



Professional Development for Reading Achievement: Results from the Collaborative Language and Literacy Instruction Project (CLLIP)

Citation

Porche, Michelle V., Daniel H. Pallante, and Catherine E. Snow. 2012. "Professional Development for Reading Achievement." *The Elementary School Journal* 112 (4) (June): 649–671.
doi:10.1086/665008.

Published Version

10.1086/665008

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:34785389>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

PROFESSIONAL DEVELOPMENT FOR READING ACHIEVEMENT

Results from the Collaborative Language and Literacy Instruction Project (CLLIP)

ABSTRACT

The Collaborative Language and Literacy Instruction Project (CLLIP) is a model of professional development designed to help teachers incorporate research-based practices of literacy instruction, support mastery, and sustained use of these practices through coaching, and serve as a foundation for whole-school reform efforts. We describe the model, intervention, implementation, and subsequent results from an exploratory study in which we tested student literacy outcomes for kindergartners and fourth graders in the classrooms of CLLIP teachers against a matched comparison group. Exploratory results from a rural cohort of elementary school teachers suggest support for skill building in the alphabetic principle, phonemic awareness, fluency, and vocabulary. We discuss outcomes by reflecting on central program features: CLLIP strengthens teachers' content knowledge and ties that knowledge to subject-specific content for students, has extended duration and support, is tied to state standards, and involves collective participation across a district to advance reform efforts.

Michelle V. Porche

WELLESLEY CENTERS FOR
WOMEN, WELLESLEY
COLLEGE

Daniel H. Pallante

OHIO EDUCATIONAL
DEVELOPMENT CENTER

Catherine E. Snow

HARVARD UNIVERSITY

THE reading field has achieved a high degree of consensus on the features of good primary-grades reading instruction: attention to oral language skills, phonemic awareness, the alphabetic principle, fluency, and procedures for supporting comprehension (National Early Literacy Panel, 2008; National Reading Panel, 2000; Snow, Burns, & Griffin, 1998). In practice, this translates into a

judicious balance of skills- and meaning-focus in instruction, rigorous task demands including higher-level questioning with substantial scaffolding, differentiated small-group instruction, and strong connections of literacy to curricular content, in the context of excellent classroom management (Taylor, Peterson, Pearson, & Rodriguez, 2002). Implementation of these good practices, though, is the recurrent challenge. We know that how teachers practice instruction is as important as the content they cover, and that ongoing professional development and support are crucial to ensuring that all teachers know how to implement excellent literacy instruction (Snow, Griffin, Burns, and the NAE Subcommittee on Teaching Reading, 2005). The National Staff Development Council's (2001) standards for staff development identify the organization of learning communities, district leaders who support continuous instructional improvement, and resources for teachers in the form of training and materials as key.

The challenge is to devise professional development that helps teachers incorporate effective elements into their teaching, supports their continued use of them, and then percolates the elements throughout a school and ultimately a district. The characteristics of effective professional development are well documented and promoted both within public policy and educational research. Holland argued for a focus on professional development initiatives that have direct effects on student achievement, which he identifies as those focused on "(1) how students learn particular subject matter; (2) instructional practices that are specifically related to the subject matter and how students understand it; and (3) strengthening teachers' knowledge of specific subject-matter content" (2005, p. 4). Recently, there has been a movement toward professional learning communities (DuFour, 2007; Eaker & Keating, 2008), which focus on teachers' learning as a collective and on making a group commitment to change in instructional practices that result in measurable student gains.

To date, the impact of well-designed professional development on student reading in the context of whole-school reform through collaborative learning communities is understudied (Vescio, Ross, & Adams, 2008). The Collaborative Language and Literacy Instruction Project (CLLIP) incorporates the core principles of professional learning communities: shared understanding of student learning goals, criteria for formative assessment to be used on a timely basis to identify individual student strengths and weaknesses, use of data to assess instructional effectiveness and to ensure continuous improvement, and inclusion of scientifically based knowledge in decision making (DuFour, 2007). This article describes the development of the CLLIP program and offers preliminary evidence suggesting its relationship to student learning.

Professional Development and School Reform

National policy has had increased focus on scientifically based research as a foundational criterion for effective professional development in order to ensure qualified teachers and subsequent student achievement. Thus, professional development should include expansion of theoretical understanding of literacy instruction designed for authentic improvement in practice. Yet the scientific aspects of reading research are too often ignored as teachers pick and choose strategies for classroom use (Cunningham, Zibulsky, Stanovich, & Stanovich, 2009). Commeyras and DeGroff (1998) found that teachers were more likely to read practitioner journals and

magazines than rigorous reading-research journals, and 60% of the elementary school educators in their study reported “never” reading research journals. Given the limited personal reading engagement reported by preservice literacy teachers (Nathanson, Pruslow, & Levitt, 2008), we might not expect much exposure to research publications. Nor, indeed, would reading an occasional article from even the most rigorous journals constitute a strong basis for revising one’s practice; the value of syntheses such as *Preventing Reading Difficulties in Young Children* (Snow et al., 1998) and the *National Reading Panel Report* (National Reading Panel, 2000) is precisely that they aggregate information across a wide variety of studies. Lacking the skills, interest, or time to integrate evidence-based knowledge into their teaching, teachers are unlikely to make necessary paradigm shifts in practice.

Programs that advance school improvement within the context of comprehensive school reform (Sterbinsky, Ross, & Redfield, 2006) can successfully promote effective implementation of improved or newly trained instructional practices in literacy. Correnti and Rowan’s investigation of three widely disseminated comprehensive school reform programs found that the key to changing teachers’ instructional practice rested on “delimited curricular areas, built around clear and highly specified designs for instructional practice, and backed by leaders who work assiduously in local settings to promote implementation fidelity” (2007, p. 328). The focus of instructional change itself might vary, but these components must be in place for successful change. Goldenberg pointed to the importance of explicit curriculum and instruction, strong leadership, collaboration, and shared goals; he asserted that change can only occur when “new settings are created or existing settings are changed to reflect new goals and activities aimed at attaining those goals” (2003, p. 11).

The CLLIP professional development model instantiates key conditions for teacher development described above in a number of ways. This intervention was designed by the second author to be a school reform model with a goal of having all teachers in an elementary school using a coordinated curriculum for literacy instruction. This is in contrast to disconnected one-time opportunities for workshop-based professional development that tend to be particularly ineffective in promoting student achievement (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). Training in literacy instruction practices integrates specific teaching strategies with the foundational scientific theory on which strategies are based. Rich opportunities for language development are at the core of CLLIP instructional practices (Adger, Snow, & Christian, 2002). Specific emphasis on oral language, units of language, academic language, and written language are interwoven in teacher training at all grade levels (Valdés, Bunch, Snow, Lee, & Matos, 2005). The scope and sequence of the training modules and coaching is extensive to ensure progress toward mastery in literacy instruction. Moreover, instructional practices are explicitly tied to the state’s Academic Content Standards in English Language Arts (Office of Curriculum and Instruction at the Ohio Department of Education, 2001). All of this is accomplished in what Tharp and Gallimore (1988) referred to as collaborative activity settings in which teacher performance is assisted by coaches and administrators rather than simply assessed.

The initial phases of development included a demonstration project carried out in 1999–2001, followed by several years of refinement. Participating districts were required to demonstrate student need as reflected by weak student proficiency scores and district report card ratings, as well as district and community financial need (as

is the case in the exploratory study described below). Preliminary results from pilot schools during this refinement period supported the feasibility of implementing the professional development program across an entire district with more rigorous assessment using control group comparisons. One premise was that a unified program for all elementary teachers would streamline training, thus offering them greater opportunity for support and access to greater depth of instructional knowledge (Darling-Hammond et al., 2003). There are also intangible benefits related to team building within schools and for schools within districts, as teachers and administrators are united around communally defined instructional processes and goals. A second premise was that a regular feedback loop, based on student assessment, would foster greater instructional self-efficacy, which in turn is linked to higher expectations for student outcomes (Boudett, City, & Murnane, 2005). A third premise is that on-site coaching and ongoing support are necessary to ensure that implementation of newly learned instructional strategies occurs with fidelity (Knight, 2009).

The primary purpose of this article is to describe the CLLIP professional development program and its implementation in a district with multiple elementary schools where comparisons could be made between intervention and control classrooms in the same schools. Previous pilot analyses were limited to data from urban participants compared to controls in a neighboring district. For this exploratory study, we compared the performance of students in CLLIP teachers' classrooms on literacy measures with those of control teachers in the same schools in order to evaluate the potential of the model to improve student outcomes. An additional goal of this study was the development of a coaching checklist for future use to guide on-site training and support related to fidelity of implementation and sustainability of the program initiatives. CLLIP is designed for district-wide implementation across all elementary school grades. However, because of limited resources for intervention in this exploratory phase, priority was given to investigate implementation in kindergarten to strengthen emerging literacy skills and in grade 4 to boost established skills in advance of state proficiency testing. The exploratory analyses presented here tested associations with reading skills for students over one school year to address the following research questions: Do CLLIP students in the emerging (kindergarten) and established (grade 4) reading levels make greater gains compared to students in control classrooms on language and literacy outcome measures specifically addressed in CLLIP teachers' professional development training? Are gains in language and literacy skills in CLLIP classrooms moderated by student risk status, given the formal small-group instruction strategies integrated into CLLIP teachers' professional development training?

Method

Participants

District description and teachers. This study was conducted in a rural Appalachian district in an economically disadvantaged community. The ultimate goal of the intervention is a district-wide scale-up over several years; this article describes the first year of participation of its five elementary schools. In the second year, as a part of the scale-up design, district capacity was increased by providing the control teachers with formal CLLIP training; thus comparative longitudinal analysis is not possi-

ble and results are limited to this first year. A total of 27 classroom teachers across grade levels, four Title I teachers, and seven administrators attended CLLIP professional development trainings. The district agreed to implement CLLIP intervention strategies exclusively during participation and to provide financial support for the intervention in the form of release time for teacher training and payment for substitutes. Strong administrator commitment by the superintendent and principals allowed us to compare treatment and control classrooms within each of the five buildings. However, this comparison sampling also allowed for the possibility of contamination if treatment teachers shared new knowledge with control teachers. Further, teachers in the intervention group were selected by administrators from among teachers willing to participate in professional development and coaching. We have few details about selection decisions except for an acknowledgment that teachers deemed more “cooperative” were more likely to be chosen. There were no indications that students were assigned to teachers in any systematic fashion. Control group teachers were chosen from among non-CLLIP participants working in the same buildings by drawing names from a hat. The fact that the intervention teachers were not randomly selected is a potential confound, as they might have been more motivated to change instructional practices.

Prior to implementation, to assist in the development of the intervention and to assess teachers’ training needs, participants were asked to provide information about their educational background (degrees, years of experience) as well as previous experience with professional development, including the type and frequency of trainings they had attended over the previous 5 years. At the end of each training session, participants were asked to provide satisfaction ratings and brief comments about their learning experience. At the end of the year, they were asked to complete satisfaction surveys to evaluate the quality of the literacy instruction programs they were engaged in (CLLIP or instruction as usual) during the school year.

CLLIP and control teachers were equivalent in average years of experience (17 years), while the ratio of CLLIP teachers with master’s degrees (67%) was higher than controls (22%). Overall, CLLIP and control teachers described similar professional development experiences, most reporting single-session workshops as their primary experience; few had received training across all components of literacy instruction, and many indicated no professional development training. One CLLIP teacher reported using her bachelor’s coursework in classroom practice, and another reported the use of master’s-level training in specific aspects of literacy instruction. One teacher (in the CLLIP group) reported reading journal articles (through subscription to *Reading Teacher*) as a way to enhance practice. Few teachers noted having previous exposure to assessment: one CLLIP teacher named Reading Recovery observations and another reported use of portfolios, while one control teacher reported use of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002).

Students. Across the five school buildings in the district, 122 kindergarten students (five intervention and two control classrooms) and 138 grade 4 students (four intervention and two control classrooms) completed the assessment batteries. The district did not allow us to collect individual-level student demographic information (e.g., race and ethnicity, socioeconomic status). Thus only summary student descriptions at the district level were available: all identified as White native English speakers, 27% were economically disadvantaged, 36% were on free or reduced-price lunch, and 15% had disabilities. Students completed pre- and posttest assessments over a

single school year. All human-subjects protections for students and teachers were adhered to, as stipulated by the Institutional Review Board for Wellesley College.

Identification of at-risk students for small-group differentiated instruction.

Classroom teachers provided students with differentiated instruction in the areas of reading, writing, and oral language competence. A pretest standardized assessment battery consisting of four tests per grade level (described in detail below) was used in order to determine which students were at risk for reading failure and who would receive a small-group language-based literacy intervention. Fall results provided indicators of risk status for each student on each individual assessment based on test developers' criteria for performance standards. Scoring at-risk status in one or more assessments qualified students for placement in small-group instruction. Groups comprised seven at-risk students per classroom.¹

Student Outcome Measures

At the urging of state educational agencies, two sets of standardized measures were selected. Coinciding with this study, DIBELS was beginning to be used by districts within the state in wide-scale attempts to increase accountability. Reports of ease of use and teacher enthusiasm for DIBELS, along with evidence of validity and reliability (Good & Kaminski, 2002), supported the selection of this assessment. The Woodcock-Johnson Tests of Achievement III (WJIII; Woodcock, McGrew, & Mather, 2001) was chosen because of its wide use in research studies, although there have been several critiques about its inclusion of a cloze item rather than a more authentic approach to measuring reading comprehension (Francis, Fletcher, Catts, & Tomblin, 2005). Teachers were fully trained on the two assessment batteries and tested their own students under the monitoring of CLLIP staff; control students were tested by CLLIP staff.

DIBELS (Good & Kaminski, 2002) is a set of standardized measures to assess growth in phonological awareness, the alphabetic principle, and fluency for students from kindergarten through third grade. These are designed to be easy-to-administer, 1–3-minute assessments that teachers can use for immediate feedback on student progress. Depending on the specific test, pre- and posttest administration occurred in September, January, or May. The DIBELS has established benchmark scoring levels at each testing administration to indicate risk status for each measure (i.e., scoring in the lowest 20th percentile):

- Letter Naming Fluency measures the ability of students to identify as many upper- and lowercase letters as possible presented in 1 minute (administered September and May).
- Initial Sounds Fluency measures the ability of students to identify and produce the initial sound of an orally presented word in 3 minutes (administered September and January).
- Phoneme Segmentation Fluency measures the ability of students to correctly segment three- and four-phoneme words into their individual phonemes in 1 minute (administered January and May).

WJIII (Woodcock et al., 2001) is a standardized norm-referenced test battery, with a mean of 100 and a standard deviation of 15. Four subtests of WJIII were used,

administered in September and May. Grade 4 students completed the three broad reading subtests (Letter-Word Identification, Passage Comprehension, and Fluency). In addition, both grade 4 and kindergarten students completed the vocabulary subtest. Raw scores for each test correspond to a grade equivalency rating that indicates whether the student is above average, average, limited, or very limited. Raw scores indicating limited or below enabled us to identify at-risk status for students at pre- and posttest assessment.

- Letter-Word Identification requires students to identify letters and subsequently read words of increasing difficulty.
- The Passage Comprehension measure requires students to read short passages and fill in missing key words.
- Fluency measures the ability of students to read and understand simple sentences within a 3-minute time limit.
- The Picture Vocabulary test starts as a measure of receptive vocabulary; the early items require children to point to the correct pictures after hearing a word. Later items require the child to name pictured objects or events.

CLLIP Intervention Procedures

Overview and staffing. The CLLIP professional development intervention is designed to foster prevention of reading difficulties rather than relying on remedial strategies at later stages in students' development. Teachers were not only trained in instructional approaches informed by scientifically based reading research but also learned how to assess students' literacy skills and use those assessments for diagnosis and prescription of intervention strategies to build strengths. Training in small-group, differentiated instruction was a critical component of the professional development. Administrators were strongly encouraged to attend all trainings so that they could keep pace with teaching staff, assist coaches in monitoring fidelity, and maintain sustainability of CLLIP practices. On-site coaching throughout the year was another major component meant to ensure fidelity of implementation in whole-classroom and small-group settings. Consistent with key coaching qualifications identified in the literature (McCombs & Marsh, 2009), CLLIP staff possessed a high level of content knowledge and teaching skills in the classroom, as well as ability to teach and mentor adult learners. The CLLIP director provided training and supervision to the assistant director (the director and assistant director are licensed language speech pathologists) and a full-time coach (hired for expertise in literacy instruction and extensive experience in elementary school classrooms). Altogether, these three CLLIP staff, hereafter identified as CLLIP coaches, led the training modules and provided the on-site coaching described below. In addition, a full-time, on-site literacy coordinator was provided by CLLIP (trained and supervised by the CLLIP director), as support staff were housed in the participating district in order to provide assistance with the day-to-day management of the intervention and support for teachers.

Professional development curriculum. Over the course of the first year, teachers and administrators attended a comprehensive series of six different professional development modules that included extensive training materials (professional books, research articles, assessment materials, classroom manipulatives; see App.

Table A1) and research-based training in effective literacy instruction in this sequence:

- Strategic Assessment: training in the DIBELS and WJIII assessments, interpretation of results, and identification of areas needing additional instruction
- Targeted Instruction for Small-Group Intervention: training in use of formal intervention practices (described below) as well as guidance in developing strategies to identify and address needs of students with similar concerns through informal small-group methods
- Word Reading: overview of underlying principles and research surrounding phonological and phonemic awareness, the alphabetic principle, and phonics instruction
- Fluency: introduction to the specific components of fluency, including phrasing, expression/volume, smoothness, and pace, as well as specific modeling strategies to build students' decoding and fluency skills
- Vocabulary: broader understanding of research related to vocabulary and its relation to the development of reading skill and comprehension, instructional practices to facilitate gains in student vocabulary knowledge and skills using roots, deep processing, inference of meaning from text, and rich instruction
- Writing: overview of writing research, informal assessment and evaluation, understanding of writing processes, traits, and standards-based skills
- Comprehension: While phonemic awareness, alphabetic principle, and guided oral reading are presented individually, reading comprehension is integrated throughout the other modules. This approach enables CLLIP instructors to help teachers make language-based connections between comprehension and specific literacy content gradually and thoroughly during the school year, through explicit instruction and modeling of comprehension strategies (Duke & Martin, 2008) across modules.

Teachers learned during the word-reading training module, for example, how weak alphabetic principle and phonemic awareness skills can have a negative influence on decoding skills, which in turn affect how well students understand some words they are trying to read (Stahl, 2003). Similarly, during fluency training, teachers learned that students with weak decoding skills are often adversely affected by slow rates of word reading (e.g., fewer words read correctly per minute), a measurable characteristic of reading fluency believed to be linked to a student's level of reading comprehension (Rasinski, 2003). As part of vocabulary training, participants observed CLLIP instructors modeling comprehension strategies during read-alouds in order to discuss and rehearse comprehension strategies such as self-monitoring, summarizing, visualizing, and making personal connections to text. In this way, CLLIP instructors were able to train teachers in comprehension gradually over time, reinforcing how comprehension is measured and taught as it relates to other literacy skills. Content standards and guidelines for differentiated instruction were delineated for sessions in which teachers were separated by grade level groupings (K–2, grades 3–6). Each training module encompassed a full day of instruction, taking place during designated professional development time and during the regular school day, with substitutes filling in for teachers in the program.

All trainings began with an overview of general goals of the particular module as well as specific expectations for learning for each section of the module, followed by theoretical and practical instruction, including the development of lesson plans. For instance, the first half of the Word Reading module included a didactic approach to presenting theoretical background in language and phonology, expectations for instruction of phonological awareness within state content standards, information on assessment and diagnoses of phonemic awareness and decoding difficulties, and research-based approaches to instruction. For the second half of the day, teachers participated in more experiential hands-on activities where specific instructional strategies were modeled and teachers were given time to practice using these strategies with each other. As part of these practice activities, teachers worked together to develop lesson plans (all teachers shared their lesson plan with all other teachers working in the same grade level) so that when they returned to the classroom they were ready to implement what they had learned. Developing lesson plans also included creating manipulatives that were used in the classroom, for example, word sorts for the Word Reading module.

Teacher training attendance and satisfaction. Attendance was recorded for each training module; the range for teachers was 82%–100%, whereas administrators had a more varied range of attendance (see Table 1). At the end of each professional development module, CLLIP participants rated their satisfaction with various aspects of training content, organization, and value for classroom practice on a 4-point scale (1 = very poor, to 4 = very good). On average, teachers and administrators gave a rating of good ($M = 3.6$ across the six modules; see Table 1). Although teachers were also invited to comment on each of the trainings, most did not. The few responses related to the first training experience ranged from enthusiasm (“the training was professional and informative—I feel ready to jump in”) to trepidation, as participants noted they were “ill-at-ease,” “uncomfortable,” or “overwhelmed.” While response rates for comments on the subsequent trainings were similarly limited, comments tended to be positive, albeit urging greater attention to how-to rather than why: “Less of the theoretical aspect and more examples on how to use it in our classrooms.” In addition, brief comments on teacher satisfaction were collected at the end of the first year. CLLIP teachers were able to name specific features of reading instruction that were provided in the CLLIP trainings, for instance: “understanding the phonemic awareness piece of sound segmentation,” “the children have developed skills and strategies to attack words that are unfamiliar to them,” and “my class

Table 1. Attendance Records for Training Modules Listed in Order of Delivery with Timing Indicated

	Satisfaction Score <i>M (SD)</i>	Attendance Teachers %	Attendance Administrators %
Strategic assessment (August)	3.28 (.34)	100	100
Targeted instruction for small-group intervention (September)	3.69 (.26)	100	100
Phonological awareness: K–3 (October)	3.46 (.18)	92	57
Fluency (November)	3.90 (.07)	82	88
Vocabulary (December)	3.54 (.12)	88	100
Writing (February)	3.62 (.17)	88	56
Mean	3.60 (.17)	93	86

reading program has changed from whole class instruction to small groups with a much greater emphasis on vocabulary.” Control group teachers pointed to the need for more training (“I’d like to have more information about teaching vocabulary and fluency”) or were vague in noting areas of improvement (“all students benefit in many areas”).

Targeted instruction for small-group intervention. For teachers and administrators, the identification of at-risk students was an innovative, direct use of assessments as diagnostics “for learning,” in contrast to established practices of testing for district accountability “of learning” (Edwards, Turner, & Mokhtari, 2008). The professional development module for assessment introduced connections between scoring on standardized tests and interpretations of the scores as reflections of the presence or absence of specific literacy skills. The second module on small-group intervention provided direction in how to address identified areas of concern. These connections were reinforced at each subsequent module through the fall semester, where teachers would bring their students’ baseline test results to serve as discussion points to guide plans for appropriate instructional interventions. For most teachers, doing their own testing and scoring was a novel experience that led them to take a more analytical approach to the data on their students and helped them become more skilled as literacy diagnosticians.

CLLIP coaches and the literacy coordinator worked with teachers to begin developing systematic intervention strategies for their students. A core language-based intervention strategy used was the SOAR to Success program (Cooper, Boschken, McWilliams, & Pistochini, 1997) for students in grades 3–6. The SOAR program is designed to place children into groups of seven for guided reading sessions and to strengthen language-based strategies targeting syntactic, semantic, and orthographic reading difficulties. Students were immersed in language-based literacy instruction in which strategies modeled by teachers and other students included summarizing, clarifying, asking questions, predicting events in stories being read, and taking turns reading aloud. A similar program, Reading Intervention for EARLY SUCCESS (Houghton Mifflin, 2003), was used for grade 1 and 2 students. These programs were chosen because they align with the foundational interactive language strategies at the core of the CLLIP approach. Students in the SOAR small groups were provided leveled reading books and then taught how to clarify, summarize, predict, and question text they read from the set of books they used in the 18-week series. Using these strategies on a regular basis to discuss text with a teacher is similar to the explicit comprehension instruction embedded in the professional development provided to teachers as part of the CLLIP. In addition, these programs required teachers to work in small groups with students who were lacking specific literacy skills, directly aligning with the emphasis placed on differentiated instruction in the CLLIP. Teachers were taught from their first training that instruction needs to stem from assessment results that reflect the performance of their students. Thus, teachers learned how to begin thinking about which students need the most assistance, the specific skills they have or may be lacking, and then how to differentiate the instruction students require to meet their needs. These programs were used to model small-group instruction, so that all teachers would gain the same skills in using differentiated instruction as well classroom management when working

with small groups during the regular school day, rather than using pull-out programs.

Additional small-group strategies for kindergarten children focused on children's comprehension and production of academic discourse. These were developed by the CLLIP director, as suggested by the Committee on the Prevention of Reading Difficulties in Young Children (Snow et al., 1998). Oral language interventions concentrated on skills necessary for students to comprehend fiction and nonfiction information presented orally, the way in which much information is provided in school settings.

Coaching and support. CLLIP coaches provided teachers and administrators with regularly scheduled on-site coaching and supervision through scaffolded instructional opportunities. These were supplemented by teacher-initiated phone and e-mail check-ins. In addition, the on-site literacy coordinator met with teachers to support practices learned in professional development modules through weekly team meetings and individual coaching meetings several times a month. CLLIP coaches faced several challenges in successfully providing structured support. They had to build trusting relationships with teachers so that teachers felt comfortable being observed in their classrooms; at the same time, they had to foster the integration of newly acquired knowledge of scientifically based reading research into practice, all while building the foundation of a coordinated effort toward school reform (Walpole, McKenna, Uribe-Zarain, & Lamitina, 2010). To facilitate the learning of instructional practices, coaches also modeled lessons for teachers at the teacher's request if she or he felt unsure how to attempt a new instructional approach.

The many demands of the packed school calendar (Al Otaiba, Hosp, Smartt, & Dole, 2008) allowed for 8 and 10 formal coaching visits to each teacher during the school year; each visit lasted approximately 30 minutes. Coaches alternated between observation and conferencing visits. Observations consisted of limited interaction with teachers, but extensive targeted note taking. Over the course of this exploratory study, coaching checklists were developed to record the use of materials and methods introduced in trainings as well as demonstrated strengths and need for improvement. As part of this study, the checklists were developed and refined as they were used to guide conferencing visits, but they were not systematically tested as a data-observation tool. Within a week of each observation, the coach distributed a short memo to the teacher and principal outlining strengths and suggestions for improvement. During conferencing visits, CLLIP coaches and teachers engaged in discussions regarding the coaching memos in order to increase fidelity of implementation of instructional practices presented in trainings. Administrators were regularly encouraged to participate in classroom observations in order to increase knowledge and supervision of expected best practices in literacy instruction (Klingner, 2004), thus building sustainability of the program once formal CLLIP participation had ended.

Analysis

Descriptive statistics on student assessments were computed in order to show the range of achievement and risk status. Correlations between reading and vocabulary pretest measures were tested by grade level. To answer the first research

question, SAS PROC MIXED was used to fit two sets (kindergarten and grade 4) of multilevel classroom-effects models (Singer, 1998) to determine whether students in the CLLIP classrooms made greater gains in the literacy assessments compared to students in control classrooms. Student-level reading outcomes (DIBELS and WJIII assessments) were tested as a function of both level 1 (student) and level 2 (classroom) predictors. Level 1 predictors included risk status and individual student pretest score (grand mean centered) for each outcome. Gender was the only student-level demographic measure provided by the district and was included here because of differences found in earlier pilot work. Level 2 predictors included teacher education level, given the higher percentage of CLLIP teachers with graduate degrees; the mean classroom pretest score for each outcome was included as a baseline control variable. To answer the second research question, interaction terms were included for each set of models to assess whether gains in language and literacy skills in CLLIP classrooms were moderated by student risk status. Regression results are reported by grade level. The fixed-effects results describe inferences about the sample pool; the random-effects outcomes provide inferences about variation in the target population, for example, describing whether differences in outcomes reflect significant variation by classroom, by individual students within classrooms, or both.

Results

Descriptive Results

Correlation analyses of kindergarten students' pretest literacy measures indicated that the reading measures had strong (Phoneme Segmentation Fluency and Letter Naming Fluency: $r = .69, p < .001$) to moderate (Phoneme Segmentation Fluency and Initial Sound Fluency: $r = .46, p < .001$; Initial Sound and Letter Naming Fluency: $r = .48, p < .001$) correlations with each other. Vocabulary was moderately related to reading skills (Phoneme Segmentation: $r = .28, p < .01$; Initial Sound Fluency: $r = .31, p < .01$; Letter Naming Fluency: $r = .43, p < .001$). Analyses for grade 4 students indicated that the three reading measures were strongly related to each other (Word Reading and Comprehension: $r = .80, p < .001$; Word Reading and Fluency: $r = .78, p < .001$; Fluency and Comprehension: $r = .77, p < .001$) and moderately related to vocabulary (Word Reading: $r = .58, p < .001$; Comprehension: $r = .61, p < .001$; Fluency: $r = .46, p < .001$). As students made gains in literacy skills over the year, the percentage of CLLIP students who scored at or below the at-risk cutoff scores also decreased, more markedly for grade 4 than kindergarten students (see Table 2). Testing of intraclass correlations (ICCs) for the posttest outcomes suggested some significant variance explained at the classroom level and greater additional variance explained at the student level for three of the four kindergarten measures (Letter Naming: $ICC_{\text{student(class)}} = .44, ICC_{\text{class}} = .18$; Initial Sound: $ICC_{\text{student(class)}} = .18, ICC_{\text{class}} = .28$; Phoneme Segmentation: $ICC_{\text{student(class)}} = .35, ICC_{\text{class}} = .27$; Picture Vocabulary: $ICC_{\text{student(class)}} = .37, ICC_{\text{class}} = .04$). For all the grade 4 posttest outcomes, only the student-level variance was significant (Letter-Word Identification: $ICC_{\text{student(class)}} = .86, ICC_{\text{class}} = .14$; Passage Comprehension: $ICC_{\text{student(class)}} = .72, ICC_{\text{class}} = .08$; Fluency: $ICC_{\text{student(class)}} = .81, ICC_{\text{class}} = .08$; Picture Vocabulary: $ICC_{\text{student(class)}} = .66, ICC_{\text{class}} = .09$).

Table 2. Pretest and Posttest Scores on Standardized Literacy Measures Including Percent at Risk

	Pretest Mean Score (% at Risk)		Posttest Mean Score (% at Risk)	
	CLLIP	Control	CLLIP	Control
Kindergarten:				
Letter Naming Fluency	16.48 (18)	19.39 (15)	45.00 (18)	40.42 (26)
Initial Sound Fluency	12.69 (9)	13.38 (24)	24.04 (10)	19.30 (17)
Phoneme Segmentation Fluency	24.78 (8)	21.23 (19)	39.14 (6)	32.84 (3)
Picture Vocabulary	17.01 (27)	17.82 (24)	18.82 (14)	18.45 (23)
Grade 4:				
Letter-Word Identification	47.85 (51)	48.94 (43)	52.89 (28)	51.93 (42)
Passage Comprehension	27.26 (49)	26.72 (54)	30.14 (36)	29.27 (40)
Fluency	38.39 (48)	38.80 (46)	47.16 (20)	43.33 (38)
Picture Vocabulary	24.45 (38)	24.96 (21)	26.52 (24)	26.07 (29)

Kindergarten Student Outcomes

In answer to our first research question, we found no main effect of CLLIP for kindergarten students. The second research question tested whether the intervention was moderated by the level 1 predictor, risk status. At-risk students scored lower on all assessments compared to their low-risk counterparts; however, at-risk CLLIP students showed greater gains on two measures (see Table 3).

Table 3. Kindergarten Fixed Effects and Variance Estimates for Posttest Outcomes ($n = 122$ Students in 7 Classrooms)

	Letter Naming Parameter (SE)	Initial Sound Parameter (SE)	Phoneme Segmentation Parameter (SE)	Picture Vocabulary Parameter (SE)
Fixed effects:				
Intercept	27.81* (13.06)	22.94* (11.32)	44.39*** (5.85)	17.69*** (1.78)
Level 1:				
Student pretest score centered	.86 (.73)	.72 (1.56)	.63* (.31)	.04 (.07)
Gender (male = 1)	.56 (2.40)	-1.98 (1.72)	-3.97* (1.98)	1.15* (.47)
Risk status (1 = at risk)	-11.32* (4.96)	-10.72** (3.48)	-13.05** (4.00)	-2.29* (.95)
Level 2:				
Teacher education level	.43 (6.16)	2.01 (10.07)	-2.09 (4.58)	.60 (.69)
Mean classroom pretest score	.59*** (.11)	-.40 (1.57)	-.26 (0.33)	.03 (.02)
CLLIP	5.27 (8.26)	2.13 (7.03)	1.73 (5.61)	-.15 (.93)
CLLIP \times risk status	5.85 (5.25)	6.12 ⁺ (3.72)	9.88* (4.37)	.50 (1.01)
Random effects (variance components):				
Classroom	24.47 (26.80)	20.56 (20.61)	16.21 (18.43)	.21 (.46)
Student	137.95*** (19.38)	72.67*** (10.18)	99.71*** (13.96)	5.54*** (.78)
-2LL	855.3	785.86	821.9	515.3

⁺ $p < .10$.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

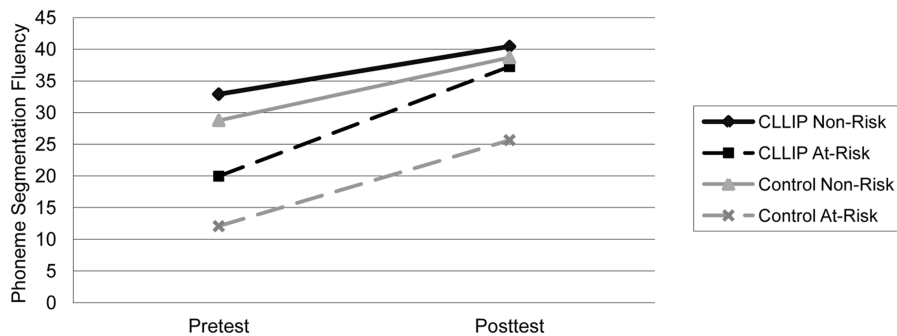


Figure 1. Kindergarten phoneme segmentation fluency gains by intervention group moderated by risk status.

For Phoneme Segmentation Fluency, at-risk CLLIP students scored almost 10 points higher at posttest, on average, compared to at-risk control students (see Fig. 1). For Initial Sound Fluency, the association between CLLIP intervention status \times risk status approached significance; at-risk CLLIP students had a trend toward greater gains (6.1 points) compared to at-risk control students. In both cases, gains for students were moderated by risk status, with posttest scores of at-risk CLLIP students similar to nonrisk CLLIP and control peers. For Phoneme Segmentation Fluency, individual student pretest score was positively related to posttest score (1 point difference above the pretest grand mean was associated with a 0.6 point difference in posttest score), while boys tended to score 4 points lower than girls at the end of the year. For Letter Naming Fluency, risk status was a significant predictor (at-risk students scoring 11.3 points lower, on average), and (level 2) mean classroom pretest score was positively related to posttest score (0.6 points higher for every one point difference in mean classroom score).

Grade 4 Student Outcomes

Results for the first research question showed a significant main effect of the intervention (level 2) for one out of the four assessments. CLLIP students scored significantly higher (3 points) on Letter-Word Identification at posttest (Table 4). For the second research question, testing moderation, there were two interaction effects of intervention status \times risk status. At-risk CLLIP students scored slightly higher than both at-risk control and nonrisk students on Fluency posttest (4.3 points), on average (see Fig. 2). Second, at-risk CLLIP students made significantly greater vocabulary gains at posttest (1.5 points) compared to at-risk control peers, ending the year with scores similar to nonrisk students (see Fig. 3). Individual pretest score on Vocabulary was associated with a 2.2 difference in posttest, on average, and at-risk students scored 2.1 points lower than nonrisk students. Individual student pretest scores for Word Recognition, Comprehension, and Fluency were positively related to those posttest measures.

Across these two sets of multilevel models (students nested in kindergarten and grade 4 classrooms), the level 2 predictors of teacher education level and mean classroom pretest score were not associated with outcomes. In each case, results indicated

Table 4. Grade 4 Fixed Effects and Variance Estimates for Posttest Outcomes ($n = 138$ Students in 6 Classrooms)

	Letter-Word Identification Parameter (SE)	Passage Comprehension Parameter (SE)	Fluency Parameter (SE)	Picture Vocabulary Parameter (SE)
Fixed effects:				
Intercept	49.64 *** (2.31)	29.06 (26.62)	42.47 *** (4.67)	25.33 *** (1.07)
Level 1:				
Student pretest score				
centered	2.56 * (1.04)	.80 *** (.06)	.65 * (.32)	2.16 * (.77)
Gender (male = 1)	.79 (.62)	.33 (.51)	-.39 (1.03)	.73 * (.36)
Risk status (1 = at risk)	-1.37 (1.24)	-1.02 (.03)	-2.51 (2.03)	-2.09 *** (.62)
Level 2:				
Teacher education level	.58 (1.08)	-.40 (1.51)	1.72 (2.71)	.58 (.47)
Mean classroom pretest				
score	-1.59 (1.03)	.06 (.96)	.41 (.32)	-1.33 (.77)
CLLIP	3.00 * (1.51)	.97 (1.53)	.53 (2.25)	.40 (.56)
CLLIP \times risk status	.96 (1.38)	-.22 (1.13)	4.31 * (2.24)	1.51 * (.74)
Random effects (variance components):				
Classroom	-.02 (.56)	.65 (1.02)	.67 (2.14)	-.11 (.06)
Student	12.43 *** (1.56)	8.28 *** (1.04)	32.09 *** (4.03)	3.68 *** (.46)
-2LL	722.0	671.3	849.5	559.2

* $p < .05$.** $p < .01$.*** $p < .001$.

additional student-level random effects that could be explained by unobserved measures, but no additional classroom-level variance was indicated.

Discussion

Results from this exploratory study allowed us to examine whether gains in targeted areas of literacy instruction for students in emerging (kindergarten) and established (grade 4) reading levels were related to a main effect of CLLIP professional development. A review of mean differences in scores for kindergarten showed that while average gains in Letter Naming, Initial Sound, Phoneme Segmentation, and Vocabulary were consistently higher for CLLIP students, they were not large enough to be

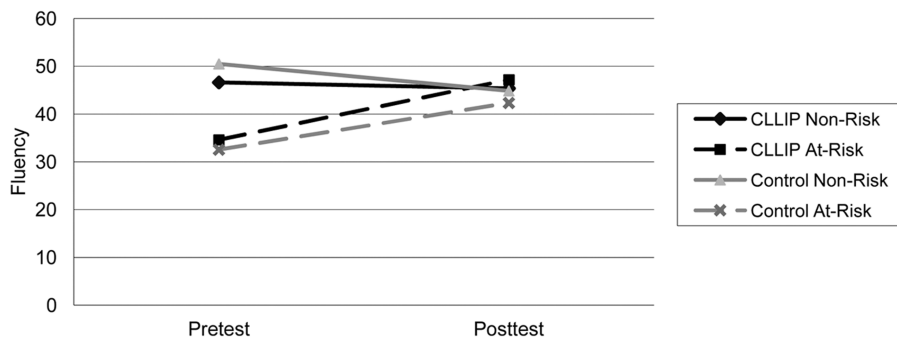


Figure 2. Grade 4 fluency gains by intervention group moderated by risk status.

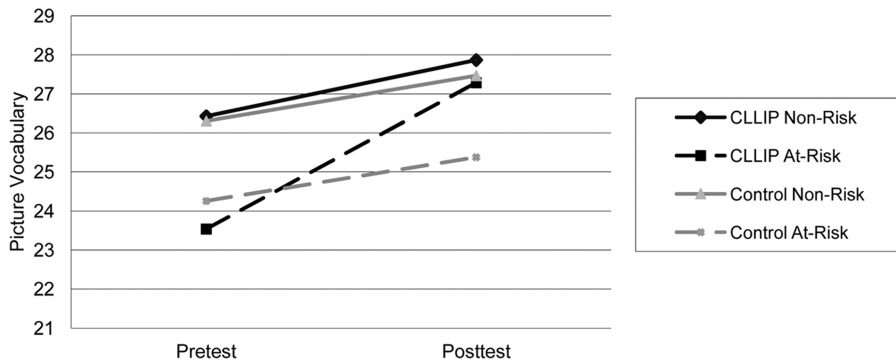


Figure 3. Grade 4 picture vocabulary gains by intervention group moderated by risk status.

statistically significant. Results for grade 4 students showed that average gains were significantly higher for CLLIP students in one of the four assessments: Word Recognition. The lack of evidence for a main effect may reflect the influence of scoring for low-risk students who had limited room for improvement compared to at-risk peers; regression to the mean was exemplified in the grade 4 fluency results.

Second, results revealed the extent that language and literacy gains for intervention students were moderated by risk status, thus suggesting a positive effect of CLLIP small-group instructional strategies. Greater benefits for at-risk students compared to low-risk peers may reflect the substantial programmatic focus on differentiated instruction. Kindergarten CLLIP students most at risk showed greater gains in Phonemic Awareness and Initial Sounds Fluency compared to peers whose teachers were not receiving CLLIP professional development training and coaching. Students from low-income families are often at a disadvantage compared to their more affluent peers, as they tend to start out in formal schooling with more limited language and literacy skills, often reflecting a lack of resources in the home for literacy support and lack of access to quality preschool programs (Lindjor, 2003). Thus, kindergarten is a critical period for literacy instruction that can mitigate against a looming possibility of the Matthew effect (i.e., starting school at an academic disadvantage that leads to reduced access to literacy experiences, thus generating a cumulative disadvantage; Stanovich, 1986).

CLLIP at-risk students in grade 4 made significant gains in fluency and vocabulary measures compared to peers in control classrooms. That we found no effect of the intervention on grade 4 comprehension may suggest the need for a stand-alone comprehension module in addition to the existing integration of comprehension instruction strategies within the other modules, or further refinement of existing comprehension training within the other modules. Making strong progress in grade 4 is important not just because of the pressures of high-stakes testing, but because this is a time when many students encounter a “fourth-grade slump” (Chall, Jacobs, & Baldwin, 1990) that can predict ultimate drop-out or academic underperformance.

Results of the regression models suggested that classrooms were similar, with little additional variance explained by classroom-level predictors. We found no evidence that CLLIP teachers with advanced degrees produced improved student outcomes. However, there was much unexplained variance in unmeasured student-level predictors. Early literacy influences in the home have been found to be strongly related

to reading outcomes and to have long-term consequences; even in a sample that appears to be homogeneous in socioeconomic status, there exists variation in financial and social support resources that can have an impact on literacy development (Snow, Porche, Tabors, & Harris, 2007).

Descriptive and exploratory results suggest the potential for effective implementation of CLLIP as a strategy toward establishing whole-school literacy reform to address the needs of at-risk students. This study documents the extensive commitment necessary to unite a district toward this common goal. The CLLIP intervention requires time and focus; teachers and administrators might need to learn and practice new and unfamiliar skills and instructional strategies, often disrupting their comfortable, established routines. Moreover, they are asked to integrate the theoretical basis of these instructional strategies into their own learning. A challenge to professional development is resistance from teachers who comment that training modules include too much “dry” theory and not enough demonstration of “activities.” By allowing for conflict and allowing teachers to give voice to their reservations, past disappointments, and cynicism, which is an essential part of the teaching and learning community (Achinstein, 2002), CLLIP coaches provided space for teachers to prove to themselves that scientifically based practices work. By doing this as a group, they also became more skilled in working together as a team. Descriptive survey results found only one teacher who reported an influence of advanced coursework on daily teaching practices, underscoring the need for continued on-site support of teaching strategies learned in advanced degree programs. Regular interactions with the CLLIP coaches were critical to introducing theoretical knowledge into regular classroom practice, although the extent of teacher change facilitated by ongoing support provided by the literacy coaches was not assessed.

Limitations

Resistance from districts to randomization of the intervention and to providing information about individual student characteristics increases the likelihood that other unobserved variables influenced change in student improvement—for instance, differences in teacher background and motivation, or student configurations within classrooms (e.g., free or reduced-price lunch status). The quasi-experimental design of this study introduces the possibility of selection bias at both the district and teacher levels. Because we were unable to assess growth over multiple years, we need to continue to monitor teachers and students to evaluate the long-term changes in both teacher practices and student learning; we also need to be able to discern the impact of CLLIP versus a potential Hawthorne effect. Because of the sampling limitations for these studies, and because multiple analyses were conducted with a limited number of participants, exploratory results must be interpreted with caution. More rigorous investigation of promising exploratory results requires a randomized control trial with a greater number of schools and with access to more exhaustive student- and teacher-level data; this would require strong support from state and district leadership. Ultimately, the evaluation of socially complex service interventions with a multifaceted organization, such as a school district, will necessarily involve challenges as a result of staffing arrangements and differently motivated populations (Wolff, 2000).

That the intervention teachers administered their own assessments (due to financial constraints of implementation) is also a limitation. Ideally, tests should be administered by independent researchers blind to group assignment. Additionally, critiques about the DIBELS assert that emphasis on assessment of constrained skills with known ceiling effects (letter knowledge, phonics, concepts of print) versus unconstrained, open-ended skills (vocabulary, comprehension) may lead teachers to misinterpret student reading skill and subsequently lead to errors in diagnosis of reading difficulties (Paris, 2005). Because the DIBELS, in particular, emphasizes measurement of constrained skills and focuses on speed of processing rather than depth of knowledge, it may actually limit effective reading instruction if teachers use it to narrowly guide curriculum (Goodman, 2006). With these caveats in mind, we acknowledge the limitations of these assessments and their administration, but assert that teachers' use of specific tools was secondary to broader training in the understanding of reading development and scientifically based practices in reading instruction.

Next Steps

Additional research is needed to identify specific mechanisms of change related to CLLIP training and coaching. Prochaska and DiClemente's (1983) stages-of-change model can be applied to understanding teacher change, in particular, resistance to doing something new as opposed to using the same ineffective strategies year after year. CLLIP teacher participation, as reflected in satisfaction ratings and comments regarding training, included strong resistance to change at first (precontemplation), coupled with concern about the lack of progress in student learning and a search for solutions (contemplation), along with a willingness to engage in professional development trainings (preparation). These are patterns found in other research on instructional change related to teacher training and coaching (DuFour, 2007; Klingner, 2004). Professional development trainings may be viewed as the easy fix, especially if teachers prefer brief trainings with a "cookbook" type of instructional approach. Coaching or other forms of sustained support can help teachers do the difficult work of implementing more rigorous change with the goal of school reform as they modify and learn new instructional practices (action) and build them into a regular routine (maintenance). These last two stages are contingent on the teacher becoming comfortable with being observed and reflecting on teaching practice in collaboration with a more experienced other who assists performance (Tharp & Gallimore, 1988). The role of a literacy coach is not that of another supervisor, as Toll explained, but rather a partner in teacher change, providing "one-on-one conferences with teachers, facilitation of small-group discussions among teachers, and demonstration lessons in teachers' classrooms, which build on the conversations that coaches have with teachers" (2009, p. 65). Exposure to professional development without a coaching component undermines the possibility that teachers would progress beyond the preparation stage of change in their instructional practices.

Supovitz, Sirinides, and May (2010) demonstrated the influence of principal leadership on change in teachers' instructional practice, which in turn mediates the influence of leadership on student achievement. Thus efficacy in whole-school reform is limited by the degree to which administrators actively participate in training (gaining theoretical knowledge while also setting an example of commitment to reform) and observations

with coaches (gaining supervisory and literacy coaching skills). For this study, all teachers and administrators attended the first two trainings, but administrator attendance was less consistent for the next four modules. All administrators received vocabulary training, but almost half missed phonological awareness and writing, thus potentially limiting their understanding of how these various components are interrelated in reading instruction. Future research should explore the connections between administrator involvement and student literacy outcomes.

Presently, most high-stakes tests in Ohio, such as the Ohio Achievement Test, Stanford Achievement Tests, and other major benchmark assessments, are administered by teachers and then mailed to a central distribution center. The results of these assessments are then returned to schools and teachers in the form of composite results. In contrast, the assessments carried out as part of CLLIP are an integral component of professional development for teachers who utilized test results as a basis for implementing differentiated instruction in their classrooms. This shift in assessment practice is one example of how to maximize the utility of research in the classroom, making it more meaningful and directly beneficial for teachers compared to the models of mandated proficiency testing they are accustomed to (Boudett et al., 2005). One of the most dramatic changes for teachers was in their use of differentiated instruction in small-group settings. Teachers were initially resistant to this change, and their preference for whole-class instruction may have also reflected their reliance on methods in which they were trained. More rigorously designed studies should examine whether improvement in student literacy scores can be attributed to this lengthy process of teachers becoming more comfortable in attending to individualized student needs in small groups (Foorman & Torgesen, 2001).

Formal standardized assessment of improvement in teacher instructional practices using new observational tools (e.g., Walpole et al., 2010) is essential to understanding the impact of coaching on instruction. If we are truly confident that we have identified best practices in reading instruction, then it seems most appropriate that teacher observation is used for accountability in order to ensure that those demonstrably effective practices are fully implemented. Ideally, observation teams comprising an independent observer and a school or district administrator would document both teacher practice in the classroom and supervisor knowledge of literacy instruction. Yet, in school districts where CLLIP has been implemented (as is the case nationwide), only student proficiency results are directly tied to state funding, and thus other measures have little significance.

Moats (1999) argued that access to systematic lessons in reading from a knowledgeable teacher using a well-designed approach based on evidence about the most effective reading instruction is denied to many students. She suggested that the responsibility of using evidence-based and agreed-upon methods is as important in education as it is for health care. CLLIP shows promising results as a school reform intervention, especially in working with students at risk. This highlights the program's potential for translating research into practice through professional development training that is supported through continuous on-site coaching. Previous studies have documented the effects of such professional development on teacher beliefs and teacher practices (Englert, Raphael, & Marriage, 1998), but we have also attempted to document potential effects on student learning. These exploratory results warrant continued investigation of the CLLIP intervention as a scientifically based design that could support a collaborative staff approach to comprehensive literacy reform.

Appendix A

Table A1. Selected Examples of Materials and Scientifically Based Research Included in Training Modules

Module	Selected Examples of Training Materials
Orientation and strategic assessment	K–3: Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002) K–6: Woodcock Language Proficiency Battery—Revised (WLPB-R; Woodcock et al., 2001) <i>Preventing Reading Difficulty in Young Children</i> (Snow et al., 1998) <i>What Teachers Need to Know about Language</i> (Adger et al., 2002)
Targeted instruction for small-group intervention	1–2: Reading Intervention for EARLY SUCCESS (Houghton Mifflin, 2003) 3–6: SOAR to Success (Cooper et al., 1997)
Phonological awareness	“Phonemic Awareness Instruction Helps Children Learn to Read: Evidence from the National Reading Panel’s Meta-Analysis” (Ehri et al., 2001) <i>Teaching Reading Is Rocket Science: What Expert Teachers of Reading Should Know and Be Able to Do</i> (Moats, 1999).
Fluency	<i>The Fluent Reader: Oral Reading Strategies for Building Word Recognition, Fluency, and Comprehension</i> (Rasinski, 2003)
Vocabulary	<i>Bringing Words to Life: Robust Vocabulary Instruction</i> (Beck, McKeown, & Kucan, 2002)
Writing	<i>Writing Workshop: The Essential Guide</i> (Fletcher & Portalupi, 2001)

Notes

Funding for the Collaborative Language and Literacy Instruction Project (CLLIP) was provided to the project developer (second author) by the 124th General Assembly of Ohio and the Ohio Department of Education. A portion of these funds was used to support an independent evaluation of CLLIP by the first author. This evaluation was conducted free from any outside influence of CLLIP, the Ohio General Assembly, or the Department of Education, and was done with no vested interest in the outcome. Address all correspondence to Daniel H. Pallante, Ohio Educational Development Center, 1915 Newark-Granville Rd., Granville, OH 43023; e-mail: dpallante@roadrunner.com.

1. In one grade 4 classroom, nine students were identified as at risk and divided into two groups, five in a group run by the classroom teacher and four in another group run by a Title I teacher. Both groups received the same small-group intervention.

References

Achinstein, B. (2002). Conflict amid community: The micropolitics of teacher collaboration. *Teachers College Record*, 104, 421–455. doi:10.1111/1467-9620.00168

Adger, C. T., Snow, C. E., & Christian, D. (Eds.). (2002). *What teachers need to know about language*. McHenry, IL: Delta Systems and the Center for Applied Linguistics.

Al Otaiba, S., Hosp, J. L., Smartt, S., & Dole, J. A. (2008). The challenging role of a reading coach, a cautionary tale. *Journal of Educational & Psychological Consultation*, 18, 124–155. doi:10.1080/10474410802022423

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York: Guilford.

Boudett, K. P., City, E., & Murnane, R. (Eds.). (2005). *Data wise: A step-by-step guide to using assessment results to improve teaching and learning*. Cambridge, MA: Harvard Education Press.

Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.

- Commeyras, M., & DeGroff, L. (1998). Literacy professionals' perspectives on professional development and pedagogy: A United States survey. *Reading Research Quarterly*, **33**, 434–472. doi:10.1598/RRQ.33.4.5
- Cooper, J. D., Boschken, I., McWilliams, J., & Pistochini, L. (1997). *Project Success: A study of the effectiveness of an intervention program designed to accelerate reading for struggling readers in the upper grades*. Unpublished manuscript.
- Correnti, R., & Rowan, B. (2007). Opening up the black box: Literacy instruction in schools participating in three comprehensive school reform programs. *American Educational Research Journal*, **44**, 298–338. doi:10.3102/0002831207302501
- Cunningham, A. E., Zibulsky, J., Stanovich, K. E., & Stanovich, P. J. (2009). How teachers would spend their time teaching language arts: The mismatch between self-reported and best practices. *Journal of Learning Disabilities*, **42**, 418–430. doi:10.1177/0022219409339063
- Darling-Hammond, L., Hightower, A. M., Husbands, J. L., LaFors, J. R., Young, V. M., & Christopher, C., (2003). *Building instructional quality: "Inside-out" and "outside-in" perspectives on San Diego's school reform. A research report*. Seattle: Center for the Study of Teaching and Policy.
- DuFour, R. (2007). Professional learning communities: A bandwagon, an idea worth considering, or our best hope for high levels of learning? *Middle School Journal*, **39**, 4–8.
- Duke, N. K., & Martin, N. M. (2008). Comprehension instruction in action: The elementary classroom. In C. C. Block & S. R. Parris (Eds.), *Comprehension instruction: Research-based best practices* (2nd ed., pp. 241–257). New York: Guilford.
- Eaker, R., & Keating, J. (2008). A shift in school culture: Collective commitments focus on change that benefits student learning. *Journal of Staff Development*, **29**, 14–17.
- Edwards, P. A., Turner, J. D., & Mokhtari, K. (2008). Balancing the assessment "of" learning and "for" learning in support of student literacy achievement. *Reading Teacher*, **61**, 682–684.
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, **36**, 250–287.
- Englert, C. S., Raphael, T. E., & Mariage, T. V. (1998). A multi-year literacy intervention: Transformation and personal change in the community of the Early Literacy Project. *Teacher Education and Special Education*, **21**, 255–277.
- Fletcher, R., & Portalupi, J. (2001). *Writing workshop: The essential guide*. Portsmouth, NH: Heinemann.
- Foorman, B. R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities: Research & Practice*, **16**, 203–212.
- Francis, D. J., Fletcher, J. M., Catts, H. W., & Tomblin, J. B. (2005). Dimensions affecting the assessment of reading comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 369–394). Mahwah, NJ: Erlbaum.
- Goldenberg, C. (2003). Settings for school improvement. *International Journal of Disability, Development and Education*, **50**, 7. doi:10.1080/1034912032000053304
- Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Education Achievement. Retrieved from <http://dibels.uoregon.edu/>
- Goodman, K. (Ed.). (2006). *The truth about DIBELS: What it is, what it does*. Portsmouth, NH: Heinemann.
- Holland, H. (2005). Teaching teachers: Professional development to improve student achievement. *Research Points*, **3**, 1–4.
- Houghton Mifflin. (2003). *Reading Intervention for EARLY SUCCESS*. Boston: Houghton Mifflin.
- Klingner, J. K. (2004). The science of professional development. *Journal of Learning Disabilities*, **37**, 248–255.
- Knight, J. (2009). Coaching: The key to translating research into practice lies in continuous, job-embedded learning with ongoing support. *Journal of Staff Development*, **30**, 18–20.
- Lindjord, D. (2003). The start of a new school year: Research-based classroom practices for improving student achievement in children from high poverty and minority families. *Journal of Early Education and Family Review*, **11**, 4–5.

- McCombs, J. S., & Marsh, J. A. (2009). Lessons for boosting the effectiveness of reading coaches. *Phi Delta Kappan*, **90**, 501–507.
- Moats, L. C. (1999). *Teaching reading is rocket science: What expert teachers of reading should know and be able to do*. Washington, DC: American Federation of Teachers.
- Nathanson, S., Pruslow, J., & Levitt, R. (2008). The reading habits and literacy attitudes of inservice and prospective teachers: Results of a questionnaire survey. *Journal of Teacher Education*, **59**, 313–321. doi:10.1177/0022487108321685
- National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel. A scientific synthesis of early literacy development and implications for intervention*. Jessup, MD: National Institute for Literacy.
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.
- National Staff Development Council. (2001). *NSDC's standards for staff development*. Retrieved from <http://www.nsdc.org/standards/index.cfm>
- Office of Curriculum and Instruction at the Ohio Department of Education. (2001). *Academic content standards: K–12 English language arts*. Retrieved from http://www.ode.state.oh.us/academic_content_standards/pdf/ENGLISH.pdf
- Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, **40**, 184–202. doi:10.1598/RRQ.40.2.3
- Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology*, **51**, 390–395. doi:10.1037/0022-006X.51.3.390
- Rasinski, T. V. (2003). *The fluent reader: Oral reading strategies for building word recognition, fluency, and comprehension*. New York: Scholastic.
- Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, **24**, 323–355. doi:10.2307/1165280
- Snow, C. E., Burns, S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Snow, C.E., Griffin, P., Burns, M.S., & the NAE Subcommittee on Teaching Reading. (2005) *Knowledge to support the teaching of reading: Preparing teachers for a changing world*. San Francisco: Jossey-Bass.
- Snow, C. E., Porche, M. V., Tabors, P. O., & Harris, S. R. (2007). *Is literacy enough? Pathways to academic success for adolescents*. Baltimore, MD: Brookes.
- Stahl, S. A. (2003). Teaching phonics and phonological awareness. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (Vol. 1, pp. 333–347). New York: Guilford.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, **21**, 360–407.
- Sterbinsky, A., Ross, S. M., & Redfield, D. (2006). Effects of comprehensive school reform on student achievement and school change: A longitudinal multi-site study. *School Effectiveness and School Improvement*, **17**, 367–397.
- Supovitz, J., Sirinides, P., & May, H. (2010). How principals and peers influence teaching and learning. *Educational Administration Quarterly*, **46**, 31–56. doi:10.1177/1094670509353043
- Taylor, B. M., Peterson, D. S., Pearson, P. D., & Rodriguez, M. C. (2002). Looking inside classrooms: Reflecting on the “how” as well as the “what” in effective reading instruction. *Reading Teacher*, **56**, 270–279.
- Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context*. New York: Cambridge University Press.
- Toll, C. A. (2009). Literacy coaching. In J. Knight (Ed.), *Coaching: Approaches and perspectives* (pp. 56–69). Thousand Oaks: Corwin.
- Valdés, G., Bunch, G., Snow, C., Lee, C., & Matos, L. (2005). Enhancing the development of students’ language(s). In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do*. (pp. 126–168). San Francisco: Jossey-Bass.

- Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education*, **24**, 80–91. doi:10.1016/j.tate.2007.01.004
- Walpole, S., McKenna, M. C., Uribe-Zarain, X., & Lamitina, D. (2010). The relationships between coaching and instruction in the primary grades. *Elementary School Journal*, **111**, 115–140. doi:10.1086/653472
- Wolff, N. (2000). Using randomized controlled trials to evaluate socially complex services: Problems, challenges and recommendations. *Journal of Mental Health Policy and Economics*, **3**, 97–109. doi:10.1002/1099-176X(200006)3:2<97::AID-MHP77>3.0.CO;2-S
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III: Tests of achievement*. Itasca, IL: Riverside.
- Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement: Issues & answers*. REL 2007-No. 033. Regional Educational Laboratory Southwest.