language and literacy

development for early childhood educators: A mixed-methods study of coursework and coaching. *The Elementary School Journal, 111*(1), pp. 63-86.

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

Sailors, M. & Shanklin, N.L. (2010). Introduction: Growing Evidence to Support Coaching in Literacy and Mathematics. *The 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*(2), 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: Routledge, Taylor & Francis Group and the Association of Teacher Education.

Stoddart, T. Solis, J., Tolbert, S., & Bravo, M. (2010). A framework for the effective science teaching of English language learners in elementary schools. In D. W. Sunal, C.S., Sunal, & E. L. Wright (Eds.), *Teaching science with Hispanic ELLs in K-16 classrooms* (pp. 151-181). Charlotte, NC: Information Age Publishing, INC.

Teemant, A. (2014). A mixed methods investigation of instructional coaching for teachers of diverse learners. *Urban Education*, *49*(5), 574-604. DOI:10.1177/0042085913481362

Teemant, A., & Hausman, C. S. (2013, April 15). The relationship of teacher use of critical sociocultural

practices with student achievement. *Critical Education, 4*(4). Retrieved

Teemant, A., & Reveles, C. (2012). Mainstream ESL instructional coaching: A repeated measures replication study. *INTESOL Journal, 9*(1), 17-34.

Teemant, A. Wink, J., & Tyra, S. (2011). Effects of coaching on teacher use of sociocultural instructional practices. *Teaching and Teacher Education, 27*(4), *683-693*.

Tharp, R. G. (1982). The effective instruction of comprehension: Results and description of the Kamehameha Early Education Program. *Reading Research Quarterly, 17*(4), 503-527.

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

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

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

Vogt, F., & Rogalla, M. (2009). Developing Adaptive Teaching Competency through Coaching. *Teaching and Teacher Education: An International Journal of Research and Studies, 25*(8), 1051-1060.

Walsh, W.B. & Betz, N.E. (1990). *Tests and Assessment* (2nd ed.). Englewood

Cliffs, NJ: Prentice Hall.

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

Yamauchi, L. A., Taum, A. K. H., Wyatt, T. R. (2006). Teachers' use and understanding of instructional conversations in four high school classrooms. Retrieved from http://www.bridgingworlds.org/modelsandmetrics/Yamauchi_etal_Teachers.pdf.