4-11-2019

Supporting teachers' use of data-based instruction to improve students' early writing skills

Kristen L. McMaster

Erica S. Lembke

Jaehyun Shin

Apryl L. Poch

Alex R. Smith

See next page for additional authors

Authors
Kristen L. McMaster, Erica S. Lembke, Jaehyun Shin, Apryl L. Poch, Alex R. Smith, Pyung-Gang Jung, Abigail A. Allen, and Kyle Wagner

Supporting Teachers' Use of Data-Based Instruction to Improve Students' Early Writing Skills

Resubmitted February 14, 2019

Abstract

We examined the effects of a professional development (PD) system designed to support teachers' use of data-based instruction (DBI) to improve early writing outcomes for children with intensive needs. The PD system, called DBI-TLC, provided *tools* for early writing assessment and intervention, *learning modules* including face-to-face workshops followed by classroom application, and ongoing *coaching* to support DBI implementation. Special education teachers in 19 classrooms in two Midwestern districts were assigned randomly to receive DBI-TLC or to a business-as-usual control group. All teachers completed pre- and posttests of DBI knowledge and skills and self-efficacy, and DBI-TLC teachers' fidelity to DBI was assessed. Fifty-three students (two to three from each classroom) completed pre- and posttests of early writing using Curriculum-Based Measures (CBM) and the Test of Early Written Language-3 (TEWL-3; Hresko, Herron, Peak, & Hicks, 2012). DBI-TLC teachers outperformed controls at posttest on DBI knowledge and skills (Hedge's $g = 2.88$), and reported a more explicit writing instruction orientation compared to controls ($g = 1.63$). DBI fidelity varied (on average, 84% for assessment, 79% for intervention, and 52% for decision-making). Students whose teachers implemented DBI showed a pattern of stronger early writing performance compared to control students on CBM, with effect sizes of 0.23 to 0.40, but not on the TEWL-3 (0.02 to 0.13). We discuss the promise of DBI-TLC to improve teacher practice and student outcomes, as well as the need to continue to explore ways to support teachers' implementation of DBI with fidelity.

Key words: data-based instruction, early writing, professional development, coaching

Educational Impact and Implications Statement

Many teachers experience difficulty individualizing instruction for children with the most intensive academic needs, particularly in the area of writing. In this randomized control trial, we examined the effects of providing elementary special education teachers with tools, learning opportunities, and coaching to support their implementation of data-based instruction (DBI) in early writing. Teachers who received these supports improved their knowledge and skills related to DBI in early writing compared to controls, and their students showed promising early writing outcomes.

Supporting Teachers' Use of Data-Based Instruction to Improve Students' Early Writing Skills

Many students benefit from research-based standard intervention protocols; however, a small proportion of students—many of whom are at risk or identified with disabilities—does not. These students require more intensive, individualized instruction (Fuchs, Fuchs, & Stecker, 2010; Wanzek & Vaughn, 2009). Yet, educators may find it difficult to individualize instruction effectively, in part because many have not received adequate preparation to do so (cf. Roehrig, Duggar, Moats, Glover, & Mincey, 2008). Thus, researchers and policymakers have called for efforts to improve educators' knowledge and skills related to effective instructional individualization (e.g., Lemons, Al Otaiba, Conway, & Mallado, 2016; Zumeta, 2015).

Effective individualization requires a framework that guides teachers to decide when and how to modify a student's instruction. One such framework—originally termed Data-Based Program Modification (Deno & Mirkin, 1977) and more recently referred to as Data-Based Instruction or Data-Based Individualization (DBI; Fuchs et al., 2010; National Center on Intensive Intervention [NCII], n.d.)—is well suited for this purpose. DBI entails a systematic, hypothesis-driven cycle of assessment and intervention using research-based instruction and adaptations (NCII, n.d.). Previous research has shown that teachers' use of DBI can lead to improved student outcomes (Jung, McMaster, Kunkel, Shin, & Stecker, 2018; Stecker, Fuchs, & Fuchs, 2005), particularly when professional development (PD) and support is in place. In this paper, we report effects of a PD system to support teachers' use of DBI for children with intensive early writing needs.

Importance of Professional Development to Support Early Writing Development

National efforts to improve teachers' individualization of instruction have primarily focused on reading and mathematics, with much less attention on writing. Yet, learning to write is an essential part of students' literacy development (Biancarosa & Snow, 2004). Writing is strongly related to reading (Graham & Hebert, 2010), is important for integrating knowledge and thinking critically (Shanahan, 2004), and is the primary means for students to communicate what they know (Graham & Perin, 2007). Lack of writing proficiency can significantly limit postsecondary and employment opportunities (Graham & Perin, 2007), and students with disabilities are particularly at risk. For example, in the 2011 National Assessment of Educational Progress (NAEP), only 5% of eighth graders with learning disabilities (LD) reached proficiency in writing and 60% performed below basic (National Center for Education Statistics, 2012).

Whereas early identification and effective, individualized writing instruction can prevent long-term negative consequences for many students (Berninger, Nielson, Abbott, Wijsman, & Rasking, 2008), few teachers are prepared to deliver such instruction. In a national survey of elementary teachers, more than half cited their teacher education programs as poor to only adequate in preparing them to deliver effective writing instruction (Cutler & Graham, 2008). Further, Troia et al. (2015) reported that typical classroom writing instruction does not include evidence-based practices. And, in a study of the relation between teacher knowledge and time allocation in literacy instruction (Spear-Swerling & Zibulsky, 2014), many general and special education teachers allocated little to no time to assessment or instruction in basic writing skills, despite evidence that such skills are foundational for writing proficiency (Berninger et al., 2008).

It is clear that teachers require better preparation to provide high-quality early writing instruction and individualized intervention for children who struggle. Thus, our research team set out to develop a PD system to support teachers' use of DBI to improve outcomes for children with intensive early writing needs, as part of a three-year development (Goal 2) project funded by the Institute of Education Sciences. We called this system "DBI-TLC" because it provides Tools, Learning modules, and Collaborative supports for teachers as they implement DBI.

DBI-TLC Components

During the first two years of the project, we iteratively developed and examined the usability and feasibility of DBI-TLC (see Lembke et al., 2018; Poch, McMaster, & Lembke, 2018). Below, we describe the DBI-TLC components and their theoretical and empirical bases.

DBI Tools

Manual. The DBI manual (McMaster & Lembke, 2014), modeled after work conducted by Deno and Mirkin (1977), Fuchs and Fuchs (2007), and NCII (intensiveintervention.org/), includes step-by-step instructions for implementing the DBI process: (1) establish student's present level of performance, (2) set a reasonable but ambitious long-term goal, (3) implement high-quality research-based instruction with fidelity, (4) monitor progress toward the goal, (5) use decision rules to determine when instructional changes are needed, (6) generate hypotheses about how to individualize instruction, (7) change instruction based on hypotheses, and (8) continue steps 4-7. Each step is illustrated by a case example of a teacher's implementation of DBI with a student with intensive writing needs, along with practice activities, resources, and sample materials.

Content in the DBI manual is supported by research showing the efficacy of DBI to improve teacher practice, such as more frequent and appropriate instructional changes targeting students' specific needs, as well as student outcomes in basic academic areas including reading and math (Stecker et al., 2005). In a recent meta-analysis, Jung et al. (2018) found an overall DBI effect size of $g = 0.37$. Further, the efficacy of DBI has recently been expanded to writing for young elementary students: Jung, McMaster, & delMas (2017) found that students in Grades

1-3 who received research-based writing intervention delivered in a DBI framework outperformed controls in early writing skills, with effect sizes of g = 0.45 to 1.36.

Assessment tools. Assessment tools include three different Curriculum-Based Measurement (CBM; Deno, 1985) tasks for beginning writers (word dictation, picture word, and story prompts). We provide 20 forms of each task, along with administration and scoring directions and graphing tools. We also provide a rubric to aid teachers in selecting an appropriate CBM task by thinking through a student's writing strengths, needs, and long-term goals to choose a task that would reflect meaningful progress for that student.

The three CBM tasks, administered for 3 min each, are supported by empirical research (see McMaster, Ritchey, & Lembke, 2011 for a review). Word dictation is designed to measure word-level writing, picture word is designed to measure sentence-level writing, and story prompts are designed to measure passage-level writing. Student responses are scored for total number of words written (WW), words spelled correctly (WSC), correct letter sequences (CLS; any two adjacent letters correctly placed according to the correct spelling of the word), correct minus incorrect letter sequences (CILS), correct word sequences (CWS; any two adjacent words spelled and used correctly in the context of the sentence; Videen, Deno, & Marston, 1982) and correct minus incorrect word sequences (CIWS). These measures have evidence of reliability, criterion validity, and sensitivity to growth in grades 1-3 (see Measures for more details).

Instructional tools. We developed tools to support teachers' implementation of high-quality, research-based writing instruction, including: (a) diagnostic tools that guide teachers to identify students' writing strengths and needs, align these needs with standards, and match needs to appropriate interventions; (b) skills-based mini-lessons and materials; and (c) a Writing Instructional Plan (WIP)—a template for creating individualized writing instructional programs.

Development of these tools was guided by a theoretical model called the Simple View of Writing. This model specifies three main components of early writing development: transcription, text generation, and self-regulation, each of which are constrained by attention and memory, such that difficulty in any one area may limit overall writing proficiency (Berninger & Amtmann, 2003). A comprehensive review of early writing intervention research (McMaster, Kunkel, Shin, Jung, & Lembke, 2018) revealed that instruction focusing on each component can improve composition skills for students at risk or with disabilities: Explicit, systematic handwriting and spelling interventions have improved students' transcription skills as well as composition quantity (Hedge's $g = 0.46$) and quality ($g = 0.17$), and interventions targeting a combination of text generation and self-regulation skills have improved composition quantity ($g = 0.53$-1.61) and quality ($g = 0.55$-1.18).

Diagnostic tools. Diagnostic tools consist of a series of checklists that align with the three main components of the Simple View of Writing (transcription, text generation, and self-regulation). Each checklist encourages the teacher to observe both the student's writing *process* (behavior during writing) and *product* (the result of the student's work) to gain insights into the student's strengths and needs. For example, the transcription checklist includes processes and products related to both handwriting (e.g., "Does the student hold the pen or pencil comfortably? Write fluidly without excessive scribbling or erasing?) and spelling (e.g., "Does the student consistently use the correct consonant at the beginning and end of words?"). The teacher can then use this information to determine the focus and content of instruction and match mini-lessons to the student's needs (e.g., the teacher might select spelling lessons for a student with relative strengths in handwriting and weaknesses in spelling).

Mini-lessons. Table 1 provides an overview of research-based mini-lessons that align with the transcription and text-generation components of the Simple View of Writing. Transcription lessons focus on handwriting and spelling, and text generation lessons focus on sentence construction and combining, along with fluency building. For text generation at the passage level along with self-regulation, we provided all teachers with a copy of *Powerful Writing Strategies for All Students* by Harris, Graham, Mason, and Friedlander (2008), because the lessons in this text are based on the widely researched Self-Regulated Strategy Development approach (see Baker, Chard, Ketterlin-Geller, Apichatabutra, & Doabler, 2009 for a review), and we saw no need to duplicate an existing compilation of research-based instructional lessons.

Writing Instructional Plan (WIP). The mini-lessons described above were designed to be used either in isolation or as part of a more comprehensive instructional plan. We designed a WIP template that teachers could use to construct a plan customized to individual student needs. For example, if a student's needs were primarily related to handwriting and spelling, the teacher might compile a WIP that included the set of transcription mini-lessons. Or, if a student had a combination of needs related to spelling and sentence construction, the teacher might compile a WIP that included the spelling-focused transcription lessons along with sentence-level text-generation lessons. In this way, the teacher could use research-based interventions *and* customize the overall instructional program according to individual student needs. Development of the WIP was guided by the diagnostic checklists and supported by a coach assigned to the teacher.

Decision-making tools. In their review of teachers' use of CBM, Stecker et al. (2005) identified application of data-based decision rules as a critical component influencing student outcomes. Researchers have recommended applying decision rules after collecting three to 10 data points, with seven data points used most frequently (Ardoin, Christ, Morena, Cormier, &

Klingbeil, 2013). In this study, we recommended that teachers collect at least eight data points before making their first decision, and every six to eight data points thereafter, based on evidence that at least eight data points are needed to obtain a reliable and stable slope of progress (McMaster et al., 2011).

Decision-making tools include a Decision-Making Rubric that guides teachers to examine graphed CBM data to decide whether to *raise the goal* when the student's progress is greater than expected in relation to the long-term goal, *keep instruction as-is* when the student is making expected progress, or *change instruction* when the student's progress is less than expected. If the decision is to *change instruction*, additional guidance is provided to support the teacher's generation of hypotheses about what type of change is needed. Specifically, the teacher is prompted through a series of self-questions to determine whether (a) instruction was implemented with fidelity (if not, fidelity should be corrected before making additional changes), (b) the student needs a change in *focus* (i.e., the teacher may need to reconsider whether to focus on transcription, text generation, and/or self-regulation) or *content* (i.e., the teacher may need to reassess what letters, words, topics, and so on to include in lessons), or (c) a change in instructional *delivery* is needed (e.g., a change in dosage, group size, explicitness, or motivational components). The teacher forms a hypothesis, makes a corresponding instructional change, and documents this change on the student's graph and (for this study) in a Decision Log.

Learning Modules

Learning modules, designed to provide teachers with knowledge and skills to implement DBI, are delivered face-to-face via a series of four day-long workshops. Each module includes learning objectives, an overview of the research base supporting the relevant DBI steps, modeling, and opportunities to apply the steps through case studies, discussion questions, practice exercises, and classroom application. We worked with school partners to determine a feasible and logical format and schedule for the workshops, as follows:

(1) Module 1: Introduction to DBI; how to administer, score, and graph data using CBM in early writing. This module is implemented right before the beginning of the school year. Teachers learn about the DBI process and receive extensive opportunities to discuss and practice administering and scoring the three CBM tasks.

(2) Module 2: This module is implemented right before the beginning of the school year. Implementing research-based early writing instruction with fidelity (Part 1). Teachers learn about the Simple View of Writing and how it aligns with research-based writing instruction. They also learn the general structure and components of the transcription, text generation, and self-regulation lessons, along with modeling (through videos and live demonstrations) and extensive opportunities to review, practice, and ask questions about the lessons and instructional materials.

(3) Module 3: Implementing research-based early writing instruction with fidelity (Part 2). This module is implemented approximately one month into the school year, after teachers have collected baseline CBM data for their students. Teachers learn to diagnose student strengths and needs, and to use this information to construct WIPs. They are encouraged to bring their students' data and writing samples, and are provided support in developing WIPs for their students during the workshop.

(4) Module 4: Data-based decision-making. This module is implemented about 8 weeks after Module 3, when teachers have sufficient data (a minimum of 8 data points) to make decisions about student progress. Teachers learn to use decision rules to determine whether an instructional change is needed, and to develop hypotheses

about specific changes to make. Teachers bring their graphed CBM data to the workshop, and receive support in using those data to make instructional decisions.

Desimone's (2009) core principles of PD guided our development of the learning modules. Specifically, the modules (a) *focus on content*—the knowledge and skills related to early writing assessment, intervention, and decision-making that teachers need to implement DBI; (b) provide *active learning opportunities over a sustained duration* to ensure sufficient time to learn, practice, and apply new content with support and feedback; (c) emphasize *coherence* of teacher learning, knowledge, beliefs, and existing instructional routines; and (d) provide *collective participation* via collaboration with researchers, peers, and coaches.

Collaborative Support

Collaborative support includes opportunities for teachers to ask questions, share data, be observed implementing DBI, receive feedback, problem solve, and identify interventions in collaboration with a coach and peers. In early phases of the project, we reviewed the coaching literature (McMaster et al., 2017) and derived five Principles of Coaching: it should (a) be teacher oriented, (b) build mastery, (c) be observable and measurable, (d) emphasize alignment with existing curriculum and instructional practices, and (e) support sustained implementation.

After each module, teachers complete a performance assessment to gauge mastery of DBI content, and receive feedback and support as needed via face-to-face (at least bi-weekly) and virtual (email, phone, Skype) coaching. Each coaching session follows a basic protocol that aligns with the five Coaching Principles: (a) celebrate and commiserate (to foster a positive relationship–*teacher oriented*); (b) set objectives (to give the meeting focus and purpose–*build mastery, be observable and measurable*); (c) review DBI steps (to build on previous learning–

build mastery, emphasize alignment); (d) review and discuss student data (to provide support in data-based decision making–*build mastery*); and (e) plan for next steps (*sustain implementation)*.

Theory of Change

Our study was guided by a theory of change based on the assumptions that teachers' knowledge and skills, efficacy, and writing orientation influence their practice; that participation in DBI-TLC will improve teacher knowledge, skills, and self-efficacy, and influence their writing orientation; and that these characteristics will lead to fidelity of DBI implementation, which will ultimately lead to improved student outcomes.

Teachers' Knowledge & Skills, Efficacy, and Writing Orientation Influence Their Practice

Critical to teachers' successful use of DBI are their *knowledge* in selecting appropriate instruction to promote student learning, *skill* in implementing that instruction (Cunningham, Perry, Stanovich, & Stanovich, 2004), and *self-efficacy*, or "confidence that they can perform the actions that lead to student learning" (Graham, Harris, Fink, & MacArthur, 2001, p.178). Further, teachers' *writing orientation* likely will influence how and what they decide to teach (Graham, Harris, MacArthur, & Fink, 2002; Troia, Lin, Cohen, & Monroe, 2011). Strong evidence indicates that explicit instruction is critical to support children who experience difficulties learning to write (Berninger et al., 2008); thus, the extent to which a teacher's practice is oriented toward explicit instruction will likely influence DBI implementation.

Participation in DBI-TLC will Improve Teacher Outcomes and Lead to Fidelity of DBI

By participating in in-depth PD focused on assessment, intervention, and decision-making related to early writing, teachers' knowledge and skills should increase in these areas (cf. Poch, Smith et al., 2018). Further, by receiving ongoing support that fosters teachers' success in applying knowledge and skills to their practice, teacher efficacy and orientation toward explicit instruction should increase, which in turn should lead to implementation of DBI with fidelity (Roehrig et al., 2008).

DBI Implemented with Fidelity will Improve Student Outcomes

When teachers collect, score, and graph data accurately; implement high-quality, research-based instruction with fidelity; and make timely, appropriate instructional decisions based on data, prior research indicates that student outcomes should improve (e.g., Jung et al., 2017; Jung et al., 2018; Stecker et al., 2005).

Purpose and Research Questions

The purpose of this study was to determine whether DBI-TLC shows promise to improve teacher outcomes, DBI implementation, and students' early writing skills. In line with specifications of an IES-funded Goal 2 development and innovation project, our primary aim was to collect preliminary evidence of the promise of DBI-TLC. Thus, we conducted an "underpowered efficacy [study] (e.g., randomized controlled [study] with a small number of classrooms or schools that provide unbiased effect size estimates of practical consequence which can stand as evidence of promise while not statistically significant)" (Institute of Education Sciences, 2012, p. 45). The following research questions guided the study:

(1) What is the effect of DBI-TLC on teachers' DBI knowledge and skills, self-efficacy for writing instruction, and writing orientation?

(2) To what extent do teachers who receive DBI-TLC implement DBI with fidelity?

(3) What is the effect of DBI-TLC on students' early writing outcomes?

Method

Setting and Participants

Sites. This study was conducted in two public school districts in two Midwestern states during the 2015-16 school year. Approval to conduct research in these sites was secured from the human subjects review boards of the principal investigators' (PIs') institutions, as well as the two districts. Site 1 was a large urban district serving 35,717 students; 66% were students of color, 22% English Learners (ELs), 63% received free/reduced lunch (FRL), and 18% received special education. Site 2 was a mid-sized city district serving 17,243 students; 38% were students of color, 5% ELs, 44% received FRL, and 10% received special education.

Teachers. To be eligible to participate, teachers had to provide direct support to elementary students at risk or with disabilities who had specific early writing needs, and had to have at least two years of teaching experience (to avoid overwhelming new teachers). We excluded teachers who had participated in earlier phases of the development project to avoid confounds related to experience with DBI. At each site, the co-PI worked with a district liaison to identify and invite eligible teachers to participate. We aimed to recruit 16 to 20 teachers per site; however, our eligibility criteria limited the number of teachers available within each district. Thus, a total of 20 teachers were recruited. In Site 1, one set of co-teachers shared a classroom and were thus considered one unit for assignment to condition. Nine teachers (8 units) in 7 schools participated in Site 1, and 11 teachers in 8 schools participated in Site 2. Demographic data for teachers by site and condition (described below) are in Table 2. Chi-square and *t*-tests revealed no significant between-group differences on any demographic variable.

Students. Participating teachers nominated students on their caseloads in need of intensive early writing intervention. Our primary target was students in Grades 1-3, but we permitted inclusion of older students (Grades 4-5) who were receiving intervention focused on beginning writing skills and who met screening criteria. Nominated students with parental consent completed two forms each of CBM word dictation and picture word (described in Measures). Word dictation was scored for CLS and picture word was scored for CWS, given these metrics have the most consistent evidence of reliability and criterion-related validity (Lembke, Allen, & Poch, 2015). Research staff scored CBM protocols, entered the data into a spreadsheet, and selected the three students per teacher who scored lowest on both measures.

Based on this process, 25 students from District 1 and 32 students from District 2 were identified as "target students" (total $N = 57$). Complete data (after attrition; described in Results) were available for 53 students. Demographic data for students by site and condition, along with chi-square and *t*-statistics, are presented in Table 3. There were no statistically significant differences between groups on any demographic variables except on race/ethnicity; more students of color were in the control group than in the treatment group.

Study Design and Conditions

We employed a pretest/posttest randomized control group design. Within each district, teachers were stratified by school (or when there was only one teacher in a school, by demographically-similar schools) and assigned randomly to treatment (DBI-TLC) or control.

DBI-TLC. DBI-TLC teachers received all components of DBI-TLC (tools, learning modules, and collaborative supports), and implemented DBI with their target students for 20 weeks, beginning in late August/early September and continuing through March or early April (not counting breaks). We recommended that teachers devote at least three times per week for 20 to 30 min per session for writing intervention; however, teachers ultimately decided when and how much intervention to provide, based on individual students' needs.

Control. Control teachers conducted their usual writing assessment and instruction. They did not receive any DBI tools, training, or support until the study's completion, at which point they were given all materials and invited to attend training on the learning modules. To guard against treatment contamination, DBI-TLC teachers were asked to not share DBI-TLC information or materials with control teachers in their buildings during the study period.

Measures

Corresponding to our theory of change, we administered measures to assess teacher, implementation, and student outcomes.

Teacher outcomes. Teachers completed a pre- and posttest measure of DBI Knowledge and Skills, developed in earlier phases of the project. The test includes 40 multiple-choice questions related to the purpose of DBI (e.g., "*All of the following are assumptions of DBI except…*"); specific DBI steps and components (e.g., "*Which of the following is a critical factor in DBI?*"); writing development and instruction (e.g., *"An applied example of transcription instruction would be…*"); and skills related to using CBM data to make instructional decisions (e.g., "*When a student's trend line falls below the goal line after 8 data points, the teacher should…*"). The score is the number of items correct. For this sample, Cronbach's α coefficients were .58 at pretest and .78 at posttest.

Teacher efficacy in writing was assessed using Graham et al.'s (2001) modified version of Gibson and Dembo's (1984) Teacher Efficacy Scale (TES). The teacher responds to 16 statements such as, "When students' writing performance improves, it is usually because I found better ways of teaching" (1 = strongly disagree; 6 = strongly agree). A factor analysis based on a national sample of teachers indicated two dimensions: personal and general teaching efficacy, with Cronbach's α of .84 and .69, respectively (Graham et al., 2001). Cronbach's α coefficients for this study were .82 for personal efficacy and .83 for general efficacy at pretest, and .87 and .63, respectively, at posttest.

Writing orientation was assessed using Graham, Harris, McArthur et al.'s (2002) Writing Orientation Scale (WOS), a 13-item survey using a 6-point Likert scale, with questions such as, "*A good way to begin writing instruction is to have children copy good models of writing*." The questions align with the sub-scales *Natural Writing*, which reflects an emphasis on incidental and informal learning, *Correct Writing*, which emphasizes spelling, grammar, copying models, and using Standard English; and *Explicit Instruction*, which reflects teaching skills overtly and systematically. Internal consistency for these subscales have been reported as Cronbach's α of .60, .70, and .64, respectively (Graham, Harris, MacArthur et al., 2002). For our sample, Cronbach's α at pre/posttest was .44/.76. (Natural), .78/.62 (Correct), and .61/.65 (Explicit).

Implementation outcomes. Implementation outcomes included DBI-TLC teachers' fidelity of DBI implementation and all teachers' self-report of instructional changes.

Fidelity of DBI. We modified the Accuracy of Implementation Rating Scales (AIRS), originally created by Fuchs, Deno, and Mirkin (1984), to assess DBI fidelity. The AIRS consists of three core DBI components: CBM, Writing Instruction, and Decision-Making (sample forms are provided in the online supplemental materials). Each includes a checklist of critical steps involved in DBI. Fidelity is recorded as the number of steps observed over the total number of applicable steps.

Instructional changes. At posttest, all DBI-TLC and control teachers completed a questionnaire about instructional changes they made for target students during the project period. Specifically, they reported whether they had made any changes, the frequency and types of changes, types of information that prompted the changes, and resources used to inform the changes. Questions were asked in a general way (e.g., "*What types of changes did you make for your target students?*" with a list from which they could select multiple options and write in their own) rather than asking them to recall specific changes made at specific points in time.

Student outcome measures. At pretest, students completed two forms each of CBM word dictation and picture word (mean scores were recorded) and the Test of Early Written Language-3 (TEWL-3; Hresko, Herron, Peak, & Hicks, 2012). At posttest, students completed the same measures plus two CBM story prompts; again, mean scores were recorded. For each CBM task, multiple scoring procedures were used to capture varying levels of complexity of students' writing (from simple counts and number of correct words, to more fine-grained indices that account for spelling development [CLS] and grammar [CWS]).

CBM word dictation. This task is designed to capture word-level writing skills. Word dictation is administered individually for 3 min. The examiner dictates words (based on spelling patterns identified in the Common Core State Standards; National Governors Association, 2010) with one repeat, and students write each word. Scores include WW, WSC, CLS, and CILS. For grades 1-3, alternate-form reliability for 3-min samples has been reported as r = .89 to .95, and criterion validity with quantitative scores from writing samples and standardized tests has ranged from r = .29 to .75 (Lembke et al., 2015; Lembke et al., 2003). Prompts administered weekly for 8 weeks have been shown to be sensitive to growth (Hampton & Lembke, 2016). Test-retest reliability coefficients for this sample were r = .94 to .96 at pretest and r = .92 to .96 at posttest.

CBM picture word. This task is designed to capture sentence-level writing skills. Each group-administered prompt consists of words with a corresponding picture above each word. Students complete a practice item, and then write sentences using the prompts for 3 min. Responses are scored for WW, WSC, CWS, and CIWS. For grades 1-3, alternate-form reliability for 3-min samples has ranged from r = .81 to .91; criterion validity has ranged from r = .50 to .60 (Lembke et al., 2015; McMaster, Du, & Petursdottir, 2009), and weekly prompts have produced reliable slopes that are sensitive to growth within 8 weeks (McMaster et al., 2011). Test-retest reliability coefficients for this sample were r = .67 to .91 at pretest and r = .85 to .90 at posttest.

CBM story prompts. This task is designed to capture passage-level writing skills and can be group administered. Story prompts have simple vocabulary and sentence structure and are designed to reflect experiences that students in U.S. schools can relate to. Each prompt is printed at the top of a page with lines to write on. Students are given 30 seconds to think and 3 min to write. Responses are scored for WW, WSC, CWS, and CIWS. For grades 1-3, alternate-form reliability has ranged from r = .74 to .88; criterion validity has ranged from .50 to .65 (Lembke et al., 2015). Weekly prompts have produced reliable slopes within 8 weeks that are sensitive to growth (McMaster et al., 2011; McMaster et al., 2017). Test-retest reliability coefficients for this sample were r = .71 to .89 (this task was administered at posttest only).

Standardized writing test. The TEWL-3 (Hresko et al., 2012) is designed to assess writing for children ages 4-0 through 11-11. Students completed the Basic and Contextual Writing subtests at pre- and posttest. Basic Writing assesses students' knowledge of purposes of writing, letter formation, spelling, capitalization and punctuation, language functions (e.g., nouns, verbs), and sentence combining. The test is untimed, and responses are scored 0, 1, or 2. For Contextual Writing, students construct a story about a picture prompt within a 30-min limit. Responses are scored from 0 to 3 on dimensions including story structure, cohesion, and ideation. Raw scores were used for analysis. Hresko et al. (2012) reported alternate-form reliability as $r \geq .90$ and criterion validity as $r \geq .70$ with other standardized writing tests.

Procedures

Coach training and support. Five graduate research assistants (GRAs; all advanced doctoral students in school psychology or special education) and four project coordinators (PCs; all with advanced degrees in special education, school psychology, or administration) were each assigned as coaches to one or two DBI-TLC teachers. Coaches were trained by the co-PIs via a two-hour "Coaches' Institute," in which they learned, through modeling and guided practice, (a) the definition and Principles of Coaching, (b) positive coaching behaviors, (c) how to implement a sequence of coaching protocols, and (d) how to differentiate coaching for individual teachers. Throughout the study, coaches attended weekly meetings with the PIs to engage in problem solving needed to facilitate teachers' successful DBI implementation.

Test administration and scoring training. The GRAs and PCs were trained to administer and score all CBM tasks by the two co-PIs. All examiners had extensive experience administering and scoring CBM as part of prior studies. Prior to testing in the schools, examiners had to demonstrate 95% accuracy of administration on the CBM-AIRS. CBM tasks were scored immediately after data were collected at pretest and again at posttest. At each time-point, each examiner had to reach 85% inter-scorer agreement with the PCs on two student samples of each CBM task (using point-by-point agreement, in which each scored item was compared and counted as an agreement or disagreement; agreement was calculated as the number of agreements divided by agreements plus disagreements and multiplied by 100). Once each scorer reached criterion, additional protocols were assigned. PCs randomly checked 30% of additional scored protocols to confirm ongoing agreement. Final agreement among scorers for almost all tasks and scoring procedures was high (99% to 100%); agreement was somewhat lower for picture-word IWS (94%), likely because there are fewer instances of this score, and word sequences are more subjective to score than words written or spelled correctly.

PCs trained the GRAs to administer the TEWL-3 subtests and checked each examiner's accuracy using a checklist of administration procedures. Administration accuracy ranged from 91% to 100%; all errors were corrected immediately. PCs also trained GRAs to score the Basic Writing subtest, which is scored during administration. PCs checked 10% of each scorer's test forms to ensure that GRAs had scored according to the standardized directions. If a scorer did not reach 90% agreement, the test was rescored when possible. Contextual Writing was scored only by the PCs. Average interrater agreement was 96% at pretest and 98% at posttest.

Pre- and post-testing. Teachers completed pretests of DBI Knowledge and Skills, and the TES and WOS via Qualtrics (an online survey system) before any other study activities began, and completed posttests following 20 weeks of DBI implementation. At posttest, the Instructional Changes questionnaire was added to Qualtrics for teachers in both conditions.

After we identified target students based on CBM word dictation and picture word scores, the students completed the two TEWL-3 subtests. Then, after 20 weeks of DBI implementation, all target students completed the CBM word dictation, picture word and story prompts and TEWL-3 subtests. Examiners administered posttests to students in classes of teachers with whom they had *not* been in frequent contact (e.g., through Coaching or pretesting) during the study, and were blind to those students' study conditions, in order to reduce the threat of testing bias.

DBI-TLC implementation. Before school began in August, DBI-TLC teachers attended workshops for Modules 1 and 2 (overview of DBI, CBM, and writing intervention), delivered across two full days. Teachers were given access to the DBI tools via Google Drive and also in printed form in binders. At the end of Module 1, coaches met with assigned teachers to exchange contact information and build rapport. At the end of Module 2, coaches assessed teachers' CBM scoring reliability and provided feedback and additional scoring guidance as needed.

After school began in both sites (mid- to late-August), research staff worked with teachers to screen and identify target students (as described in Participants). After screening and pre-testing, coaches supported their teachers in selecting a CBM task and observed teachers' CBM administration. Then, in mid-September, teachers attended Module 3 to learn to develop WIPs. At the end of Module 3, coaches supported teachers in setting long-term goals, justifying selection of specific mini-lessons, and creating WIPs.

After Module 3, teachers implemented DBI with target students, including monitoring progress on a weekly basis and implementing their WIPs as designed. Coaches held bi-weekly face-to-face meetings with teachers, using an ongoing coaching protocol that included observing CBM administration and checking scoring reliability, ensuring the teacher was graphing data, observing writing instruction, and supporting the teacher as needed. Virtual coaching was provided as needed in weeks when they did not meet face-to-face.

The Module 4 workshop was held once teachers had the opportunity to collect eight CBM data points. Teachers were asked to bring graphed student data to the workshop, which focused on using data to make instructional decisions. After Module 4, bi-weekly (every other week) coaching continued and included supporting teachers' data-based decisions as needed.

DBI and coaching logs. All DBI-TLC teachers maintained logs of time spent conducting DBI activities per target student per week, including number of minutes spent preparing, administering, and scoring CBM; graphing and examining CBM data; developing hypotheses; creating intervention materials and implementing intervention; and conducting the entire DBI process. Overall, teachers spent, on average, 3.25 hours per child per week in September, with considerable time devoted to preparing CBM and intervention materials, scoring CBM data, and delivering intervention. In the remaining months, teachers reported spending around 2 hours or slightly less per child per week, with most of this time spent on intervention activities. Coaches also logged time spent in each bi-weekly face-to-face coaching session with each teacher. Across the study period, coaching sessions averaged 40 min per teacher ($SD = 16.10$).

Fidelity. Fidelity observations were completed for each DBI component (CBM, writing instruction, and decision making), as well as for each TLC (tools, learning, and collaborative support) component. Fidelity of DBI was considered to be an implementation outcome for this study, while fidelity of the TLC components was established to be sure that the PD intervention was implemented as intended.

DBI fidelity. The two co-PIs and PCs at each site observed teachers' CBM administration and writing instruction using the AIRS-CBM and Writing Instruction checklists. We observed each teacher at least once, early in DBI implementation (September to November) and, when possible, a second time (January or February). In addition, coaches re-scored at least 10% of teachers' scored CBM protocols (including each type of task the teacher was using) each month and calculated interrater agreement. Teachers received feedback to maximize fidelity.

At the end of the study, we collected teachers' graphs and Decision Logs to determine fidelity of decision making. The two co-PIs and one GRA examined each graph to determine (a) the number of opportunities for a decision to be made (every 6 to 8 data points), and (b) what the decision should be at each opportunity (raise goal, keep as-is, or change instruction). These criteria were compared with the actual timing and decision made by the teacher, as recorded in the Decision Log. The number of timely and appropriate decisions were each summed and divided by the number of opportunities to make a decision. We also recorded whether teachers noted the type and rationale for each decision in the Decision Log.

Fidelity of tools access (Fidelity-T). A critical component of DBI implementation is teachers' access of appropriate DBI tools. Thus, we developed Fidelity-T, which consisted of a questioning guide that required an observer to direct the teacher to "show me" where or how to find specific materials. During the first coaching visit (after Module 1), coaches used Fidelity-T to determine whether teachers could access all of the DBI tools via the Google Drive. The percentage of items teachers accessed independently ranged from 81% to 100% (mean = 94%).

Fidelity of learning modules (Fidelity-L). For each learning module, we identified key components that must be addressed through (a) explanation, (b) modeling, (c) guided practice with feedback, and (d) practice and application. At each workshop, a research staff member who was not presenting information to teachers used Fidelity-L to note whether each component was observed. The percentage of components observed ranged from 94% to 100%.

Fidelity of coaching (Fidelity-C). We also identified key components that should be included in each coaching session. While specific coaching activities varied based on the DBI steps the teacher was currently implementing, each session should have included: (a) a brief rapport-building time, (b) review of objectives, (c) review of DBI steps, (d) discussion of student data, and (e) planning for next steps. PIs and PCs observed each coach in fall (October-November) and winter (January-March). They noted whether each component was observed or not observed on the Fidelity-C checklist. Fidelity-C ranged from 71% to 100% (mean = 94%).

Control observations. PIs and PCs observed control classrooms to describe writing instruction, taking detailed field notes about what was occurring. In general, students were taught in small groups, for 20 to 40 min. Writing activities included writing sentences about a theme (e.g., what would you eat on Thanksgiving), writing sentences using information (e.g., who, what, where) based on a model sentence, and making inferences using evidence from a picture (e.g., I infer ___because the picture says ___). Most teachers provided modeling, reinforced students' writing (e.g., with praise or stickers), and sometimes gave individualized feedback. In a few cases, teachers addressed some basic writing skills—typically incidentally (e.g., the teacher reminded students to use capital letters and punctuation in sentences).

Data Analysis

Pre-treatment analyses. For teacher and student outcomes, we first examined whether there were pretest differences between groups. Given that most correlations among teacher variables ranged from weak to moderate (*r*s = -.55 to .65), we conducted a *t*-test for Knowledge and Skills, and separate multivariate analyses of variance (MANOVAs) for the TES and WOS. Given moderate to strong correlations among student variables (*r* = .35 to .96), we conducted a MANOVA for all student pretest variables.

Post-treatment analyses. Differences in posttest mean scores for teacher outcomes were analyzed using *t*-tests for Knowledge and Skills and separate MANOVAs for the TES and WOS. Student outcomes comparing the DBI-TLC and control groups were analyzed using hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002), given that students were nested within teachers. All multilevel analyses were conducted using HLM software ver. 6.0 (Raudenbush, Bryk, & Congdon, 2004). Prior to the HLM analyses, we calculated the intra-class correlation (ICC) to determine the amount of variance in posttest CBM and TEWL-3 scores between teachers to confirm the need for HLM. This analysis was a baseline model with no predictors in Level 1 or Level 2. The ICC results indicated that a substantial amount of variance in posttest CBM scores (35.7% - 50.6%) and TEWL-3 scores (53.1% - 57.7%) was between teachers. Given the large ICCs (Hox, 2010), we conducted HLM analyses with pretest scores as covariates at Level 1 and condition (DBI-TLC or control) at Level 2. The final model is shown below.

$$Y_{ij} = \beta_{0j} + \beta_{1j} Pretest\ scores_{ij} + r_{ij} \qquad \text{(Level 1)}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Condition + u_{0j} \qquad \text{(Level 2)}$$

$$\beta_{1j} = \gamma_{10}$$

Given that the preliminary analysis indicated the effect for the pretest covariate, β_{1j}, did not vary across teachers at Level 2, the effect of the pretest covariate was constrained to be identical (fixed) for each Level 2 unit, except for TEWL-3 Basic Writing. Therefore, all models except for Basic Writing did not include random variance (u_{1j}) in the final two-level models. In addition, the difference between estimated means of DBI-TLC and control groups, adjusted for pretest scores, was captured by γ_{00} (adjusted mean of control group) and γ_{01} (treatment effect).

We calculated effect sizes using Hedges' *g* (Hedges, 1981) for teacher outcomes. For student outcomes, effect sizes between the DBI-TLC and control group were calculated based on HLM results, using the within and between cluster variances to compute a standard deviation (Reis et al., 2008). That is, posttest mean differences were divided by the square root of the sum of within (Level 1) and between (intercept) variances in the unconditional models.

Fidelity of DBI implementation and instructional changes. Because DBI fidelity data were collected only for DBI-TLC teachers and not control teachers, these data were summarized descriptively (means, *SD*s, and ranges). Data from the instructional changes questionnaire (collected from both DBI-TLC and control teachers) were also summarized descriptively; chi-square was computed to determine whether there was a significant difference in the number of DBI-TLC versus control teachers who reported making instructional changes.

Results

In this section, we present results from analyses of the effects of DBI on teacher outcomes (knowledge and skills, self-efficacy, and writing orientation), implementation outcomes (DBI fidelity, teachers' instructional changes) and student early writing outcomes.

Attrition

A few teachers and students were not included in final analyses due to attrition or missing data. Specifically, one control teacher was excluded from final analyses for the WOS and another control teacher from both TES and WOS analyses because they did not complete those measures. Out of 57 students, three (one DBI-TLC and two control) moved before post-testing. In addition, three control students were excluded from final CBM analyses, because two were missing pre- or posttest data, and one was an outlier with scores 4 *SD* above the mean. Two students were excluded from final TEWL-3 analyses because they did not complete all posttests.

Overall attrition rates were 5% and 10% for the TES and WOS, respectively, 10.5% for CBM, and 8.7% for TEWL-3. For all measures, data were missing at random. To determine whether results were sensitive to missing data, separate analyses were conducted for teacher and student pretest data with and without the missing participants. Independent *t*-tests for each pretest measure confirmed that results with and without attrition were not different for both teacher and student outcomes (all *p*-values were above .89).

Teacher Outcomes

Our first research question was whether teachers' participation in DBI-TLC would affect their DBI knowledge and skills, self-efficacy for writing instruction, and writing orientation. First, we confirmed whether basic assumptions for normality and homoscedasticity (equal error variance) were met. Skewness and kurtosis for all teacher variables ranged from -1.19 to .35 and from -1.38 to 2.79, respectively. To meet the normality assumption, the recommended range of the absolute value is within 3 for skewness and 10 for kurtosis (Kline, 2015). Levene's test revealed that the error variance of all teacher variables was not statistically different between groups. These results showed that the data met assumptions for further analyses.

Pre- and posttest means, *SD*s, and ranges on all teacher outcome variables are in Table 4. For pretest, the main effect for condition was not significant for Knowledge and Skills ($t = .07$, $p = .946$), TES (Wilks' Lambda = .908, $F = .86$, $p = .44$), or WOS (Wilks' Lambda = .912, $F = .52$, $p = .678$), confirming that DBI-TLC and control teachers' pretest scores were similar.

At posttest, however, there was a statistically significant main effect for Knowledge and Skills ($t = -6.79$, $p < .001$); DBI-TLC teachers outperformed controls ($g = 2.88$). There was also a significant main effect for WOS (Wilks' Lambda = .510, $F = 4.48$, $p = .02$). Follow-up univariate ANOVAs using a Bonferroni correction (given six comparisons, we adjusted the *p*-value to .008) revealed a significant mean difference ($p = .003$): DBI-TLC teachers reported a more Explicit orientation ($g = 1.63$), and controls reported a more Natural orientation ($g = -.97$).On the TES, no reliable mean group differences were found (Wilks' Lambda = .844, F = 1.48, $p = .26$); effect sizes for General and Personal Efficacy were $g = -.08$ and .78, respectively.

Implementation Outcomes

Fidelity of DBI implementation. Table 5 shows results of fidelity observations for each DBI component, broken down by the major elements on each AIRS. For CBM, overall fidelity was 83.5% (range = 50% to 100%). Some teachers scored lower on giving directions and demonstrating procedures because they tended to reduce or omit these elements after multiple administrations. On average, teachers scored CBM samples with 95% reliability on the 10% of student samples collected each month. For most tasks and scoring procedures, reliability was consistently high across teachers over time (mode = 100%), with occasional dips (usually on letter or word sequences, which are more subjective than the other scoring indices).

For Writing Instruction, overall fidelity was 79.1% (range = 50% to 93%). Lower scores were typically obtained on the introduction and wrap-up components (e.g., some teachers did not present learning objectives at the beginning of the lesson or include an explicit wrap-up). Fidelity of each activity (in which explicit instruction using modeling, guided practice, and corrective feedback were documented) was generally high. For Decision-Making, overall fidelity was 52.1% (range = 0% to 100%). For many teachers, Decision-Making fidelity was "all or nothing" (evidenced by the 0% to 100% range for each element)—if they did not make timely decisions, decisions were also not likely to be appropriate or indicated on the student graph.

Instructional changes. Although Decision-Making fidelity was low, teachers reported that they did make decisions. On the instructional changes questionnaire, significantly more DBI-TLC teachers ($n = 10$) reported implementing changes during the study period than did control teachers ($n = 1$; $\chi^2 = 12.74$, $p = .001$). Most teachers reported making changes less than once per month. Most DBI-TLC teachers (91%) reported making instructional changes when data indicated a student wasn't on track to meet the goal. The one control teacher who reported making any changes reported using intuition as the basis for those changes. Specific types of instructional changes that teachers reported are summarized in Table 6. When asked about the sources of information teachers used to inform changes, all teachers reported using graphed progress monitoring data. In addition, nine DBI-TLC teachers and the control teacher reported using student writing samples and their own professional judgment, and four DBI-TLC teachers reported using peer or professional recommendations to determine instructional changes.

Student Outcomes

Our third research question was whether teachers' participation in DBI-TLC would affect students' posttest outcomes on CBM and the TEWL-3 (Hresko et al., 2012). First, descriptive analyses were conducted for all pre- and posttest scores of CBM and TEWL-3 (see Table 7). Tests of assumptions indicated that the student outcome data met normality and homoscedasticity assumptions for further analyses.

Pretest results. For CBM at pretest, the MANOVA revealed that the main effect for condition was not significant (Wilks' Lambda = .842, $F = .99$, $p = .460$). However, follow-up univariate tests indicated significant effects of condition for word dictation WW ($F = 4.83$, $p = .033$) and CLS ($F = 4.90$, $p = .032$); and for picture word WW ($F = 4.13$, $p = .048$). Because of these possible differences, pretests were used as covariates in posttest CBM analyses. For TEWL-3, the MANOVA indicated that the main effect of condition was not significant (Wilks' Lambda = .926, $F = 1.89$, $p = .163$). Follow-up tests revealed no between-groups difference for Contextual Writing ($F = 2.56$, $p = .116$), but the effect for Basic Writing approached significance ($F = 3.85$, $p = .055$). Thus, pretest was used as a covariate in posttest TEWL-3 analyses.

Posttest results. Next, we conducted separate HLM analyses to examine differences between groups on students' posttest scores on CBM tasks and on TEWL-3 raw scores. Pretest scores were used as covariates to improve precision given significant group differences on the pretest CBM word dictation and picture word and near-significant group differences on the pretest TEWL-3 Basic Writing subtest. Given moderately high correlations with posttest scores, WW was used as the covariate for the analysis of posttest word dictation scores, and CWS was used for picture word and story prompt. Basic Writing was used as the covariate for the analyses of posttest TEWL-3 scores. Posttest mean differences between DBI-TLC and control students on all CBM tasks and scoring procedures are shown in Tables 8-10, and posttest mean differences between groups on the TEWL-3 subtests are shown in Table 11. To guard against Type I error, a Bonferroni correction was used. Given three CBM tasks with four scoring indices, there were 12 comparisons, so *p* was set at .004. For the two TEWL-3 subtests, *p* was set at .03.

Regarding CBM (Tables 8-10), HLM results indicated that students' pretest CBM scores were significantly associated with posttest CBM scores. Specifically, pooled within-teacher regression coefficients for the Level-1 covariate (γ_{10}) ranged from .29 to 2.77 for word dictation, from .75 to 1.1 for picture word, and from .41 to .64 for story prompt. Effects of pretest CBM scores were significant across almost all CBM tasks and scoring procedures (most ps $< .004$). Group differences were not statistically significant; however, the DBI-TLC group's adjusted posttest mean scores were consistently higher than the control group's scores across CBM tasks and scoring procedures, with small to moderate effect sizes (ES = .23-.40), except for story prompt WSC, CWS, and CIWS. For random effects, significant differences in posttest CBM scores existed across teachers at Level-2 (τ_{00}), indicating that, overall, there was significant variance in mean posttest CBM scores across teachers, suggesting a potential teacher effect.

For TEWL-3 (Table 11), results indicated that students' pretest Basic Writing scores were significantly associated with posttest scores. Pooled within-teacher regression coefficients for the Level-1 covariate (γ_{10}) were .94 for Basic Writing and .85 for Contextual Writing. These pretest effects were significant across both TEWL-3 subtests (all ps $< .001$). Regarding group mean differences, there were no statistically significant treatment effects (γ_{01}) for either Basic Writing ($p = .929$) or Contextual Writing ($p = .515$), with small effect sizes (ES = .015 for Basic Writing, ES = .132 for Contextual Writing). For random effects, posttest Basic and Contextual

Writing scores did not significantly vary across teachers (*p*s = .053 and .373, respectively). In contrast to CBM results, there was no clear teacher effect on posttest TEWL-3 scores.

Discussion

The purpose of this study was to examine the promise of a PD system developed to support teachers' use of DBI to improve outcomes for students with intensive early writing needs. Below, we discuss findings in relation to our three research questions, study limitations, and implications for further research and practice.

Effects of DBI-TLC on Teachers' Knowledge and Skills, Efficacy, and Writing Orientation

Our theory of change specified that teachers' participation in DBI-TLC would lead to improved teacher outcomes. To some degree, this happened. With respect to teacher knowledge and skills, DBI-TLC teachers significantly outperformed control teachers at posttest with a large effect size ($g = 2.88$), indicating that teachers gained critical knowledge and skills about DBI in early writing. This finding was expected given our attempts to model the learning modules and collaborative supports on Desimone's (2009) core principles of PD, particularly by placing strong emphasis on *content* (knowledge and skills teachers need to implement DBI in early writing) and *active opportunities to learn* and practice this content over an extended duration.

We were less successful in changing teacher efficacy. The lack of statistically significant differences between DBI-TLC and control teachers may be because, in general, teachers reported relatively high levels of self-efficacy. This finding is consistent with previous research (e.g., Graham et al., 2001; Troia et al., 2011; Ritchey, Coker, & Jackson, 2015), and might reflect that teachers who are motivated to participate in research already have a sense of their capacity to effect improved student outcomes. It is also likely that self-efficacy takes a longer time to change than we were able to observe in a relatively short period. We did observe that, despite no reliable differences, DBI-TLC teachers appeared somewhat stronger on personal efficacy (their belief in their own ability to teach writing effectively) at posttest (g = .78).

In terms of writing orientation, DBI-TLC teachers reported a more explicit orientation at posttest compared to control teachers (g = 1.63), whereas control teachers reported a more natural writing orientation (g = -.97). This finding is plausible given that DBI-TLC teachers learned (from both theoretical and empirical perspectives) that students who experience difficulties with writing are likely to benefit from explicit instruction targeting basic early writing skills in transcription and text generation (e.g., Berninger et al., 2008).

Overall, given that DBI-TLC teachers reported a more explicit writing orientation (which should influence how and what they decide to teach; Graham, Harris, MacArthur et al., 2002), and that they gained critical knowledge and skills related to specific DBI practices (which should serve to improve student outcomes; Cunningham et al., 2004), our findings provide promising evidence that DBI-TLC prepared teachers for successful DBI implementation. Further research is needed to determine the extent to which DBI-TLC affects the self-efficacy of teachers, particularly with samples with a broader range of initial self-efficacy.

Extent to Which DBI-TLC Teachers Implemented DBI with Fidelity

Our theory of change specified the assumption that improved teacher outcomes should lead to implementation of DBI with fidelity. We did not have a large enough sample to statistically test this relation; however, we did examine DBI-TLC teachers' fidelity of each DBI component. This examination revealed that fidelity varied considerably across components, suggesting that the promising teacher outcomes related to knowledge, skills, and writing orientation did not lead to uniformly strong implementation, and that some components were easier to implement with fidelity than were others.

For the CBM component, on average, teachers' implementation fidelity was relatively high (84%), with most instances of lower fidelity reflecting shortcuts teachers took given the repetitive nature of the task. These shortcuts actually seemed reasonable and possibly a better approach to CBM administration, because repeating unnecessary directions could waste valuable instructional time and frustrate students. Thus, our conclusion was that the AIRS-CBM tool should be revised to reflect a shortened (yet still accurate) version of a fidelity assessment that reflects a more realistic expectation for ongoing administration.

Fidelity of writing instruction was somewhat lower (79% on average), indicating that teachers' implementation of writing instruction did not always match our expectations. Writing instruction fidelity tended to be higher on implementation of instructional activities and lower on the introduction and wrap-up components of the lessons, underscoring the need for further evidence regarding which components are most important for improved student outcomes (and thus should be given more weight on the fidelity tool). At the same time, we see a need to improve teachers' writing instruction fidelity, which could be addressed in learning modules (e.g., through more extensive modeling and practice of instructional activities) and through coaching (e.g., through more frequent observations, specific feedback, and classroom modeling when needed). We also believe that more consideration should be given to the extent to which *flexibility* of implementation can be tolerated, particularly in the context of individualized instruction (cf. Harn, Parisi, & Stoolmiller, 2013; Johnson & McMaster, 2013). Further research is needed to determine which aspects of writing instruction are essential for implementation, and which components can be modified to suit specific student needs and instructional contexts.

Fidelity of decision making was lowest of the three DBI components (52%, on average), with few teachers meeting our "timely and appropriate" criteria (a decision was deemed "timely" if it was made within 6 to 8 weeks of data collection and "appropriate" if the trend of the data compared to the goal line was correctly interpreted based on prescribed decision rules). This finding revealed another area where teacher learning opportunities and ongoing supports need to be strengthened, leading us to consider additional instruction and practice within the DBI learning modules as well as more targeted coaching activities that emphasize the decision-making process. At the same time, an important question is: How critical is our definition of "timely and appropriate?" Findings from the instructional changes questionnaire revealed that DBI-TLC teachers reported making many more instructional changes than did controls, that these decisions were based on student data, rather than intuition, and that the decisions varied from changes to setting or format, instructional delivery, content, and focus. More research is needed to determine the importance of decisions that conform to the "timely and appropriate" definition, or whether it is sufficient that teachers are simply being responsive by looking at data and making changes, even if those changes do not adhere perfectly to prescribed decision rules.

Effect of DBI-TLC on Students' Early Writing Outcomes

Finally, our theory of change specified that teachers' implementation of DBI with fidelity would lead to improved student outcomes. DBI-TLC teachers' students did not significantly outperform controls at posttest; however, patterns in CBM data, with small to moderate effect sizes (ES = .23-.40) suggest promise, particularly given that the study was underpowered due to the developmental focus of this work, and that the primary focus of the study was on teacher outcomes. Also, students were performing, on average, around the 27th percentile on the TEWL-3 Overall Writing (39th percentile for Basic Writing, and 18th percentile for Contextual Writing) at pre-test, indicating that target students had significant needs in early writing. Improving outcomes for children with intensive needs can be particularly challenging, but is of critical importance (Fuchs et al., 2010; Fuchs, Fuchs & Compton, 2012).

These findings tentatively support previous work indicating the promise of teachers' use of DBI to improve students' outcomes (Jung et al., 2018; Stecker et al., 2005) including in writing (Jung et al., 2017), particularly when ongoing supports are in place (Stecker et al., 2005). However, these tentatively positive results are related only to CBM outcomes, which may be considered more proximal to the intervention, given that DBI includes ongoing progress monitoring using CBM in writing. Further, the positive patterns observed in CBM data were particular to word dictation and picture word prompts. Based on teachers' WIPs that they submitted to their Coaches, Coaches' classroom observations and discussions with teachers, and notes gathered during writing instruction fidelity observations, we noted that all teachers focused instruction on word-level transcription and sentence-level text generation activities, rather than passage-level text generation and self-regulation activities (i.e., no teacher used the *Powerful Writing Strategies* text with any target student). This instructional focus makes sense, given that our target sample was students with intensive early writing needs. It is possible that word- and sentence-level CBM tasks were more sensitive to growth in these areas in a relatively brief time period. Thus far, effects on more distal writing outcomes remain unclear. It may be that the intervention was simply not sufficient to affect more generalized writing proficiency for students with intensive needs within 20 weeks, but it could also be that the TEWL-3 was not sensitive to any gains made in this time frame.

The implementation outcomes described above also suggest that we need to find ways to improve teacher fidelity; we are optimistic that such improvements are possible through enhanced learning opportunities and ongoing supports. Also, we observed a possible teacher effect related to CBM outcomes, suggesting the need to further explore specific teacher characteristics and contextual factors that might contribute to improved outcomes. Such characteristics include knowledge, skills, and self-efficacy, but might also include other characteristics related to teachers' educational backgrounds and experiences, or contextual variables such as school climate and leadership.

Limitations and Directions for Research

Our findings must be interpreted in the context of the following limitations, which also have implications for further research. First, our sample size was small, which was deliberate given the developmental nature of this project, but limited power to find statistically significant differences, and also limited the generalizability of findings. In addition, there were possible pre-treatment differences between students whose teachers received DBI-TLC or served as controls, and there were more students of color in the control group than in the treatment group. Whereas we attempted to control for pre-treatment differences, it would be ideal to start out with more balanced groups. Also, students with a wide range of disabilities were included in the sample, and the distribution of disability types varied by site (e.g., Site 1 included more students with autism and "students needing alternative programming"—a special category used in Site 1 that encompasses a range of academic and behavioral needs, including students who would typically be identified with learning disabilities—whereas Site 2 included more students identified with learning disabilities and language impairments). The sample size was not sufficient to examine possible moderating effects of disabilities. Research is needed with larger and more representative samples of teachers and students with intensive writing needs.

A second limitation is the multicomponent nature of the DBI-TLC package. Without a specific component analysis, it is not clear which DBI-TLC features are necessary and sufficient to lead to improved student outcomes in writing. For example, completing weekly CBM probes in writing might have led to a practice effect, or teachers' implementation of research-based early writing interventions might have been sufficient to improve student outcomes. Further research could address this question by comparing individual components to the total package (e.g., a study could include a comparison of some combination of the following conditions: teachers' use of CBM in writing, teachers' implementation of early writing intervention, teachers' use of CBM and intervention in a DBI framework, and teachers' use of DBI with TLC). However, such research would be resource intensive. Given findings from the DBI literature in general (that it is more effective when teachers receive support; Stecker et al., 2005), and evidence indicating that, despite the existence of research-based instruction in early writing (e.g., McMaster et al., 2018), teachers feel under-prepared to implement such instruction (Cutler & Graham, 2008), the costs and benefits of doing such an analysis should be carefully considered.

A third limitation is that, despite fairly intensive PD and coaching, teachers' DBI fidelity varied considerably, and was particularly low for the decision-making component. It is not clear whether student outcomes would have been stronger if fidelity had also been stronger. Further research is needed to address the following issues: (1) Do measures of DBI fidelity sufficiently capture teachers' DBI implementation? In current work, we are devoting considerable effort to enhancing these tools as well as considering other ways to capture teachers' DBI implementation, including the quality of implementation and their capacity to intensify intervention based on student responsiveness to intervention. (2) Can we improve DBI fidelity by improving DBI learning modules and coaching? Although we believe we provided high-quality PD and support, teacher learning takes time (cf. Poch, McMaster et al., 2018), and the learning modules and coaching could be further strengthened to allow teachers to become more fluent in their implementation of each DBI component. (3) How does fidelity change over time, particularly in the context of ongoing coaching? (4) To what extent do specific teacher characteristics, including DBI knowledge, skills, and self-efficacy, influence fidelity? And, (5) to what extent is a balance between fidelity and flexibility to adapt to specific classroom contexts and student needs most beneficial for improving student outcomes (cf. Harn et al., 2013; Johnson & McMaster 2013)?

A final limitation is that this study was conducted over a relatively short time period (20 weeks). Given that our aim is to change student outcomes via changes in teachers' practice, we recognize that it is ambitious to expect large effects on student measures in a short period of time as a result of PD and support for teachers. Indeed, few PD studies have revealed strong student effects (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007); yet, we do not believe that modest findings should cause us to abandon such efforts. More longitudinal research is needed to answer questions about the long-term effects of teachers' use of innovations such as DBI, as well as what types and amounts of support are needed to ensure teachers' sustained use of these innovations in ways that lead to long-term improvements in students' learning.

Implications for Practice

Whereas results thus far should be considered preliminary, they are promising, and support the following implications for practice. First, teachers who need to individualize instruction for students with intensive early writing needs should be encouraged to use the DBI framework, which has general support in improving student outcomes in reading, mathematics, spelling, and writing (Jung et al., 2018; Stecker et al., 2005). Second, teachers interested in using DBI in early writing might increase their knowledge and skills for implementing DBI under the

PD and support conditions that were implemented in this study (content-oriented learning modules with embedded coaching). Third, implementing DBI with PD and support may improve students' early writing outcomes, at least on proximal measures. In doing so, it is critical to implement research-based early writing intervention with fidelity, conduct ongoing progress monitoring, and make timely and appropriate instructional decisions. Such systematic, hypothesis- and data-driven instruction is likely key to improving outcomes for students with the most intensive early writing needs.

References

Ardoin, S. P., Christ, T. J., Morena, L. S., Cormier, D. C., & Klingbeil, D. A. (2013). A systematic review and summarization of the recommendations and research surrounding curriculum-based measurement of oral reading fluency (CBM-R) decision rules. *Journal of School Psychology*, *51*, 1-18. doi: 10.1016/j.jsp.2012.09.004

Baker, S. K., Chard, D. J., Ketterlin-Geller, L. R., Apichatabutra, C., & Doabler, C. (2009). Teaching writing to at-risk students: The quality of evidence for self-regulated strategy development. *Exceptional Children*, *75*, 303-318.

Berninger, V. & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice. In H.L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of research on learning disabilities* (pp. 345-363). New York: Guilford.

Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Writing problems in developmental dyslexia: Under-recognized and under-treated. *Journal of School Psychology, 46*, 1-21. doi: 10.1016/j.jsp.2006.11.008

Biancarosa, G., & Snow, C. (2004). *Reading next—A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York*. Washington, DC: Alliance for Excellence in Education.

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Stanovich, P. J. (2004). Disciplinary knowledge of K-3 teachers and their knowledge calibrations in the domain of early literacy. *Annals of Dyslexia, 54*, 139-167. doi: 10.1007/s11881-004-0007-y

Cutler, L., & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of educational Psychology, 100*, 907-919. doi: 10.1037/a0012656

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*, 219-232. doi: 10.1177/001440298505200303

Deno, S. L., & Mirkin, P. K. (1977). *Data-based program modification: A manual*. Reston VA: Council for Exceptional Children.

Desimone, L. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*, 181–199. doi: 10.3102/0013189X08331140

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2012). Smart RTI: A next-generation approach to multilevel prevention. *Exceptional Children*, *78*, 263-279. doi: 10.1177/001440291207800301

Fuchs, D., Fuchs, L. S., & Stecker, P. (2010). The "Blurring" of special education in a new continuum of general education placements and services. *Exceptional Children, 76*, 301-323. doi: 10.1177/001440291007600304

Fuchs, L. S., & Fuchs, D. (2007). Using CBM for progress monitoring in written expression and spelling. Retrieved on 8/5/2019 from https://files.eric.ed.gov/fulltext/ED519251.pdf.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. (1984). The effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement, and student awareness of learning. *American Educational Research Journal*, *21*, 449-460.

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, *76*, 569-582. doi:10.1037/0022-0663.76.4.569

Graham, S., & Hebert, M. (2010). Writing to read: Evidence for how writing can improve reading. New York: Carnegie Corporation.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students.

Journal of Educational Psychology, 99, 445-476. doi: 10.1037/0022-0663.99.3.445

Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write? Treatment of handwriting problems in beginning writers. *Journal of Educational Psychology*, *92*, 620-633.

Graham, S., Harris, K. R., & Fink-Chorzempa, B. (2002). Contributions of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology, 94*, 669-686. doi: 10.1037/0022-0663.94.4.669

Graham, S., Harris, K. R., MacArthur, C., & Fink, B. (2002). Primary grade teachers' theoretical orientations concerning writing instruction: Construct validation and a nationwide survey. *Contemporary Educational Psychology*, *27*, 147-166.

Graham, S., Harris, K., Fink, B., & MacArthur, C. A. (2001). Teacher efficacy in writing: A construct validation with primary grade teachers. *Scientific Studies of Reading, 5*, 177-202. doi: 10.1207/S1532799Xssr0502_3

Hampton, D. D., & Lembke, E. S. (2016). Examining the Technical Adequacy of Progress Monitoring Using Early Writing Curriculum-Based Measures. *Reading & Writing Quarterly*, *32*, 336-352. doi: 10.1080/10573569.2014.973984

Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing fidelity with flexibility and fit: What do

Harris, K. R., Graham, S., Mason, L. H., & Friedlander, B. (2008). *Powerful writing strategies for all students*. Baltimore, MD: Brookes.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107-128. doi: 10.3102/10769986006002107

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd Ed.). New York:

Routledge.

Hresko, W. P., Herron, S. R., Peak, P. R., & Hicks, D. L. (2012). *Test of Early Written Language-3*. Austin, TX: ProEd.

Institute of Education Sciences (2012). Request for applications: Special Education Research Grants, CFDA Number: 84.324A. Washington, DC: U.S. Department of Education.

Johnson, L., & McMaster, K. L. (2013). Adapting research-based practices with fidelity: Flexibility by design. In B. G. Cook, M. Tankersley, & T. J. Landrum (Eds.), *Advances in learning and behavioral disabilities* (Vol. 26). Bingley, UK: Emerald.

Jung, P., McMaster, K. L., & delMas, R. (2017). Effects of early writing intervention delivered within a data-based instruction framework. *Exceptional Children, 83*, 281-297. doi: 10.1177/0014402916667586

Jung, P., McMaster**,** K., Kunkel, A., Shin, J., & Stecker, P. (2018). Effects of data-based individualization for students with intensive learning needs: A meta-analysis. *Learning Disabilities Research & Practice, 33*, 144-155.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. (4th Ed.). New York: Guilford Press.

Lembke, E., Allen, A., & Poch, A. (2015). Technical report #1 for the DBI-TLC project: Curriculum-Based Measurement screening study 1. June 1, 2015.

Lembke, E., Deno, S., & Hall, K. (2003). Identifying an indicator of growth in early writing proficiency for elementary school students. *Assessment for Effective Intervention, 28*, 23-35. doi: 10.1177/073724770302800304

Lembke, E., McMaster, K., Smith, R., Allen, A., Brandes, D., & Wagner, K. (2018). Professional Development for Data-Based Instruction in early writing: Tools, Learning, and

Collaborative Support. *Teacher Education and Special Education, 41*(2), 106-120. DOI 0888406417730112.

Lemons, C. J., Al Otaiba, S., Conway, S. J., & Mellado De La Cruz, V. (2016). Improving professional development to enhance reading outcomes for students in special education. In B. Foorman (Ed.), *Challenges to implementing effective reading intervention in schools. New Directions for Child and Adolescent Development*, *154*, 87–104. doi: 10.1002/cad.20177

McMaster, K. L., & Lembke, E. (2014). *Data-Based Instruction in Beginning Writing: A Manual*. Unpublished manual.

McMaster, K. L., Du, X., & Petursdottir, A. (2009). Technical features of curriculum-based measures for beginning writers. *Journal of Learning Disabilities, 42*, 41-60. doi: 10.1177/0022219408326212

McMaster, K. L., Du, X., Yeo, S., Deno, S. L., Parker, D., & Ellis, T. (2011). Curriculum-based measures of beginning writing: Technical features of the slope. *Exceptional Children, 77*, 185-206. doi: 10.1177/001440291107700203

McMaster, K. L., Kunkel, A., Shin, J., Jung, P., & Lembke, L. (2018). Early writing intervention: A best-evidence synthesis. *Journal of Learning Disabilities, 51*, 1-18. doi: 10.1177/0022219417708169

McMaster, K. L., Ritchey, K. D., & Lembke, E. (2011). Curriculum-based measurement for beginning writers: Recent developments and future directions. In T.E. Scruggs and M.A. Mastropieri (Eds.), *Assessment and intervention: Advances in learning and behavioral disabilities* (Vol. 24). Bingley, UK: Emerald.

McMaster, K. L., Shin, J., Espin, C., Jung, P., Wayman, M., & Deno, S. L. (2017). Monitoring elementary students' writing progress using curriculum-based measures: grade and gender differences. *Reading and Writing*, 1-23. doi: 10.1007/s11145-017-9766-9

National Center for Education Statistics *(2012) Digest of Education Statistics (US Department of Education, Washington).*

National Center on Intensive Intervention (n.d.). Retrieved on 12/14/2017 from https://intensiveintervention.org/

National Governors Association Center for Best Practices, Council of Chief State School Officers (2010). *Common Core State Standards (English Language Arts).* National Governors Association Center for Best Practices, Council of Chief State School Officers, Washington D.C.

Parker, D. C., Dickey, B. N., Burns, M. K., & McMaster, K. L. (2012). An application of brief experimental analysis with early writing. *Journal of Behavioral Education*, *21*, 329-349.

Poch, A. L., McMaster, K., & Lembke, E. (2018). Teachers' perceptions of the usability and feasibility of DBI for developing writers. Manuscript submitted for publication.

Poch, A. L., Smith, A., Wagner, K., Jung, P., Lembke, E., & McMaster, K. (2018). Using coaching to support teachers' use of data in academic instruction: A review. Manuscript in preparation.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd Ed.). Thousand Oaks, CA: Sage Publications.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2004). HLM 6 for Windows [Computer software]. Lincolnwood, IL: Scientific Software International.

Reis, S. M., McCoach, D. B., Coyne, M., Schreiber, F. J., Eckert, R. D., & Gubbins, E. J. (2007). Using planned enrichment strategies with direct instruction to improve reading fluency, comprehension, and attitude toward reading: An evidence-based study. *The Elementary School Journal, 108*, 3-23. doi: 10.1086/522383

Ritchey, K. D., Coker, D. L., & Jackson, A. F. (2015). The relationship between early elementary teachers' instructional practices and theoretical orientations and students' growth in writing. *Reading and Writing, 28*, 1333-1354. doi: 10.1007/s11145-015-9573-0

Roehrig, A. D., Duggar, S. W., Moats, L., Glover, M., & Mincey, B. (2008). When teachers work to use progress monitoring data to inform literacy instruction: Identifying potential supports and challenges. *Remedial and Special Education, 29*, 364-382. doi: 10.1177/0741932507314021

Saddler, B., & Graham, S. (2005). The effects of peer-assisted sentence-combining instruction on the writing performance of more and less skilled young writers. *Journal of Educational Psychology*, *97*, 43-54.

Schumaker, J. B., & Sheldon, J. B. (2005). *Fundamentals in the sentence writing strategy*. Lawrence, KS: Edge Enterprises, Inc.

Shanahan, T. (2004). Overcoming the dominance of communication: Writing to think and to learn. In T. L. Jetton & J. A. Dole (Eds.). *Adolescent literacy research and practice* (pp. 59–73). New York: Guilford.

Spear-Swerling, L., & Zibulsky, J. (2014). Making time for literacy: teacher knowledge and time allocation in instructional planning. *Reading and Writing, 27*, 1353-1378. doi: 10.1007/s11145-013-9491-y

Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools, 42*, 795-819. doi: 10.1002/pits.20113

Troia, G. A., Lin, S. J. C., Cohen, S., & Monroe, B. W. (2011). A year in the writing workshop: Linking writing instruction practices and teachers' epistemologies and beliefs about writing instruction. *The Elementary School Journal, 112*, 155-182. doi: 10.1086/660688

Troia, G. A., Olinghouse, N. G., Mo, Y., Hawkins, L., Kopke, R. A., Chen, A., ... & Stewart, K. A. (2015). Academic standards for writing: To what degree do standards signpost evidence-based instructional practices and interventions? *The Elementary School Journal, 116*, 291-321. doi: 10.1086/683984

Videen, J., Deno, S. L., & Marston, D. (1982). *Correct word sequences: A valid indicator of proficiency in written expression* (Vol. IRLD-RR-84). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

Wanzek, J., & Vaughn, S. (2009). Students demonstrating persistent low response to reading intervention: Three case studies. *Learning Disabilities Research & Practice, 24,* 151-163. doi: 10.1111/j.1540-5826.2009.00289.x

we really know about fidelity of implementation in schools? *Exceptional Children*, *79*, 181-193. doi: 10.1177/001440291307900204

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from http://ies.ed.gov/ncee/edlabs

Zumeta, R. O. (2015). Implementing intensive intervention: How do we get there from here? *Remedial and Special Education, 36*, 83-88. doi: 10.1177/0741932514558935

Table 1

Research-Based Mini-Lessons Aligned with Components of the Simple View of Writing

Mini-Lesson Title	Simple View of Writing Component	Description (what students do)
Transcription (TR) 1: Phonics Warm-Up[a]	Transcription	Identify letter(s) corresponding to sounds for short vowels, consonants, blends, and digraphs.
TR2: Alphabet Practice[b]	Transcription	Learn to correctly and efficiently write the letters of the alphabet.
TR3: Word Building[a]	Transcription	Practice writing words from basic word stems.
TR4: Word Study[a]	Transcription	Learn a strategy for studying spelling words.
TR5: Alphabet Rockets[b]	Transcription	Build fluency in writing target letters learned in the alphabet practice activity.
TR6: Word Sort[a]	Transcription	Practice identifying and discriminating between pairs of word sounds or endings.
Text Generation (TG) 1.1: Sentence Construction[c]	Text Generation	Learn and practice basic grammatical and syntactical elements of complete sentences, specifically capitalization and end punctuation.
TG1.2: Sentence Construction[c]	Text Generation	Learn additional elements of complete sentences, specifically subjects and verbs.
TG1.3: Sentence Construction[c]	Text Generation	Learn additional elements of complete sentences, specifically overall sentence coherence.
TG2: Sentence Combining[d]	Text Generation	Use grammatical strategies to combine simple and compound sentences.
TG3: Writing Goals[e]	Text Generation	Build handwriting automaticity through the use of performance-based incentives.
TG4: Repeated Writing[e]	Text Generation	Build handwriting automaticity through repeated practice.
TG5: Writing[a]	Text Generation	Integrate transcription and text generation skills to create a story or essay using a writing prompt.

Note. Lessons were adapted from the following intervention studies: a = Graham, Harris, & Fink-Chorzempa (2002); b = Graham, Harris, & Fink (2000); c = Schumaker & Sheldon (2005); d = Saddler & Graham (2005); e = Parker, Dickey, Burns, & McMaster (2012).

Table 2

Teacher Demographics by Condition and Site

	DBI-TLC (n = 11)			Control (n = 10)			
	Site 1 (n = 5) n (%)	Site 2 (n = 6) n (%)	DBI-TLC total	Site 1 ($n = 5^*$) n (%)	Site 2 (n = 5) n (%)	Control total	χ^2 (p-value)
Gender							n/a**
Female	5 (100%)	6 (100%)	11 (100%)	5 (100%)	5 (100%)	10 (100%)	
Ethnicity							2.43 (.296)
Asian American/Pacific Islander	-	-	-	-	-	-	
Black/African American	-	-	-	1 (20%)	-	1 (10%)	
Hispanic/Latino(a) American	-	-	-	-	1 (20%)	1 (10%)	
White/European American	5 (100%)	6 (100%)	11 (100%)	4 (80%)	4 (80%)	8 (80%)	
Age							1.78 (.620)
20-29	1 (20%)	1 (16.67%)	2 (18.18%)	1 (20%)	2 (40%)	3 (30%)	
30-39	2 (40%)	1 (16.67%)	3 (27.27%)	2 (40%)	1 (20%)	3 (30%)	
40-49	1 (20%)	4 (66.67%)	5 (45.45%)	1 (20%)	1 (20%)	2 (20%)	
50-59	1 (20%)	-	1 (9.09%)	1 (20%)	1 (20%)	2 (20%)	
Highest Degree							.76 (.683)
Bachelor's	3 (60%)	1 (16.67%)	4 (36.36%)	2 (40%)	2 (40%)	4 (40%)	
Master's	1 (20%)	2 (33.33%)	3 (27.27%)	2 (40%)	2 (40%)	4 (40%)	
Master's + Coursework	1 (20%)	3 (50%)	4 (36.36%)	1 (20%)	1 (20%)	2 (20%)	
Current Job Title							1.16 (.283)
Special Education Teacher	5 (100%)	5 (100%)	10 (100%)	5 (100%)	4 (80%)	9 (90%)	
English as a Second SL TeacherLanguage	-	-	-	-	1 (20%)	1 (10%)	
	M (Range)	*M* (Range)	*M* (Range)	*M* (Range)	*M* (Range)	*M* (Range)	t (p)
Years in current position	2.6 (1-5)	5 (1-20)	3.91 (1-20)	3.7 (2-4.5)	2.2 (1-4)	2.95 (1-4.5)	11.12 (.085)
Years teaching elementary	7.2 (1-15)	10 (2-20)	8.73 (1-20)	5.1 (3-10)	9 (1-25)	7.05 (1-25)	11.31 (.502)
Years teaching special education	9.8 (1-22)	9.6 (2-20.5)	9.68 (1-22)	4.3 (2-7.5)	7.25 (1-17)	5.61 (1-17)	15.29 (.226)

Note. DBI-TLC = Teachers who received Data-Based Instruction Tools, Learning, and Collaborative Support
*Control n for Site 1 includes two co-teachers who were considered one "unit" for this study.
**Chi-square test for gender was not conducted because all teachers (both DBI and control) were female.

Table 3

Student Demographics by Condition and Site

	DBI-TLC (n = 31)			Control (n = 22)			
	Site 1 (n = 14) n (%)	Site 2 (n = 17) n (%)	DBI-TLC total	Site 1 (n = 9) n (%)	Site 2 (n = 13) n (%)	Control total	χ^2 (p)
Gender (Female)	8 (57%)	9 (53%)	17 (55%)	3 (30%)	7 (54%)	10 (46%)	
Grade							2.02 (.733)
1	3 (21%)	1 (6%)	4 (13%)	3 (33%)	0	3 (14%)	
2	2 (14%)	7 (41%)	9 (29%)	2 (22%)	4 (31%)	6 (27%)	
3	3 (21%)	3 (18%)	6 (19%)	2 (22%)	4 (31%)	6 (27%)	
4	4 (29%)	3 (18%)	7 (23%)	2 (22%)	4 (31%)	6 (27%)	
5	2 (14%)	3 (18%)	5 (16%)	0	1 (8%)	1 (5%)	
Ethnicity							11.32* (.045)
Asian American/Pacific Islander	-	-	-	-	1 (8%)	1 (5%)	
Black/African American	2 (14%)	2 (12%)	4 (13%)	3 (33%)	3 (23%)	6 (27%)	
Hispanic/Latino(a) American	1 (7%)	1 (6%)	2 (6%)	1 (11%)	3 (23%)	4 (18%)	
White/European American	11 (79%)	11 (65%)	22 (71%)	4 (45%)	5 (38%)	9 (41%)	
American Indian	-	-	-	1 (11%)	1 (8%)	2 (9%)	
Multiracial	-	3 (17%)	3 (10%)	-	-	-	
Free/Reduced Lunch	11 (79%)	10 (59%)	21 (68%)	7 (78%)	9 (69%)	16 (73%)	.15 (.697)
English Language Learners	4 (29%)	2 (12%)	6 (19%)	2 (22%)	3 (23%)	5 (23%)	.09 (.765)
IEP/Special Education	14 (100%)	16 (94%)	30 (97%)	9 (100%)	9 (77%)	18 (82%)	1.99 (.157)
Special Education Category							8.22 (.412)
Autism	7 (50%)	1 (6%)	8 (27%)	5 (56%)	1 (10%)	6 (27%)	
Emotional/Behavioral Disorder	1 (7%)	-	1 (3%)	2 (22%)	1 (10%)	3 (14%)	
Learning Disability	2 (14%)	6 (38%)	8 (27%)	-	2 (20%)	2 (9%)	
Language Impairment	-	5 (31%)	5 (17%)	-	2 (20%)	2 (9%)	
Intellectual Disability	-	-	-	-	1 (10%)	1 (5%)	
Hearing Impairment	-	1 (6%)	1 (3%)	-	-	-	
Other Health Impairment	1 (7%)	3 (19%)	4 (13%)	-	3 (30%)	3 (14%)	
SNAP	3 (22%)	-	3 (10%)	2 (22%)	-	2 (9%)	
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	t (p)
Age	9.04 (1.59)	9.19 (1.29)	9.12 (1.41)	8.47 (1.39)	9.48 (1.45)	9.07 (1.48)	-.14 (.891)

Note. DBI-TLC = Teachers who received Data-Based Instruction Tools, Learning, and Collaborative Support; IEP = Individualized Education Program; SNAP = students needing alternative programming.
*$p < .05$.

Table 4

Pre- and Posttest Scores on Teacher Variables for DBI-TLC and Control Groups

	DBI-TLC group				Control group				Group difference	
	Mean	*SD*	Range	*n*	Mean	*SD*	Range	*n*	*t*/*F** (*p*)	*g*
Pretest										
Knowledge & Skills	25.55	4.95	16.00	11	25.67	2.78	8.00	9	.07 (.949)	-.03
Personal efficacy	4.51	.78	2.40	11	4.58	.36	.80	9	.05 (.825)	-.11
General efficacy	4.10	.76	2.66	11	4.53	.64	2.00	9	1.76 (.201)	-.58
WOS Correct	3.04	1.10	3.40	11	2.48	.91	2.00	9	1.49 (.238)	.53
WOS Explicit	4.98	.68	2.25	11	4.97	.78	1.75	9	.00 (.988)	.01
WOS Natural	4.37	.58	1.55	11	4.41	.50	1.33	9	.03 (.873)	-.07
Posttest										
Knowledge & Skills	32.32	2.26	8.00	11	23.5	3.52	10.00	9	-6.79 (.000)	2.92
Personal efficacy	4.95	.54	1.70	11	4.53	.48	1.00	8	3.02 (.100)	.78
General efficacy	4.55	.69	2.34	11	4.61	.70	1.84	8	.04 (.836)	-.08
WOS Correct	3.29	.72	2.40	11	3.03	.93	3.00	7	.45 (.512)	.31
WOS Explicit	5.75	.25	.75	11	5.11	.52	1.50	7	12.67 (.003)	1.63
WOS Natural	3.68	.71	2.56	11	4.35	.56	1.33	7	4.51 (.050)	-.97

Note. DBI-TLC = Teachers who received Data-Based Instruction Tools, Learning, and Collaborative Support; WOS = writing orientation scale.
**t*/*F* = *t*-values and *F*-values were calculated for knowledge & skills and efficacy/writing orientation variables, respectively.

Table 5

Fidelity of Data-Based Instruction (DBI) Implementation

Component	Mean Fidelity	Range
AIRS-Curriculum-Based Measurement (CBM)		
Materials	97.5%	75% - 100%
Directions	74.7%	20% - 100%
Demonstration	75.0%	0% - 100%
Timing	88.0%	75% - 100%
Overall CBM Fidelity	83.5%	50% - 100%
AIRS-Writing Instruction		
Introduction to lesson	70.8%	50% - 100%
Activity 1	97.0%	90% - 100%
Activity 2	86.0%	60% - 100%
Activity 3	93.0%	20% - 100%
Lesson wrap-up	27.5%	0% to 75%
Overall Writing Instruction fidelity	79.1%	65% to 93%
AIRS-Decision Making		
Timely decisions	39.4%	0% - 100%
Appropriate decisions	51.3%	0% - 100%
Noted change on graph	52.0%	0% - 100%
Noted type of change	65.7%	0% - 100%
Total timely & appropriate	45.3%	0% - 100%
Overall Decision Making fidelity	52.1%	0% - 100%

Note. AIRS = Accuracy of Implementation Rating Scale.

Table 6

Instructional Changes Made by DBI-TLC and Control Teachers

Category	Types of Instructional Changes	Number of Changes Made DBI-TLC	Control	Total *N* (%)
Setting or Format	Changes in motivational strategies	4	1	5 (45.45)
	Added behavioral intervention (e.g., behavior contract, school-home note system, etc.)	1	0	1 (9.09)
	Changes in time (e.g., frequency, duration of the instruction, etc.)	3	0	3 (27.27)
	Changes in group size	1	1	2 (18.18)
	Changes in format (e.g., order of activities implemented)	6	1	7 (63.64)
Delivery	More practice (e.g., repeated practice)	7	1	8 (72.73)
	More explicit instruction (e.g., modeling, guided practice, immediate and corrective feedback)	6	0	6 (54.55)
	More systematic instruction (e.g., re-teach foundational skills, break down tasks, scaffolded support)	3	0	3 (27.27)
Content	Changes in content (e.g., sounds, letters, words, vocabulary, genre)	3	1	4 (36.36)
Focus	Changes in writing focus (e.g., transcription, text generation, self-regulation)	7	0	7 (63.64)
Other	Raise the goal	1	0	1 (9.09)

Note. DBI-TLC = Teachers who received Data-Based Instruction Tools, Learning, and Collaborative Support.

Table 7

Pre- and Posttest Scores on CBM and TEWL-3 for DBI-TLC and Control Groups

CBM	DBI-TLC group (n = 31)			Control group (n = 20)			Group difference
	Mean	*SD*	Range	Mean	*SD*	Range	F (p)
Pretest							
WD – WW	19.44	9.69	38.3	13.88	7.30	27	4.83 (.033)
WD – WSC	5.58	4.99	19.7	3.65	3.41	11.5	2.29 (.137)
WD – CLS	58.20	35.41	137.2	37.73	26.49	83	4.90 (.032)
WD – CILS	24.92	36.91	202.7	14.48	25.26	82	1.23 (.273)
PW – WW	19.73	11.53	44	13.70	8.13	28	4.13 (.048)
PW – WSC	16.40	10.40	45.5	11.30	6.92	20.5	3.73 (.059)
PW – CWS	13.09	10.71	51	8.15	7.93	27	3.14 (.083)
PW – CIWS	1.74	12.83	62	-1.60	10.70	45	0.94 (.338)
Posttest							
WD – WW	24.89	9.45	41	17.35	9.57	31	(see HLM
WD – WSC	10.15	6.88	27.9	6.08	5.34	21.5	results)
WD – CLS	88.77	42.68	168	55.70	36.57	129.5	
WD – CILS	56.16	40.55	166.4	29.83	34.76	134.5	
PW – WW	29.01	11.31	43.5	22.68	11.34	36.5	
PW – WSC	26.26	10.75	43	19.83	10.49	35.5	
PW – CWS	26.29	13.67	58	17.78	12.17	42.5	
PW – CIWS	16.39	17.77	74.5	5.43	15.85	52.5	
SP – WW	23.36	13.30	52	16.35	9.81	32	
SP – WSC	17.58	11.16	39.5	13.39	8.63	30.5	
SP – CWS	11.27	8.71	40	8.46	7.09	25.5	
SP – CIWS	-3.18	10.54	59	-0.94	10.98	38.5	

TEWL-3	DBI group (n = 31)			Control group (n = 21)			Mean difference
	Mean	*SD*	Range	Mean	*SD*	Range	F (p)
Pretest							
Basic Writing	39.32	10.79	40	32.95	10.57	34	2.11 (.055)
Contextual Writing	20.35	12.19	50	15.14	10.32	31	1.61 (.116)
Posttest							
Basic Writing	43.58	11.73	57	37.38	14.09	53	(see HLM
Contextual Writing	24.61	12.46	44	17.52	12.22	49	results)

Note. CBM = Curriculum-based measure, WD = Word Dictation, PW = Picture Word, SP = Story Prompt. WW = Words Written, WSC = Words Spelled Correctly, CLS = Correct Letter Sequences, CILS = Correct minus Incorrect Letter Sequences, CWS = Correct Word Sequences. CIWS = Correct minus Incorrect Word Sequences, HLM = hierarchical linear modeling, TEWL-3 = Test of Early Written Language-3. Story prompts were administered at posttest only. TEWL-3 means and *SD*s are based on raw scores.

Table 8

Parameter Estimates for Two-Level Model for Posttest Word Dictation and Effect Sizes

WW	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	19.78	1.61	12.29	18	<.001	
Condition (γ_{01})	3.21	2.14	1.49	18	.152	
Pretest (γ_{10})	.76	.11	7.11	49	<.001	
	SD	Variance	Chi-square	*df*	*p*	.316
Random effects:						
Level-1 variance (χ^2)	6.21	38.56	-	-	-	
Intercept variance (τ_{00})	2.29	5.25	23.01	18	.190	
WSC	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	7.12	1.45	4.90	18	<.001	
Condition (γ_{01})	2.24	1.94	1.15	18	.265	
Pretest (γ_{10})	.29	.09	3.31	49	.002	
	SD	Variance	Chi-square	*df*	*p*	.342
Random effects:						
Level-1 variance (χ^2)	4.87	23.70	-	-	-	
Intercept variance (τ_{00})	2.78	7.75	32.37	18	.020	
CLS	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	64.57	8.20	7.87	18	<.001	
Condition (γ_{01})	17.20	10.99	1.57	18	.135	
Pretest (γ_{10})	2.77	.51	5.45	49	<.001	
	SD	Variance	Chi-square	*df*	*p*	.395
Random effects:						
Level-1 variance (χ^2)	27.36	748.57	-	-	-	
Intercept variance (τ_{00})	15.80	249.67	32.78	18	.018	
CILS	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	35.63	9.37	3.80	18	.002	
Condition (γ_{01})	16.08	12.56	1.28	18	.217	
Pretest (γ_{10})	1.68	.56	3.02	49	.004	.40
	SD	Variance	Chi-square	*df*	*p*	
Random effects:						
Level-1 variance (χ^2)	29.08	845.70	-	-	-	
Intercept variance (τ_{00})	19.61	384.70	38.45	18	.004	

Note. WW = words written, WSC = words spelled correctly, CLS = correct letter sequences, CILS = correct minus incorrect letter sequences

Table 9

Parameter Estimates for Two-Level Model for Posttest Picture Word and Effect Sizes

WW	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	25.20	2.55	9.90	18	<.001	
Condition (γ_{01})	2.62	3.38	.78	18	.448	
Pretest (γ_{10})	.76	.14	5.35	48	<.001	
	SD	Variance	Chi-square	*df*	*p*	.225
Random effects:						
Level-1 variance (χ^2)	7.06	49.80	-	-	-	
Intercept variance (τ_{00})	5.74	32.96	47.74	18	<.001	

WSC	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	22.44	2.42	9.29	18	<.001	
Condition (γ_{01})	2.56	3.21	.80	18	.435	
Pretest (γ_{10})	.78	.13	6.03	48	<.001	
	SD	Variance	Chi-square	*df*	*p*	.232
Random effects:						
Level-1 variance (χ^2)	6.23	38.82	-	-	-	
Intercept variance (τ_{00})	5.70	32.53	56.19	18	<.001	

CWS	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	20.41	2.97	6.87	18	<.001	
Condition (γ_{01})	4.31	3.97	1.09	18	.292	
Pretest (γ_{10})	1.01	.15	6.94	48	<.001	
	SD	Variance	Chi-square	*df*	*p*	.312
Random effects:						
Level-1 variance (χ^2)	6.55	42.89	-	-	-	
Intercept variance (τ_{00})	7.54	56.89	80.68	18	<.001	

CIWS	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	8.76	3.95	2.22	18	.039	
Condition (γ_{01})	5.84	5.23	1.12	18	.279	
Pretest (γ_{10})	1.10	.22	4.96	48	<.001	.328
	SD	Variance	Chi-square	*df*	*p*	
Random effects:						
Level-1 variance (χ^2)	11.28	127.17	-	-	-	
Intercept variance (τ_{00})	8.70	75.67	44.46	18	.001	

Note. WW = words written, WSC = words spelled correctly, CWS = correct word sequences, CIWS = correct minus incorrect word sequences

Table 10

Parameter Estimates for Two-Level Model for Posttest Story Prompts and Effect Sizes

WW	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	18.67	2.77	6.73	18	<.001	
Condition (γ_{01})	3.26	3.66	.89	18	.385	
Pretest (γ_{10})	.64	.16	3.90	48	<.001	
	SD	Variance	Chi-square	*df*	*p*	.314
Random effects:						
Level-1 variance (χ^2)	8.92	79.62	-	-	-	
Intercept variance (τ_{00})	5.44	29.64	33.86	18	.013	
WSC	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	15.55	2.21	7.02	18	<.001	
Condition (γ_{01})	.63	2.92	.22	18	.831	
Pretest (γ_{10})	.64	.13	4.94	48	<.001	
	SD	Variance	Chi-square	*df*	*p*	.061
Random effects:						
Level-1 variance (χ^2)	6.99	48.83	-	-	-	
Intercept variance (τ_{00})	4.45	19.76	35.17	18	.009	
CWS	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	10.24	1.72	5.97	18	<.001	
Condition (γ_{01})	-.19	2.27	-.09	18	.932	
Pretest (γ_{10})	.57	.09	5.84	48	<.001	
	SD	Variance	Chi-square	*df*	*p*	-.024
Random effects:						
Level-1 variance (χ^2)	4.91	24.13	-	-	-	
Intercept variance (τ_{00})	3.78	14.26	43.53	18	.001	
CIWS	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	-.04	2.85	-.01	18	.990	
Condition (γ_{01})	-3.97	3.77	-1.05	18	.307	
Pretest (γ_{10})	.42	.16	2.60	48	.013	-.369
	SD	Variance	Chi-square	*df*	*p*	
Random effects:						
Level-1 variance (χ^2)	8.11	65.75	-	-	-	
Intercept variance (τ_{00})	6.29	39.75	44.22	18	.001	

Note. WW = words written, WSC = words spelled correctly, CWS = correct word sequences, CIWS = correct minus incorrect word sequences

Table 11

Parameter Estimates for Two-Level Model for Posttest TEWL-3

Basic Writing	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	41.64	1.64	25.35	18	<.001	
Condition (γ_{01})	.19	2.10	.09	18	.929	
Pretest (γ_{10})	.94	.12	7.83	19	<.001	
	SD	Variance	Chi-square	*df*	*p*	.015
Random effects:						
Level-1 variance (χ^2)	6.18	38.23	-	-	-	
Intercept variance (τ_{00})	1.63	2.66	26.04	16	.053	
Slope variance (τ_{11})	.32	.10	33.92	17	.009	

Contextual Writing	Coefficient	*SE*	*t*	*df*	*p*	ES
Fixed effects:						
Intercept (γ_{00})	20.72	1.92	10.81	18	<.001	
Condition (γ_{01})	1.68	2.53	.66	18	.515	
Pretest (γ_{10})	.85	.11	7.69	49	<.001	.132
	SD	Variance	Chi-square	*df*	*p*	
Random effects:						
Level-1 variance (χ^2)	7.99	63.79	-	-	-	
Intercept variance (τ_{00})	1.94	3.78	19.31	18	.373	