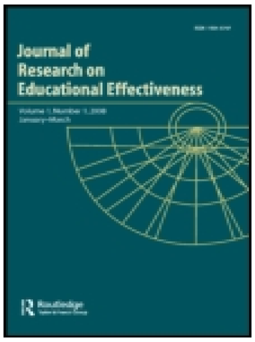

Routledge
Taylor & Francis Group

Journal of Research on Educational Effectiveness

ISSN: 1934-5747 (Print) 1934-5739 (Online) Journal homepage: https://www.tandfonline.com/loi/uree20

Effective Classroom Instruction: Implications of Child Characteristics by Reading Instruction Interactions on First Graders' Word Reading Achievement

Carol McDonald Connor , Frederick J. Morrison , Christopher Schatschneider , Jessica R. Toste , Erin Lundblom , Elizabeth C. Crowe & Barry Fishman

Journal of Research on Educational Effectiveness, 4: 173–207, 2011

ISSN: 1934-5747 print / 1934-5739 online
DOI: 10.1080/19345747.2010.510179

Routledge
Taylor & Francis Group

INTERVENTION, EVALUATION, AND POLICY STUDIES

Effective Classroom Instruction: Implications of Child Characteristics by Reading Instruction Interactions on First Graders' Word Reading Achievement

Carol McDonald Connor
Florida State University, Tallahassee, Florida, USA

Frederick J. Morrison
University of Michigan, Ann Arbor, Michigan, USA

Christopher Schatschneider
Florida State University, Tallahassee, Florida, USA

Jessica R. Toste
McGill University, Montreal, Quebec, Canada

Erin Lundblom and Elizabeth C. Crowe
Florida State University, Tallahassee, Florida, USA

Barry Fishman
University of Michigan, Ann Arbor, Michigan, USA

Abstract: Too many children fail to learn how to read proficiently with serious consequences for their overall well-being and long-term success in school. This may be because providing effective instruction is more complex than many of the current models of reading instruction portray; there are Child Characteristic × Instruction (CXI) interactions. Here we present efficacy results for a randomized control field trial of the Individualizing Student Instruction (ISI) intervention, which relies on dynamic system forecasting intervention models to recommend amounts of reading instruction for each student, taking into account CXI interactions that consider his or her vocabulary and reading skills. The study, conducted in seven schools with 25 teachers and 396 first graders, revealed that students in the ISI intervention classrooms demonstrated significantly greater reading skill gains by spring than did students in control classrooms. Plus, they were more likely to receive differentiated reading instruction based on CXI interaction guided recommended amounts than were students in control classrooms. The precision with which students received the recommended amounts of each type of literacy instruction, the distance from recommendation, also predicted reading outcomes.

Address correspondence to Carol McDonald Connor, Florida State University, Florida Center for Reading Research/Developmental Psychology, PDB 306-4304, Tallahassee, FL 32306-4304, USA. E-mail: cconnor@fcrr.org

Keywords: Literacy, classroom instruction, reading, differentiated instruction

The environments in which children live and learn have a significant impact on their adaptation and growth (Masten & Coatsworth, 1998). In particular, schools and classroom environments play a central role in shaping children's well-being and positive adjustment (Connor, Son, Hindman, & Morrison, 2005; NICHD-ECCRN, 2002, 2004). Indeed, the promotion of academic success has been asserted to play an essential role in children's development (Bronfenbrenner & Morris, 2006; Cicchetti, 1990; Reynolds & Ou, 2004; Rutter & Maughan, 2002) and the development of literacy skills is critical to this academic success (NRP, 2000; Snow, Burns, & Griffin, 1998). However, too many children in the United States fail to achieve functional literacy. Across the nation, more than 30% of students fail to achieve basic reading skills by the fourth grade, and the rate is higher, close to 60%, among children living in poverty and who belong to underrepresented minorities (NAEP, 2007). Increasingly, research demonstrates that students who are unable to read are more likely to become frustrated, overwhelmed, or disinterested, with clear consequences for their engagement in learning and their future success in schools (G. M. Morrison & Cosden, 1997). Moreover, reading difficulties have long-term implications for children's well-being including grade retention, referral to special education, dropping out of high school, and entering the juvenile criminal justice system (Reynolds, Temple, Robertson, & Mann, 2002).

Many children fail only because they do not receive the amount and type of instruction they need (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; F. J. Morrison, Bachman, & Connor, 2005; Vellutino et al., 1996). Early literacy instruction that is balanced between basic skill, or code-based instruction, and meaningful reading experiences has been shown to be more effective than instruction that focuses on one to the exclusion of the other (Mathes et al., 2005; Xue & Meisels, 2004). However, providing balanced instruction may be more complex than current models of reading instruction imply. Accumulating evidence reveals that the effect of any particular instructional strategy will likely vary with each child's language and literacy skills (Connor, Morrison, & Katch, 2004; Foorman et al., 1998; Juel & Minden-Cupp, 2000). These have been called *Child Characteristic* × *Instruction interactions* (CXI; Connor, Morrison, & Katch, 2004). They have also been called *Aptitude* × *Treatment* interactions (Cronbach & Snow, 1969, 1977). Child × Instruction interaction research demonstrates that relations among instruction, child characteristics, and outcomes are nonlinear, transactional, and dynamic (Connor, Piasta, et al., 2009), with correlational and emerging quasi-experimental evidence for CXI interactions from kindergarten through third grade (Al Otaiba et al., 2010; Connor, Jakobsons, Crowe, & Meadows, 2009; Connor, Morrison, Fishman, Schatschneider, & Underwood, 2007; Connor, Morrison, & Katch, 2004; Connor, Morrison, & Petrella, 2004; Connor, Morrison, & Underwood, 2006).

The purpose of this study was to test whether CXI interactions are causally implicated in the varying achievement observed among children within and between classrooms and schools. Explicitly, we aimed to extend previous research (Connor, Morrison, Fishman, et al., 2007) by examining the effect of differentiating children's literacy instruction on students' reading skill gains using the Individualized Student Instruction (ISI) intervention, which takes into account CXI interactions observed in longitudinal correlational studies. The ISI intervention has five components: (a) conceptualizing reading instruction across multiple dimensions, (b) student assessment, (c) Assessment-to-instruction (A2i) software, (d) professional development (PD), and (e) implementation during a dedicated literacy block utilizing flexible homogeneous, skill-based small-group instruction. The A2i software is a teacher instruction planning tool that computes recommended amounts and types of reading

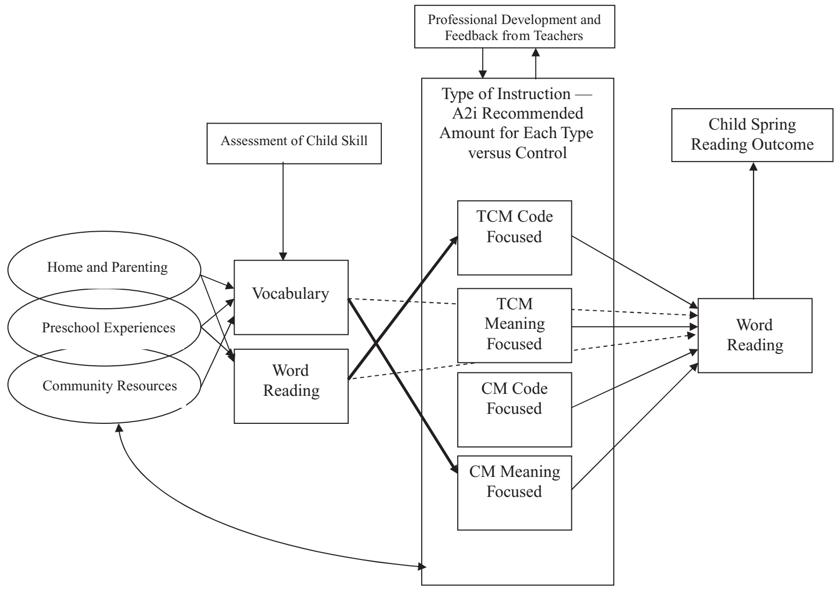

Figure 1. Theory of change for the Individualizing Student Instruction intervention. *Note.* Children enter the classroom with varying vocabulary and word reading skills. Instruction that takes into account the differential effectiveness of reading instruction strategies (e.g., teacher/child-managed [TCM] code focused, child-managed [CM] meaning focused) will likely be more effective than instruction that does not take into account Child × Instruction interactions and will predict stronger gains in word reading skills. Professional development is designed to support teachers' implementation of the Assessment-to-instruction (A2i) recommended amounts of instruction. Feedback from teachers is used to improve the A2i software and professional development protocol.

instruction for each student using dynamic forecasting intervention models (specifically computer algorithms) that consider the relations among child characteristics and types of reading instruction (Connor, in press). These dynamic forecasting intervention models, not unlike those used by meteorologists (Rhome, 2007), translate theoretically and empirically derived projections of what composes optimal reading instruction for each child based on his or her vocabulary and reading skills, using empirical findings of CXI interactions, to provide recommendations that teachers can implement in the classroom. The model informing this study is provided in Figure 1.

MULTIDIMENSIONAL VIEW OF EARLY READING INSTRUCTION

The ISI intervention conceptualization of reading instruction relies on four dimensions of instruction: (a) who is managing or focusing the students' attention on the learning activity at hand, (b) the content of the instruction, (c) context or grouping during the activity, and (d) change in duration of instruction type over time (Connor, Morrison, et al., 2009). These dimensions operate simultaneously to define specific reading instruction activities (see Table 1 and Figure 1).

First, the dimension of management, *teacher- teacher/child-* versus *child-managed instruction* captures who is responsible for focusing the child's attention on the learning activity at hand, the teacher, the teacher and child together, or the child and his or her peers. For instance, a teacher reading and discussing a book with a small group of children would

Table 1. Dimensions of instruction: Components from Individualizing Student Instruction Coding System across two dimensions, management (columns) and content (rows)

	Teacher/Child Managed	Child Managed
Code focused	Phonological awareness	Phonological awareness
	Grapheme-phoneme Correspondence	Grapheme-phoneme correspondence
	Word decoding	Word decoding
	Word encoding	Word encoding
	Morphological awareness	Morphological awareness
	Letter and word fluency	Letter and word fluency
Meaning focused	Comprehension	Comprehension
	Text fluency	Text fluency
	Text reading	Text reading
	Writing	Writing
	Print vocabulary	Print vocabulary
	Oral language	Oral language
	Print and text concepts	Print and text concepts

Note. Other dimensions include context (i.e., grouping) and change over time.

be considered a teacher/child-managed (TCM) activity because the teacher and child are actively interacting. The activity becomes child managed (CM) when the teachers asks children to break into small groups to write about what they learned from the book and expects the children to work independently or with peers (Connor, Morrison, et al., 2009). In this study, we consider TCM instruction to be explicit or direct, and it can be highly interactive, such as when teachers provide coaching or scaffolding (hence the teacher/child designation) as has been described in other studies (Graue, 2004; Taylor, Pearson, & Rodriquez, 2003). Teacher-managed (TM) instruction is what has been labeled as didactic (Smith, Lee, & Newmann, 2001). To be considered TM instruction, the activity had to last for at least 3 min where the teacher talked/instructed and did not invite the children to participate. If the teacher asked questions or gave students any opportunity to participate in the activity, then it was considered TCM. Virtually no first-grade instruction of any type was observed to be TM according to this definition (less than 30 s on average across classrooms). Hence we did not include TM instruction in this study or during PD. Instead, teachers were encouraged to use TCM instruction.

Second, we identify instruction as either *code-focused* or *meaning-focused* based on classroom observations and theories of reading (see Table 1; Hoover & Gough, 1990; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001, 2002; Scarborough, 2001; Snow, 2001). *Code-focused instruction* activities explicitly focus on helping children learn to decode (Foorman et al., 1998). This would include alphabet activities, phonological awareness, phonics, and letter and word fluency (e.g., letter and letter-sound knowledge, blending letter sounds to form words).

Meaning-focused instruction activities support students' efforts to actively extract and construct meaning from text (Snow, 2001) and include reading aloud, reading independently, writing, oral language, vocabulary, and comprehension strategies. At times, the line between code- and meaning-focused instruction blurs, for example, when using decodable books. The designation is based on what the student is intended to learn. If the purpose of the activity is to help students decode the word family, "-oat" presented in the decodable book ("The goat is rolling in oat seeds by the moat"), then the activity is code focused. If, however, the decodable book is used to allow children to practice reading-connected text, then the activity is meaning focused. Teachers in the study had no difficulty grasping this distinction.

Third, the dimension of *context* or *grouping* relies on the research findings that instruction provided to the entire class—*whole class or classroom level*—is significantly different in supporting children's early reading skill growth than is instruction provided in small groups or individually—*student level*—(Connor, Morrison, & Slominski, 2006). It also relies on the effective schools research (Taylor, Pearson, Clark, & Walpole, 2000; Wharton-McDonald, Pressley, & Hampston, 1998), which indicates that effective teachers use small homogeneous, flexible learning groups more frequently than do less effective teachers. For example, the teacher might introduce a phonics activity and provide a basic lesson to the entire class (whole class) and then ask children to break into smaller groups to play a related phonics game while she or he works with a small group of students with weaker reading skills who are unlikely to have mastered the content from the whole-class instruction alone. The ISI intervention was designed to encourage the use of small homogeneous, flexible learning groups.

Fourth, the dimension of *change-over-time* indicates that particular instructional strategies may be more important for students at certain times of the year. For example, for children with stronger vocabulary skills, frequent opportunities to read and write independently have been found to be associated with stronger reading skill gains. However, such activities have been associated with less reading skill growth for children whose vocabulary skills were below expectations. For them, less time in independent reading activities in the fall with increasing amounts over the school year have been associated with greater reading skill gains (Connor, Morrison, & Katch, 2004). This effect of timing further underscores the dynamic nature of instruction and how it differs for children depending on the skills they bring to the classroom. The ISI intervention was designed to explicitly direct teachers' attention to changing amounts and types of instruction over the course of the school year.

Again, these four dimensions operate simultaneously (see Table 1) to generate four specific types of reading instruction—TCM code-focused, TCM-meaning-focused, CM code-focused, and CM meaning-focused reading instruction that varied by group size, amount (minutes), and content level (i.e., difficulty). For example, a teacher explaining to the entire class about word families (e.g., *bat*, *cat*, *mat*, *sat*) and how to use the similarities to decode unknown words would be a TCM code-focused whole-class activity. If the teacher then asks the children to read a book with a buddy, that activity would be a CM meaning-focused small-group/pair activity. If the teacher then works with a small group of children to more explicitly teach them how to decode using onset and rime, that would be a TCM code-focused small-group activity.

As part of the ISI intervention PD, teachers learned how to use this multidimensional framework to describe research-based literacy instruction activities. The recommended amounts (in minutes/day) for each child were provided for each type of activity and could be viewed using the A2i software in the Classroom View (see Figure 2). In addition, their core literacy curriculum and the materials provided by the school were indexed according to the types of instruction. Thus, the ISI intervention did not constitute a new reading curriculum but rather a way to plan and implement reading instruction (all four types) that was differentiated for each student based on his or her reading and vocabulary skills.

CXI INTERACTIONS: INDIVIDUALIZING STUDENT INSTRUCTION

A growing research base establishes that there are multiple sources of influence on children's early literacy skills (see Figure 1), including home, preschool, disabilities (e.g.,

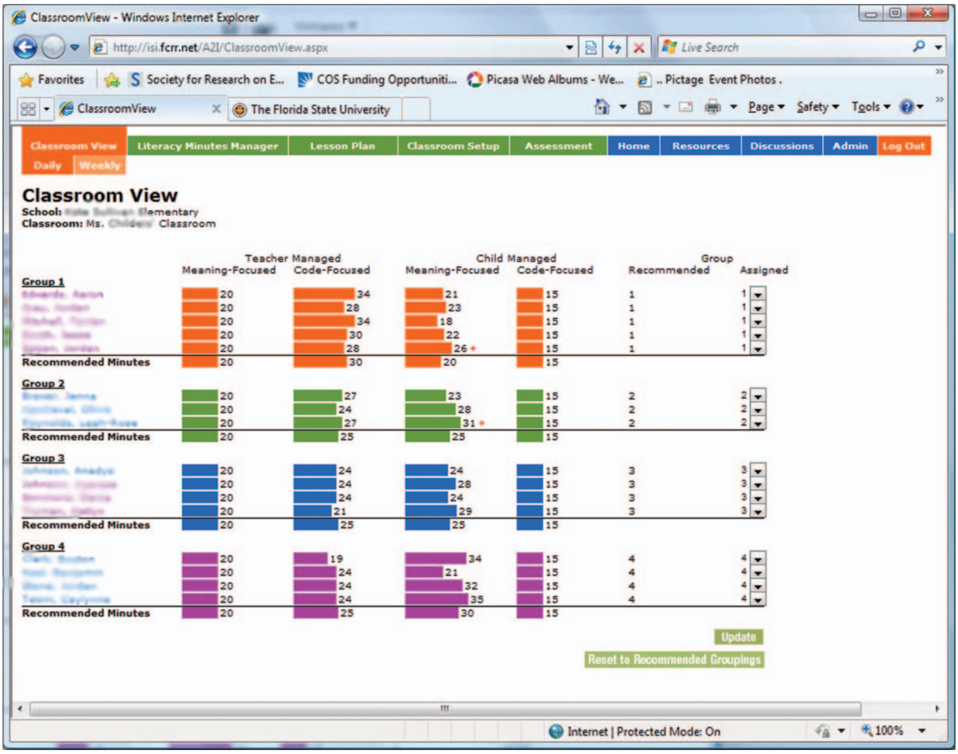

Figure 2. The classroom view from the Assessment-to-instruction (A2i) software in fall. *Note.* This is where teachers view the recommended amounts of each type of reading instruction activity. Clicking children's names leads to a pop-up window showing test scores and progress monitoring charts. (Color figure available online).

language impairment, cerebral palsy), and socioeconomic status (SES), among others (Bronfenbrenner & Morris, 2006; Connor, Son, Hindman, & Morrison, 2005; NICHD-ECCRN, 2004). As such, children begin first grade with widely different language and early literacy skills (e.g., children from low SES families frequently, but not always, start school with weaker vocabulary skills). Plus, there is emerging evidence that students with above-average reading skills who attend higher poverty schools are significantly less likely to maintain these skills than are their peers at more affluent schools (Neal & Schanzenbach, 2007). Children with disabilities are increasingly being included in general education classroom settings (Al Otaiba et al., 2008; Connor, Jakobsons, et al., 2009; Foorman et al., 2006). Altogether, this makes it necessary to consider an even greater diversity of skills within classrooms.

Research indicates that the characteristics and aptitudes (Kyllonen & Lajoie, 2003) children bring to the classroom affects the efficacy of the reading instruction they receive (see Figure 1). For example, Foorman et al. (1998) found that explicit code-based instruction was more effective in promoting students' reading skills when children began the year with weaker phonological awareness skills. CXI interactions have been found from preschool through third grade (Connor, Morrison, & Katch, 2004; Connor, Morrison, & Petrella, 2004; Connor, Morrison, & Slominski, 2006; Connor, Morrison, & Underwood, 2007), with different samples of children (Connor, Jakobsons, et al., 2009), and by numerous researchers (Al Otaiba et al., 2008; Foorman et al., 2006; Juel & Minden-Cupp, 2000).

The individualized or differentiated instruction implemented in this study relied on previous correlational research that found significant CXI interactions, which affected first-grade students' word recognition skill growth (Connor, Morrison, & Katch, 2004). Total amounts of TCM meaning-focused and CM code-focused instruction did not significantly predict students' word reading skills. Rather, children with weaker fall word recognition skills (i.e., below grade expectations) demonstrated greater progress when they were in classrooms where teachers provided greater amounts of TCM code-focused instruction than did children with similar skills in classrooms where smaller amounts of this type of instruction was observed. In contrast, children with stronger fall word recognition skills showed less progress in classrooms where greater amounts of TCM code-focused instruction were observed, whereas students with similar skills made greater progress in classrooms where less of this type of instruction was observed. The recommended amount of TCM code-focused instruction for the children in this study is depicted in Figure 3 (top).

In addition, children who started first grade with weaker vocabulary skills demonstrated greater word recognition skill growth in classrooms where they received smaller amounts of CM meaning-focused instruction in the fall with steady increases in amount until spring (i.e., change over time). This is in contrast to children with similar skills who demonstrated weaker gains when they were in classrooms that provided greater amounts of CM meaning-focused consistently all year long. Yet, in the latter classrooms, children with stronger fall vocabulary scores showed greater word recognitions skill growth when they were in classrooms where greater amounts of CM meaning-focused instruction were provided than they did in classrooms with lesser amounts of CM meaning-focused instruction. It is these specific Child × Instruction interactions that the A2i algorithms used in this study were designed to capture (see Figure 3, bottom). That is, the hierarchical linear models (HLM) used in these studies and results from other studies of first-grade reading instruction (Connor, Jakobsons, et al., 2009; Connor, Morrison, Fishman, et al., 2007) provided the foundation for the equations used in the dynamic forecasting intervention models used in the current study. We hypothesized that reading instruction, which explicitly considered CXI interactions, should be more effective than instruction that did not specifically consider CXI interactions (see Figure 1).

THE ISI INTERVENTION

Again, the ISI intervention has five components that work together to support teachers' reading instruction planning and implementation: (a) multiple dimensions conceptualization of reading instruction just described; (b) student assessment and progress monitoring; (c) A2i web-based software; (d) teacher training, including online PD resources, workshop, school, and classroom-based support; and (e) implementation in the classroom.

Student Assessment and Progress Monitoring

Using student assessment data to plan instruction is the heart of the ISI intervention. The assessments used for this research were the Woodcock–Johnson Tests of Achievement III (WJ–III) Letter-Word Identification and Picture Vocabulary subtests, which are described more fully next. Theoretically, any valid and reliable assessment of word reading and vocabulary should adequately inform instruction. Indeed, the original HLM models utilized the Peabody Individual Achievement Test battery (Markwardt, 1998). The challenge is to

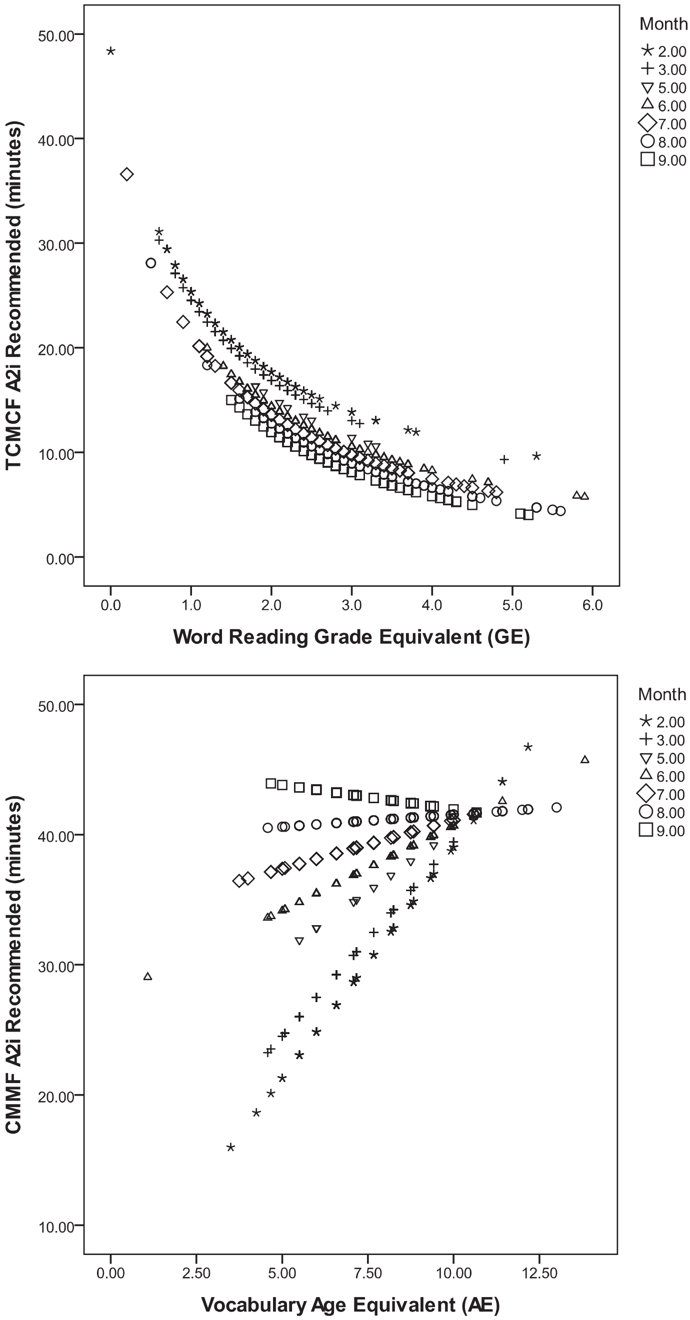

Figure 3. Top: Assessment-to-instruction (A2i) algorithm teacher/child-managed code-focused (TCMCF) instruction recommended amounts as a function of children's reading GE by month (September = 1, October = 2, etc.) assuming vocabulary scores falling at the mean AE. *Note.* A GE of 1.0 corresponds to beginning of first-grade reading achievement. A GE of 0 corresponds to a beginning of kindergarten level, and so on. Thus the recommended amount of TCMCF would be approximately 25 min/day in September for a child reading half a grade level below first-grade expectations. Bottom: A2i algorithm child-managed meaning-focused (CMMF) instruction recommended amounts as a function of children's vocabulary AE and month of the year.

facilitate interpretation of the data so that teachers can plan and implement instruction based on students' assessed strengths and weaknesses in reading and vocabulary.

When students were reassessed in the winter, the updated scores were entered into A2i and new groups and instruction recommendations were computed. These changes were carefully reviewed with teachers so that, for example, students who made strong gains were placed in the newly recommended group and received the recommended regime. Using the metric of grade and age equivalents (GE, AE) provided meaningful ways for teachers to view their students' growth. Feedback from the teachers indicated that they liked the use of GE and AE scores because it explicitly helped them interpret the scores.

A2i Software

A2i software was designed to support teachers' interpretation and use of assessment data for planning and implementing differentiated instruction. Using students' reading and vocabulary scores, A2i software computes recommended amounts of TCM code-focused and CM meaning-focused instruction for each child in the classroom based on the CXI interactions observed in the correlational studies (see Figure 3). These recommendations are provided in the classroom view (see Figure 2). Using a set target outcome, which was defined as end of first-grade achievement according to district norms (GE = 2.1), children's assessed word reading and vocabulary skills, and the month of the school year (where September = 1, October = 2, etc.), the algorithms solved for each type of reading instruction.

Because slope was implicated in the CM meaning-focused models, the recommended amount varied each month depending on students' vocabulary skills. The function for CM meaning-focused instruction for the year is provided in Figure 3 (bottom). Note how the recommended amount increases each month for students with weaker vocabulary skills but remains steady for students with higher vocabulary skills. In both parts of Figure 3, each dot represents a recommended amount for a child at a particular month in the study.

TCM meaning-focused and CM code-focused instruction amounts were set at the means observed in the Connor, Morrison, and Katch (2004) study because neither significantly predicted students' word reading growth. The recommendations for each student are displayed in the A2i online classroom view (see Figure 2). By clicking the child's name, the teacher accessed test scores and progress monitoring charts. A2i also includes literacy block scheduling tools, online PD resources, and lesson planning features (Connor, in press). The website is http://isi.fcrr.org, and the log-in for the first-grade demonstration classroom is A2idemo; the password is isi06! (include the exclamation point).

The schools' core reading curriculum and Florida Center for Reading Research center activities (http://fcrr.org/Curriculum/SCAindex.htm) were indexed to the four types of instruction (i.e., TCM code-focused, CM meaning-focused, etc.). Again, the ISI intervention was not a new curriculum but rather, using the available school resources, was intended to provide a way to implement differentiated reading instruction more effectively for each child.

Professional Development

Although this study was designed as an efficacy study, where typically instruction is provided by researchers, we relied on the classroom teachers to provide the ISI instructional regimen. Thus, teachers received intensive PD, including online resources, to help them provide the A2i recommended amounts of instruction for each child in their classroom.

A coaching or mentoring model was used (Bos, Mather, Narr, & Babur, 1999; Vaughn & Coleman, 2004), which included monthly school-level meetings and biweekly classroom-based support during literacy instruction. PD focused on how to use the A2i software for progress monitoring and planning instruction; using assessment to guide instruction; classroom management, using stations or centers effectively; differentiating instructional content according to students' reading and language skill levels; and using research to inform practices, among other topics.

PD was provided by highly trained research assistants who, except for one, had teaching experience and were fully credentialed. The one assistant who was not certified worked under the close supervision of a research assistant who was certified. Called research partners, these individuals worked with the teachers at one or two schools. During the literacy block, the research partners were active participants and would model classroom management and instruction strategies, help the teacher with classroom management, including organizing small groups, facilitating transitions, and developing daily schedules. The goal was to support teachers' efforts to provide instruction in small flexible, homogeneous groups at the teacher table while the other students worked independently or with peers.

Implementation

Designed to be implemented during a dedicated block of time devoted to literacy instruction, the ISI intervention relied on the evidence from the effective schools literature (Taylor et al., 2000; Wharton-McDonald et al., 1998) that revealed that effective schools tended to use flexible, homogeneous skill-based grouping during a dedicated block of time devoted to literacy. This was coupled with strong planning (Fuchs, Fuch, & Phillips, 1994) and classroom management (Brophy, 1983; Cameron, Connor, Morrison, & Jewkes, 2008). A similar model was also used in the Florida Reading First model at the time of the study (Connor, Jakobsons, & Granger, 2006; Crowe & Connor, 2007), which included the use of a research-based core literacy curriculum. Fully implemented, teachers provided the recommended amounts of TCM code-focused and CM meaning-focused instruction daily, differentiating not only duration but also content and skill requirements, based on the A2i recommendations and each student's assessment results. The goal was exemplary classroom management with all of the children engaged in meaningful literacy instruction throughout the literacy block. Teachers followed lesson plans developed using A2i or using the school-required lesson plan.

Of note, in this school district, a dedicated 90-min block of time devoted to literacy instruction and use of small-group instruction was mandated for all schools. Thus both treatment and control teachers were expected and were observed to provide a 90-min literacy block that included small-group instruction. The specific treatment implementation manipulation was providing the recommended amounts of TCM code-focused and CM meaning-focused reading instruction computed by the A2i software.

RESEARCH QUESTIONS

We asked the following research questions:

1. What was the effect of individualizing student literacy instruction using A2i recommended amounts (i.e., the ISI intervention) compared to literacy instruction provided

as usual (control group)? We hypothesized that students in the ISI classrooms would demonstrate greater word reading skill gains than would their peers in the control classrooms.
2. Was the ISI intervention more or less effective for children with differing characteristics, specifically initial vocabulary and reading skills, the SES of the school, and assignment to special education? Because the essence of the ISI intervention was that a specific plan for effective reading instruction was provided for each student, we anticipated no Child Characteristic × ISI interactions.
3. To what extent did teachers receiving the ISI intervention individualize their students' instruction compared to control group teachers—that is, were they more likely to provide their students the A2i recommended amounts of TCM code-focused and CM meaning-focused reading instruction? We hypothesized that teachers in the ISI intervention group would be more likely to differentiate their students' instruction following A2i recommendations than would teachers in the control classroom. The implicit assumption was that instruction delivered to the entire classroom was not differentiated. Thus, although we examined instruction during the entire literacy block, only instruction provided in small groups or individually was considered to be individualized. We also hypothesized that how precisely teachers (both treatment and control) provided the A2i recommended amounts would predict student reading outcomes (Connor, Piasta, et al., 2009).

METHOD

Participants and Random Assignment Procedures

Three hundred sixty-nine children in 25 first-grade classrooms from seven schools participated in this cluster randomized control field trial. Schools were located in an ethnically and economically diverse north Florida district. For the seven schools, matched pairs were created based on school-wide percentages of children qualifying for free or reduced price lunch (FARL), Reading First status, and third grade school mean performance on the Florida Comprehensive Achievement Test, the state-mandated test of reading (see Table 2). Because we had an odd number of schools, the middle ranked unmatched school was randomly assigned to the treatment or control condition. Thus, four schools, 11 teachers, and 174 students were assigned to the control condition and three schools, 14 teachers, and 222 students were assigned to the treatment condition. Schools ranged in percentage of children qualifying for FARL from 4 to 87%, and two schools were participating in Reading First. All schools used Open Court (https://www.sraonline.com/oc_home.html) as their core literacy curriculum (see also Crowe, Connor, & Petscher, 2009). All participating teachers were fully certified and had, at a minimum, a BA or BS degree. All of the first-grade teachers at the participating schools were invited to join the study, with 90% actually participating. No schools or teachers withdrew from the study after the onset of PD training. Except for years of experience, where control teachers had significantly more years of experience than did treatment teachers, there were no differences among treatment and control teachers on key characteristics (see Table 2).

All of the students in participating teachers' classrooms were invited to join the study and we were able to recruit approximately 86% of the students. Notably, 15% of children in the control and 14% of children in the treatment group were identified as eligible for special or exceptional student education (ESE, e.g., speech impairment, language impairment,

Table 2. Characteristics of participating schools (top) and teachers (bottom)

School	Group	School-Wide % FARL	Reading First
A	Treatment	87.0	Yes
B	Control	60.0	Yes
C	Treatment	38.0	No
D	Control	33.0	No
E	Control	12.0	No
F	Control	9.0	No
G	Treatment	4.0	No
Teachers	**Treatment**	**Control**	**Total**
Years of experience, *M* (*SD*)	14 (9.6)*	22 (8.6)	17 (10.0)
Fall teacher knowledge, *M* (*SD*)	23 (7.2)	29 (8.6)	25 (8.3)
Spring teacher knowledge, *M* (*SD*)	35 (7.4)	33 (5.6)	34 (7.4)
% Female	93%	100%	96%
% White	64%	73%	68%
% Black	36%	9%	24%
% with master's	29%	45%	36%

Note. Teacher Knowledge Score out of 45 possible. FARL = free and reduced lunch.
$^{*}p = .043$.

developmental disability, etc.). For this reason, we included their status in our models with 1 = ESE and 0 = not indentified for ESE, excluding children identified as gifted. Across groups, students were similar demographically. Fifty-four percent of the control and 51% of the treatment group were girls; 26% of the control and 32% of the treatment group were African American; 45% of both groups were White, and the remaining children across groups represented other ethnic groups (e.g., Hispanic, multiracial); 49% of the control students and 44% of the treatment students qualified for FARL. Ten percent student attrition, which is in line with mobility for the participating schools, was equally divided between treatment and control classrooms.

A subset of participating children was randomly selected from each classroom as target children for classroom observation coding. Students in each classroom were rank ordered by fall word reading scores, and 3 children from the top and middle and 4 from the bottom scores of the rankings were identified as target students. If there were fewer than 10 consented children in the classroom, all students were included as target children. This resulted in a sample of 234 students for whom classroom reading instruction was coded. Of these, 132 were in treatment classrooms and 102 in control classrooms. Multivariate comparisons of target versus nontarget students revealed no significant differences in word reading or vocabulary for fall, winter, and spring assessments, Wilks's $\Lambda = .995$, $F(4, 2458) = 1.691$, $p = .149$, for Season (fall, winter, spring) $\times$ Target Child = 1.

Assessments

Students' language and literacy skills were assessed in the fall, and again in the winter and spring, using a battery of language and literacy assessments, including tests from the WJ–III

(Mather & Woodcock, 2001). The WJ–III was selected because it is widely used in schools and for research. It is psychometrically strong for this age group (reliabilities on the tests used ranged .81–.94), and subtests are brief. All assessments were administered to children individually by a trained researcher in a quiet location near the students' classrooms.

We assessed students' letter and word reading skills using the WJ–III Letter-Word Identification subtest, which asks children to recognize and name increasingly unfamiliar letters and words out of context. Expressive vocabulary was assessed using the WJ–III Picture Vocabulary subtest, which asks children to name pictures of increasingly unfamiliar objects.

W scores, which are a variation of Rasch scores and thus have equal intervals, were used in models to evaluate the efficacy of the ISI intervention. The letter-word identification GE and the picture vocabulary AE scores were entered into the A2i software by research assistants. In combination with the end-of-year target outcome (fall reading GE + .9; minimum = 2.1 based on previous district mean), word reading and vocabulary scores were used in the computer dynamic system forecasting intervention algorithms to compute recommended amounts of TCM code-focused and CM meaning-focused instruction. Teachers also viewed these scores using the A2i software, including graphs that displayed children's progress over the school year for these measures (i.e., GE for word reading and AE for vocabulary).

The treatment group teachers received PD beginning in August 2006 and first gained access to A2i recommendations and assessment information in September 2006. They received training and used the software continuously through May 2007. The control group teachers were provided written reports of the assessments results for their students in the fall, winter, and spring of the study year and received the ISI PD and access to the A2i software the following school year.

Fidelity of Intervention Implementation

There are concerns related to randomized control field trials, which include ethical considerations, fidelity of treatment (Jo, 2002), and attrition (Cook & Payne, 2002; Gueron, 2002). As discussed, no teachers left the study, and student attrition levels provided no cause for concern. Ethical concerns were addressed as part of the research design inasmuch as ISI represented a research-based intervention, and all students' reading skills were monitored with scores provided to both treatment and control group teachers.

Evidence that teachers in the treatment group actually implemented the ISI regime was judged to be important for substantiating our claim that effective amounts and types of instruction for a particular child could be predicted. We assessed fidelity of implementation in two ways: A2i use and classroom observations. First, as part of the PD protocol, teachers' fidelity of implementation was monitored by examining the logs of A2i use, which were automatically generated by the software. Logs included the date and time teachers logged on, the pages they visited, and the amount of time spent on each page. Total amounts of time spent using A2i included all of the time for each session, from log on to log out or 1 min after the last click of the mouse if the teacher did not log out. Teachers' time included time spent viewing the online resources, using the discussion board, and accessing the assessment and planning features of A2i. In a previous study, conducted in 2005–2006 (Study 1), the more teachers used the software, the greater was their students' reading skill growth (Connor, et al., in press). In this study, we examined teachers' use regularly and, when they were not using the software, contacted them with inquiries as to why this might be. In this way, we were able to encourage and support A2i use throughout the school year. There were two teachers who did not meet the study goal of 120 min total A2i use because they were not comfortable with technology. These teachers were provided with a printout of their classroom view, which showed, for each student, the recommended amounts for each type of instruction, the recommended small groups, and reading and vocabulary scores. Our supposition was that it was providing the A2i recommended differentiated instruction in the classroom and not A2i use per se that would contribute to student reading gains.

Treatment teachers used A2i software, on average, for a total of 560 min from September 1 to May 31 (range = 12.5–1,927 min). This is considerably higher than A2i use in the previous study, when teachers used the software for an average of 180 min for the entire year.

To compare instruction in both treatment and control classrooms, we conducted videotaped classroom observations of the dedicated literacy block in the fall, winter, and spring using the ISI Coding System (Connor, Morrison, et al., 2009). These observations were conducted and subsequently coded by research assistants who were blind to teacher condition to the extent possible. Thus we were able to examine whether treatment group teachers were individualizing students' instruction to a greater or lesser extent than were the control teachers.

Observation dates were scheduled at the teachers' convenience and conducted by trained research assistants. Two digital video cameras with wide angle lenses were used to capture instruction. In addition, a research assistant recorded field notes, including descriptions of students and any activities that might be difficult to see on the camera (e.g., content of workbook pages; materials used at the teacher table).

Video was then digitized and coded in the laboratory using Noldus Observer software (Noldus Information Technology, 2001). Because we were interested in the extent to which teachers provided different amounts of types of reading instruction to students in the classroom, instruction was coded at the level of the individual student. Thus Child A might be at the teacher table with a small group of students working on phonological awareness (a small-group TCM code-focused activity) while Child B reads independently in the library corner (individual CM meaning-focused activity). The coding metric was duration (in minutes) of activities, where any activity lasting at least 15 s (.25 min) was coded. Noninstruction activities (e.g., transitions, disruptions, off-task) were also coded, so the times reported represent actual instruction duration stripped of transitions, disruptions, organization, and off-task behavior. Including the noninstruction activities was beyond the scope of this study.

Following the multiple dimensions of instruction conceptualization, each activity was assigned a content code (phonological awareness, print vocabulary, word decoding, etc.), a grouping code (whole class, small group, etc.), and a management code (TCM, CM, etc.). The major reading instruction codes are provided in Table 1, and the definitions for Word Identification/Encoding modifiers are provided in the appendix. The complete protocol and coding manual are available upon request from the first author.

All coders had to achieve a kappa of at least .70 with the lead coder before they were allowed to code independently. Reliability was checked for approximately 10% of the coded observations, and reliability ranged from .94 to .98 for duration (±10 seconds) and activity.

ANALYTIC STRATEGIES

Examining the Effect of Treatment

Because students were nested in classrooms and classrooms were, in turn, nested in schools, we used HLM (Raudenbush & Bryk, 2002). Because of the nested design, standard errors may be misestimated if shared classroom- and school-level variance among students is not considered. In addition, random assignment occurred at the school level and thus considering only student-level differences could possibly overestimate the effect of the intervention. Finally, HLM allows examination of cross level interactions (e.g., Child × School interactions), and therefore we were able to test whether the treatment was more or less effective for children with different language and literacy skill, or who qualified for special education services. To test the within level interaction (i.e., Treatment × School SES), the variable was created in SPSS software and entered into the model at Level 3. However, power was limited for testing these interactions and so we computed effect sizes (*d*) as well (Rosenthal & Rosnow, 1984).

Examining Teacher Instruction and Fidelity

Instruction variables represent the sum of the number of minutes of the component activities (e.g., comprehension, phonological awareness) for each student that were judged to be code focused or meaning focused as observed during the literacy block (see Table 1). To examine whether treatment teachers provided A2i recommended amounts of TCM code-focused and CM meaning-focused instruction more or less precisely than did control teachers, we computed distance from recommendations (DFR) for each student for each season for each type of instruction (small-group TCM code-focused, etc.). Because classrooms were observed three times per year and A2i recommended amounts changed monthly and when students were reassessed in the winter, we used HLM latent growth curve modeling with TCM code-focused and CM meaning-focused amount variables (amount and DFR) entered as time varying covariates (Raudenbush & Bryk, 2002). DFR for each instruction activity type (TCM code- and CM meaning-focused) was used first as an outcome to examine differences between treatment and control groups. DFR variables were then entered as time varying covariates with reading as the outcome in latent growth curve models to examine whether the DFRs predicted students' reading skill growth.

RESULTS

Effect of the ISI Intervention on Gains in Students' Word Reading Skills

Three-level HLM models with fall scores as the outcome revealed that there were no significant differences between treatment and control groups in students' fall reading and vocabulary scores (see Table 3 for means and standard deviations by group). Examining standard scores (standardized $M = 100$, $SD = 15$) for fall and spring revealed that, on average, both groups demonstrated grade-appropriate gains in word reading and vocabulary scores but there was variability (see Table 3). Notably, ranges for spring word reading standard scores were wide, ranging from 63 (*very low*) to 141 (*very high*) for both groups.

Examining the unconditional HLM model with spring word reading W score as the outcome and no predictor variables revealed that the intraclass correlation (ICC), the proportion of between-school variance, was .06, school-level variance = 38.05, $\chi^2(6) = 30.77$, $p < .001$. There was no significant classroom-level variance, variance = .78, $\chi^2(18) = 18.97$, $p = .393$. Student-level variance in the unconditional model was 641.19.

To build the final model, we added fall word reading and vocabulary W scores as covariates, centered at the grand mean for the sample, the other covariates, and then

Table 3. Fall and spring achievement W scores and standard scores (SS) by treatment and control condition

	M		SD		Minimum		Maximum	
Treatment	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
WJ Letter-Word W	417.41	464.86	29.64	24.98	340	371	511	514
WJ Letter-Word SS	107	112	16	14	67	66	150	141
WJ Vocabulary W	481.39		9.23		456		506	
Control								
WJ Letter-Word W	417.61	461.20	32.01	27.45	314	345	507	512
WJ Letter-Word SS	108	112	15	14	64	63	144	141
WJ Vocabulary W	481.59		14.87		448		510	
Total (used in HLM)								
WJ Letter-Word W	416.68	463.22	30.84	26.09				
WJ Vocabulary W	481.19		9.55					

Note. WJ = Woodcock–Johnson; HLM = hierarchical linear models.

the interactions. By grand mean centering covariates, the intercept may be interpreted as the fitted mean for children whose initial status falls at the mean of the sample. We added a variable for ESE status (i.e., receiving special education services) where children identified as ESE were coded 1 and all other children, including gifted, were coded 0. At the school level, we added the percentage of students school-wide who qualified for FARL. Accordingly, the intercept represents the fitted mean for children who were not ESE, who had typical fall reading and vocabulary scores, and who attended a school where about 50% of the children qualified for FARL. Adding covariates was not done to control for group differences but rather to remove variance in the outcome variable that was not of primary interest. This increases power to find differences between groups (Venter, Maxwell, & Bolig, 2002). In addition, by controlling for initial status, we could compare group gains (i.e., residualized change). The treatment variable (assignment to the treatment condition = 1; assignment to the control condition = 0) was added at the school level because random assignment was at the school level. Because there was no significant classroom-level intercept variance, adding this variable at the classroom level would yield the same results. The coefficient for the treatment variable represents the fitted mean difference between school/classroom groups in students' word reading W scores. Finally, the cross-level and within-level interactions were added to the model. The equation for the final model is provided in Table 4. Of note, in all but the final model, there was significant classroom-level variance for ESE status, suggesting that some classrooms included more ESE students than did others. In addition, there was no significant between school variance when school SES was added to the model, $\chi^2(3) = 3.64$, $p = .302$.

There was a significant effect of treatment (see Table 4). That is, students whose teachers implemented the ISI regime demonstrated significantly greater gains in word reading scores than did students whose teachers and schools were in the control group and conducted literacy instruction as usual. The effect size (d = treatment coefficient/15.8, which is the child level standard deviation) was .50, which is a moderate effect (Rosenthal & Rosnow, 1984). Considered another way, the more than 7-point difference between the ISI intervention and control group represents a 2-month difference in grade level, based on equating the fitted mean for the control group students (459 = 2.6 GE) and the intervention group students (467 = 2.8 GE) where .9 represents a 9-month school year.

Table 4. Final three-level model testing the effect of the ISI Intervention (TREAT060) on the outcome, spring word reading score, controlling for school-wide socioeconomic status (SES; % of students qualifying for FARL, FARLPERC), students' special education status (ESE), and fall Vocabulary (WJ_VOC_W) and Word Reading scores (WJ_LW_W)

LEVEL 1 MODEL

$$\text{SP\_LW\_W}_{ijk} = \pi_{0jk} + \pi_{1\text{jk}}(\text{WJ\_LW\_W}_{ijk} - \overline{\text{WJ\_LW\_W}}\ldots)$$
$$+\,\pi_{1jk}(\text{WJ\_VOC\_W}_{ijk} - \overline{\text{WJ\_VOC\_W}}\ldots) + \pi_{3jk}(ESE_{ijk}) + e_{ijk}$$

LEVEL 2 MODEL

$$\pi_{0jk} = \beta_{00k} + r_{0jk}$$
$$\pi_{1jk} = \beta_{10k}$$
$$\pi_{2jk} = \beta_{20k}$$
$$\pi_{3jk} = \beta_{30k} + r_{3jk}$$

LEVEL 3 MODEL

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(\text{FARLPERC}_k - \overline{\text{FARLPERC.}}) + \gamma_{002}(\text{TREAT060}_k)$$
$$+\,\gamma_{003}(\text{TREATXSE}_k - \overline{\text{TREATXSE.}})$$
$$\beta_{10k} = \gamma_{100} + \gamma_{101}(\text{TREAT060}_k)$$
$$\beta_{20k} = \gamma_{200} + \gamma_{201}(\text{TREAT060}_k)$$
$$\beta_{30k} = \gamma_{300} + \gamma_{301}(\text{TREAT060}_k)\_u_{30k}$$

Fixed Effects	Coefficient	*SE*	*df*	*p*
Intercept (Fitted *M* spring word reading W score)	459.58	1.91	25	<.001
Child-level variables				
Fall word reading	.64	.04	386	<.001
Fall vocabulary	.19	.13	386	.171
ESE status = 1	−6.95	5.63	5	.272
School-level variable				
Treatment group = 1	7.84	3.19	25	.021
School-wide SES	.02	.06	25	.698
Child × School Level Interactions				
Treatment × Read	−.07	.06	386	.236
Treatment × Vocab	−.11	.19	386	.550
Treatment × ESE	−4.30	7.18	5	.575
School × School Level Interaction				
Treatment × SES	−.09	.07	25	.234

Random Effects	*Variance*	*df*	χ^2	*p*
Student level (e_{ijk})	251.34			
Classroom-level (r_{0jk})	1.56	19	21.46	.311
ESE (r_{3jk})	110.23	14	22.93	.061
School-level ESE (u_{30k})	.62	4	7.70	.102

Note. The school-wide SES × Treatment variable (TREATXSE) tests the effect of the ISI intervention as a function of school-wide SES. All continuous variables are grand mean centered. *SD* for outcome at the student level = 15.85. FARL = free and reduced lunch.

Child × ISI Interactions

To examine whether the ISI intervention was more or less effective for children who had varying levels of fall vocabulary and reading skill or who were identified as eligible for ESE (1 = identified, 0 = not identified or gifted), we entered the school level treatment variables into each Level 3 equation (see Table 4). None of the Child Characteristic × Treatment interactions was significant. However, as noted previously, the study is underpowered, and so it is possible that there were interactions that were not possible to detect. Thus, we computed effect sizes (*d*). The ISI × ESE coefficient for children receiving ESE services did not reach levels of significance but was, nonetheless, fairly large. The negative effect size (*d*) –.27 suggests that ISI might be somewhat less effective for children who are receiving ESE services.

To compute effect sizes for the impact of the ISI intervention compared to control for schools with greater or smaller percentages of students qualifying for FARL we compared effect sizes for schools that fell at the penultimate maximum and minimum of our sample (60% and 9%), which are at the 25th and 75th percentile of the sample. The effect size was .44 comparing schools where 60% of the children qualified for FARL and .73 for schools where only 9% of the children qualified for FARL. Thus, there is some indication that the intervention was more effective at more affluent schools. Nevertheless, the effect size for the higher poverty school was still fairly close to the overall effect size for the study (.50).

We also computed effect sizes for students whose fall word reading and vocabulary scores fell at the 25th and 75th percentiles. For word reading, the treatment effect size for students with lower fall reading (W = 393) was .59, whereas for students with stronger fall scores (W = 435) the effect size was slightly smaller (d = .41). The same was found for vocabulary; the effect size was .54 when children had weaker fall vocabulary scores (W = 474) and was slightly smaller (.45) for children with stronger vocabulary scores (W = 487) in the fall.

Comparing Implementation of ISI Instruction Across Conditions

In our coding system, instruction can occur at the classroom level, where all the students are receiving substantially the same instruction, or at the student level, where students within the classroom are receiving different amounts and types of instruction. With regard to classroom-level instruction (i.e., whole-class instruction), HLM latent growth curve modeling, with amount of each type of instruction observed modeled over time (fall, winter, spring) at Level 1 and treatment condition entered at Level 2, revealed no significant differences between the treatment and control classrooms except for TCM meaning-focused instruction. Generally, treatment group teachers spent significantly less time than did control teachers teaching whole-class TCM meaning-focused instruction; this was a large effect (d = 2.83), with treatment teachers spending about 8.7 min per day compared to 20.2 min per day for control teachers by the end of the year. Mean mounts of classroom-level instruction (minutes/day) are as follow: TCM code focused = 14.58 (*SD* = 11.45); CM meaning focused = 7.40 (10.94); CM code focused = 1.86 (*SD* = 5.07).

Both treatment and control teachers provided student-level small group and individual instruction but, according to HLM results, the amounts provided varied by condition for TCM code-focused and CM meaning-focused instruction and time of year (see Figures 4 and 5). Teachers in the treatment condition generally provided students more time in TMC code-focused and CM meaning-focused student level instruction than did teachers in the control condition, but the differences were small for TCM code focused instruction (see Figure 4). Figure 5 compares treatment and control teachers for the components of each of the four types of instruction (see Table 1 for which components were considered code and which meaning focused).

However, just providing more time in small-group and individual instruction to students is not the ISI intervention. Rather, teachers were asked to teach the A2i recommended

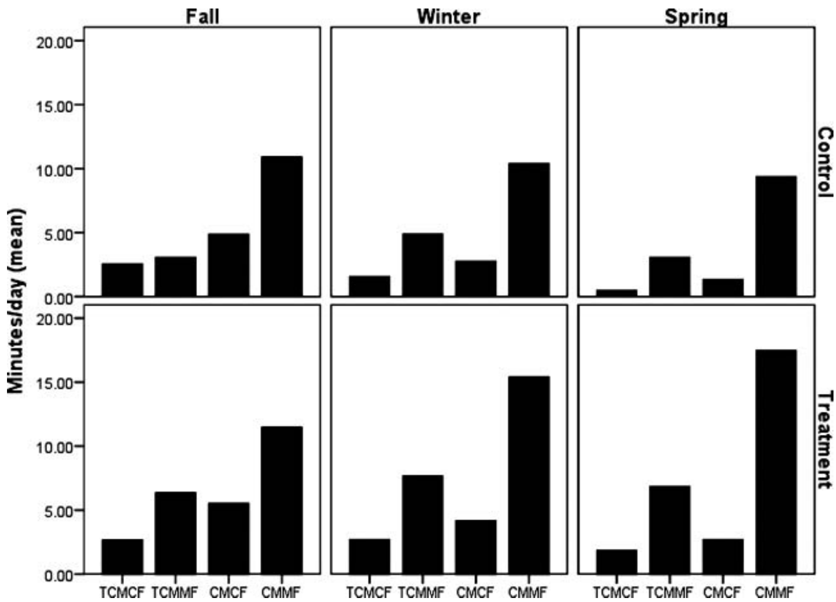

Figure 4. Mean observed amounts (minutes/day) of small-group and individually provided instruction for students in treatment and control classrooms for fall, winter, and spring observations. *Note.* Types of instruction include teacher/child-managed code focused (TCMCF), teacher/child-managed meaning focused (TCMMF), child-managed code focused (CMCF), and child-managed meaning focused (CMMF).

amounts of TCM code-focused and CM meaning-focused instruction computed using dynamic forecasting intervention model algorithms that rely on the correlational evidence of CXI interactions. To suggest that CXI interactions are causally implicated in student achievement, the treatment teachers should have been more precise in providing the A2i recommended amounts than the control teachers, and this, in turn, should predict students' gains in word reading skills. To compare the precision with which teachers implemented differentiated instruction following the A2i recommendations, we computed the DFR for each student for each observation. The DFR is the absolute value of the difference between the observed total amount (in minutes) of instruction activity type, that is, TCM code-focused and CM meaning-focused activities, and the A2i recommended amount. Figure 6 compares the observed and recommended amounts for target children in one treatment and one control classroom. So, for example, if Child AT receives 10 min, on average, of TCM code-focused instruction in either small group or individually, and the A2i recommended amount is 15 min, his DFR would be 5 min. Some teachers provided more than the recommended amounts and some less. For example, differences in TCM code-focused amounts recommended and observed ranged from –26.17 to 9.30 min. However, the negative mean, –11.61 ($SD = 6.09$) indicated that, on average, teachers were not providing the recommended amounts. The range was wider for CM meaning-focused differences (–45.70 to 30.04 min). Again, the mean was negative, –21.83 ($SD = 13.58$).

In 3-level HLM latent growth curve models, repeated observations of DFR at the student level were modeled over time, with individual students modeled at Level 2 and the treatment variable entered at Level 3, the classroom level. Results revealed that treatment teachers had significantly smaller DFRs for TCM code-focused and CM meaning-focused

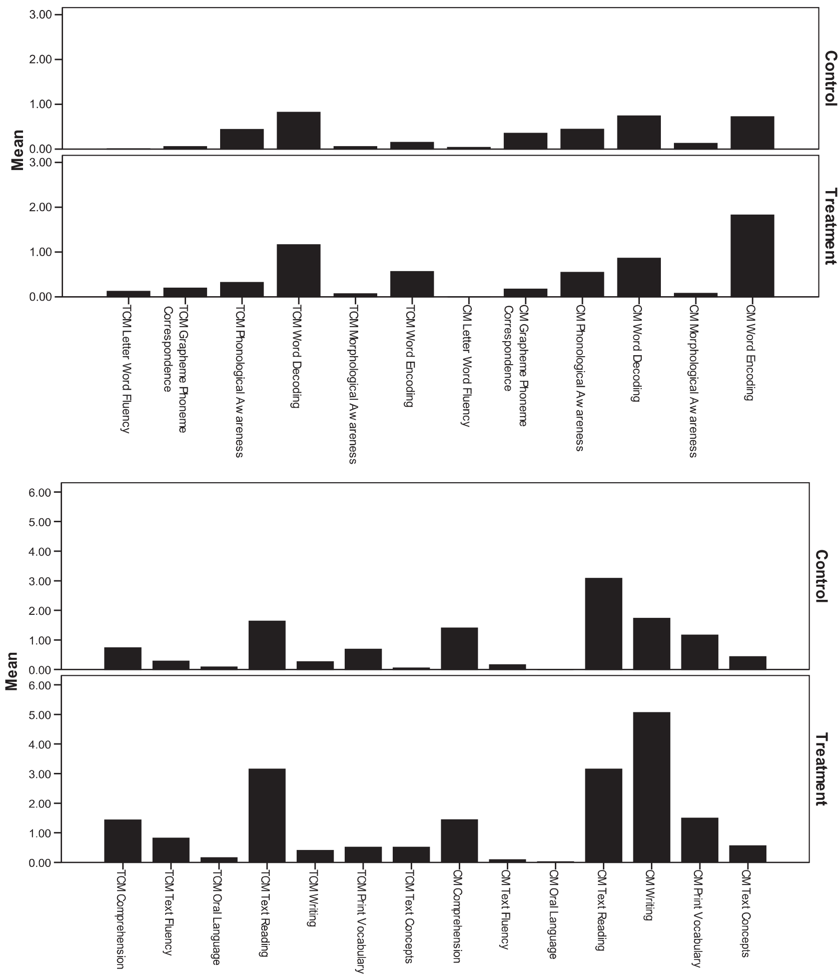

Figure 5. Student-level mean amounts (in minutes) fall, winter, and spring of subcomponents of instruction with teacher/child-managed (TCM) and child-managed (CM) code focused (top) and TCM and CM meaning focused (bottom) by condition.

instruction (see Table 5) compared to control teachers. In addition, for CM meaning-focused instruction, treatment teachers became increasing more precise over time compared to control teachers.

To further explore the possibility that the ISI intervention might be less effective for children designated as ESE because it was more difficult to meet the greater times generally recommended for these children, considering only students in treatment classrooms, we asked if there were significant differences in DFR as a function of ESE designation. Results of HLM revealed no significant differences in TCM code-focused and CM meaning-focused DFR over time as a function of ESE designation. There was a trend (.064) that teachers became increasing precise over time in meeting CM meaning-focused recommendations for ESE students compared to non-ESE peers.

We conducted similar analyses for school SES. TCM code-focused DFR significantly increased as schools served greater numbers children who qualified for FARL (coefficient =

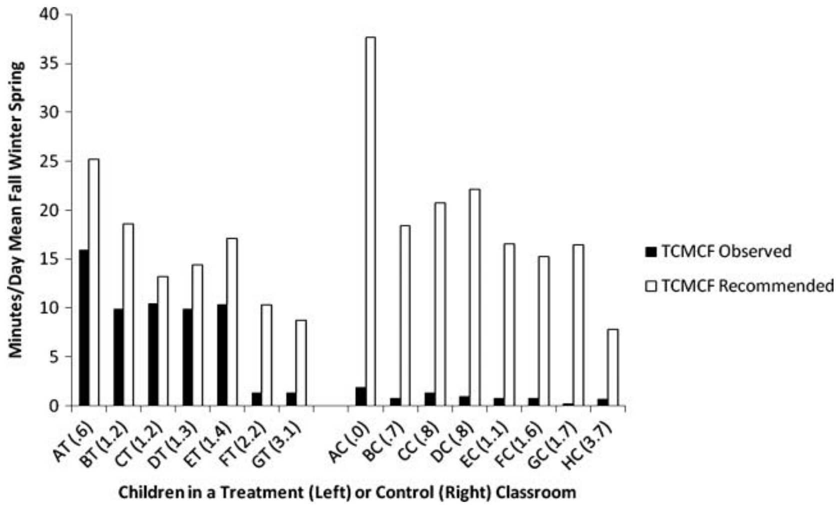

Figure 6. Mean fall, winter, and spring observations: A treatment (left) versus a control (right) classroom comparing observed provided amount of teacher/child-managed code focused (TCMCF) with recommended amount for each target student. *Note.* The fall word reading grade equivalent (GE) is provided in parentheses next to the child's name abbreviation. The recommended amounts changed in the winter depending on each student's progress (e.g., child CT made greater gains than child BT, hence the recommended amount is smaller even though they had the same GE in the fall). The mean GE fall-to-spring gain for the treatment students was 1.4; the GE gain for the control students was 1.0. This represents a 4-month GE advantage for students in this treatment classroom.

.053, $p = .010$). That is, teachers at higher poverty schools were less likely to provide the A2i recommended amounts of TCM code-focused instruction. For CM meaning-focused DFR, there was no significant effect of school SES (coefficient = –.05, standard error [*SE*] = .074, $p = .470$). Again, cautions about power to support null findings are appropriate. The trend would suggest that children at higher poverty schools would tend to have smaller DFR for CM meaning-focused instruction.

To examine whether TCM code-focused and CM meaning-focused DFR predicted students' word reading skills, centering time at the April observation with word reading as the outcome, we entered the DFR variables as time varying covariates. We did this because the DFRs changed over the course of the school year and depended on students' current word reading and vocabulary scores. Results and the model are provided in Table 6. Students with smaller DFRs for TCM code-focused and for CM meaning-focused instruction had significantly greater word reading scores by spring than did students with larger DFRs. That is, the more precisely students received the A2i recommended amounts, the better were their word reading skills by the end of the year. The effects were large, especially for TCM code-focused DFR. By the end of the year, a child who received a DFR of 6.5 min/day, which is 1 *SD* below the overall mean of 12.5 (and the mean amount provided by treatment teachers) would be predicted to achieve a word reading score of 478 W points compared to a child who received a DFR of 18.5, which is 1 *SD* above the mean, who would be predicted to achieve a word reading score of 451 W points. This 27-point difference is an effect size (*d*) of 2.3 (27/11.6). The effect size (*d*) for CM meaning-focused DFR comparing 1 *SD* above and below the mean (13–35) is .5.

Table 5. Fitted means in spring (from HLM analyses) DFR, in minutes/day, by treatment and control condition for TCM code focused and CM meaning focused

$$Y_{tik} = \pi_{0ik} + \pi_{1ik}{}^*(\text{Month Observed}_{tik}) + e_{tik}$$
$$\pi_{0ik} = \beta_{00k} + r_{0ik}$$
$$\pi_{1ik} = \beta_{10k} + r_{1ik}$$
$$\beta_{00k} = \gamma_{000} + \gamma_{001}{}^*(\text{Treatment}_k) + u_{00k}$$
$$\beta_{10k} = \gamma_{100} + \gamma_{101}{}^*(\text{Treatment}_k)$$

Y_{itk}	TCM Code Focused	CM Meaning Focused
Treatment DFR ($\gamma_{001} + \gamma_{000}$)	6.54[a,*]	27.31[a,**]
Control DFR (γ_{000})	8.37	34.03
Treatment DFR slope ($\gamma_{101} + \gamma_{100}$)	−1.61	1.27[a,***]
Control DFR slope (γ_{100})	−1.46	2.15
Random Effect	**Variance (*df*)**	**Variance (*df*)**
r_0	Fixed	1.37 (209)
r_1	.39 (233)***	Fixed
e	10.07	70.61
u_{00}	3.38 (23)***	27.65 (23)***

Note. DFR TCM code focused (HLM descriptive), $M = 12.47$ ($SD = 5.93$), range = 0–48; DFR CM meaning focused, $M = 24.57$ ($SD = 10.98$), range = 0–45. HLM = hierarchical linear models; DFR = distance from recommendation; TCM = teacher/child managed; CM = child managed.
[a]Significant difference between treatment and control.
$^*p < .05$. $^{**}p \leq .01$. $^{***}p \leq .001$.

To rule out that the impact of DFR was related to time on task, so to speak, rather than the precision with which the A2i recommendations were met, we repeated the analysis but instead included the total amounts of TCM code-focused and CM meaning-focused instruction (minutes). Results showed that greater amounts of small-group TCM code-focused instruction negatively predicted students' spring word reading scores, coefficient = –.43 (.20), $p = .035$. However, greater amounts of CM meaning-focused instruction positively predicted spring word reading scores, coefficient = .25 (.06), $p < .001$. Thus, it is unlikely that the impact of TCM code-focused instruction is simply time on task, but this might be the case with CM meaning-focused instruction.

Considering only treatment students and teachers, we found that DFR varied significantly between classrooms. The treatment teachers' ICC for TCM code-focused DFR (using growth models described previously; repeated measure over time centered in spring with DFR as time varying covariates) was .36. That is, 36% of the variability in students' TCM code-focused DFR fell between classrooms. The ICC for CM meaning-focused DFR was .51. DFR predicted word reading scores. In the HLM model for treatment students with DFRs as time varying covariates, the intercept was 489.18 (13) with a slope of 5.93 (13), $SE = .59$, $p < .001$. For TCM code-focused DFR, the HLM coefficient was –1.64 (367), $SE = .35$, $p < .001$; for CM meaning-focused DFR, the coefficient was –.26 (367), $SE = .08$, $p = .002$. Only student-level variance ($r_0 = 422.04$) was significant ($p < .001$). Thus, fidelity within the treatment group also predicted students' word reading outcomes.

Table 6. HLM results of TCMCF and CMMF DFR (minutes/day) predicting word reading where Y_{tik}, the predicted word reading outcome for child *i* at time *t* in classroom *k*, is a function of the time observed centered at spring (Month Observed_{tik}), where slope is minutes change per month, and the TCMCF DFR and CMMF DFR for each student

$$Y_{tik} = \pi_{0ik} + \pi_{1ik}{}^*(\text{Month Observed}_{tik}) + \pi_{2ik}{}^*(\text{TCMCF DFR}_{tik}) + \pi_{3ik}{}^*(\text{CMMF DFR}_{tik}) + e_{tik}$$
$$\pi_{0ik} = \beta_{00k} + r_{0ik}$$
$$\pi_{1ik} = \beta_{10k}$$
$$\pi_{2ik} = \beta_{20k} + r_{2ik}$$
$$\pi_{3ik} = \beta_{30k}$$
$$\beta_{00k} = \gamma_{000} + u_{00k}$$
$$\beta_{10k} = \gamma_{100}$$
$$\beta_{20k} = \gamma_{200}$$
$$\beta_{30k} = \gamma_{300}$$

Fixed Effect	Coefficient	*SE*	*df*	*p*
Intercept (γ_{000})	493.36	3.09	24	<.001
Slope (γ_{100})	5.09	0.316	645	<.001
TCMCF DFR (γ_{200})	−2.25	0.151	233	<.001
CMMF DFR (γ_{300})	−0.26	0.064	645	<.001

Random Effect	Variance	*df*	χ^2	*p*
Intercept, r_{0ik}	23.98	207	591	<.001
TCMCF slope, r_{2ik}	0.14	231	284	<.001
Level 1, e_{tik}	136.08			
Intercept, u_{00k}	34.03	24	44.3	.007

Note. Deviance = 5548.10. TCMCF = teacher/child-managed code focused; CMMF = child-managed meaning focused; DFR = distance from recommendation.

DISCUSSION

The purpose of this cluster randomized control field trial was to examine the causal implications of Child × Instruction interactions on student reading achievement. Thus we examined the efficacy of the ISI intervention, which was designed to incorporate observed Child × Instruction interactions (i.e., the ISI intervention), compared to literacy instruction that did not explicitly consider Child × Instruction interactions. There was a significant intent-to-treat effect. That is, students whose teachers participated in the ISI intervention made greater gains in word reading skill growth than did students in control classrooms. Students in ISI intervention classrooms achieved about a 2-month advantage in end-of-year word reading skills compared to the students in the control condition ($d = .50$).

Using classroom observations, where instruction was coded for individual students, we found that students in the ISI intervention group were more likely to receive the A2i recommended amounts of TCM code-focused and CM meaning-focused instruction, which are based on Child × Instruction interactions, than were control students. Plus, how precisely the observed amounts matched the A2i recommended amounts (see Figure 6) strongly predicted students' word reading outcomes with an effect size (*d*) of more than 2.0. For word reading (using the entire sample), a student with a DFR of 6.5 min/day, which is 1 *SD* below the mean of 12.5 min/day, would be expected to achieve a third-grade reading GE (3.2). In contrast, a student with a TCM code-focused DFR of 18.5 min/day, 1 *SD* above the mean, would end the year a full grade level behind (GE = 2.2). Keep in mind that our metric is minutes per day, and there are by conservative count about 150 instructional school days per school year. Thus a difference in DFR of only 2.5 min per day between treatment and control students actually represents about a 375 min or 6.5 hr closer to recommended amounts over the school year. A tenet of dynamic system theories (Yoshikawa & Hsueh, 2001) is that small differences can have large effects. The differences in DFR for TCM code focused and CM meaning focused are large enough to begin to explain the intent-to-treat effect observed for word reading gains (intent to treat effect = 7.84 W points; DFR treatment effect = 5.9 W points). Taken together, and keeping in mind that all teachers used the same curriculum during a dedicated literacy block, these findings indicate that Child × Instruction interactions are very likely causally implicated in first graders' response to the reading instruction they receive and that children with different profiles of vocabulary and reading skills will require different instructional regimes to reach their potential.

With the right information, taking into account documented sources of influence (e.g., instruction, home support) and constraints (e.g., previous achievement, lack of resources, documented and undocumented disabilities), we can do a better job of predicting what amounts and types of reading instruction are going to be most effective for children (Landry, Antony, Swank, & Monseque-Bailey, 2010; Torgesen, 2000). We have called the equations that predict effective instructional regimes dynamic system forecasting intervention models. The underlying assumption is that with reliable estimates of children's language and literacy skills (and perhaps other attributes such as social skills, self-regulation, or behavior), we can do a better job of predicting students' reading skill development or their potential trajectories of growth (see Figure 7; Raudenbush, 2005). Taking this a step further, we argue that carefully designed reading instruction, which is individualized taking into account children's current reading and language skills, and the ways in which these skills interact with instruction, can influence the projected path of learning (dotted lines in Figure 7, left) so that children reach the highest levels of reading skill within their potential trajectories of achievement. By monitoring students' reading skills with valid and reliable assessments, we can make better predictions regarding their end-of-grade outcomes and can redesign planned instruction so as to influence the achievement trajectory upwards. Later in the school year, the potential trajectories narrow, reflecting more precise prediction of reading outcomes (see Figure 7, right).

In this study, we employed a dynamic forecasting intervention model of reading instruction (see Figure 1) that integrated elements of complex systems theory (Yoshikawa & Hsueh, 2001) and ecological theory (Bronfenbrenner & Morris, 2006) along with current theories of reading (Rayner et al., 2001; Scarborough, 2001; Snow, 2001). Components of complex systems are interrelated, and these relations may be nonlinear. As noted, small changes can have large effects and large changes can have small effects (Buell & Cassidy, 2001). For example, missing recommended amounts of instruction each day by a small amount appeared to have large effects on student reading skills by the end of first grade. Plus, instruction in the classroom was not considered to be a closed system isolated from and independent of the environment (Bronfenbrenner & Morris, 2006). Rather, the model considered instruction as conducted in a classroom with a teacher (and sometimes an aide) and an average of 20 students, all of whom were interacting—in the case of this study—around the teaching and learning of a complex skill, reading. This is why coding of classroom activities was conducted at the level of the individual student. Moreover, we assumed that children's language and reading skills were directly and indirectly affected by home and community influences (Connor et al., 2005; NICHD-ECCRN, 2004). As a result, children brought different aptitudes and skills to the classroom that impacted the

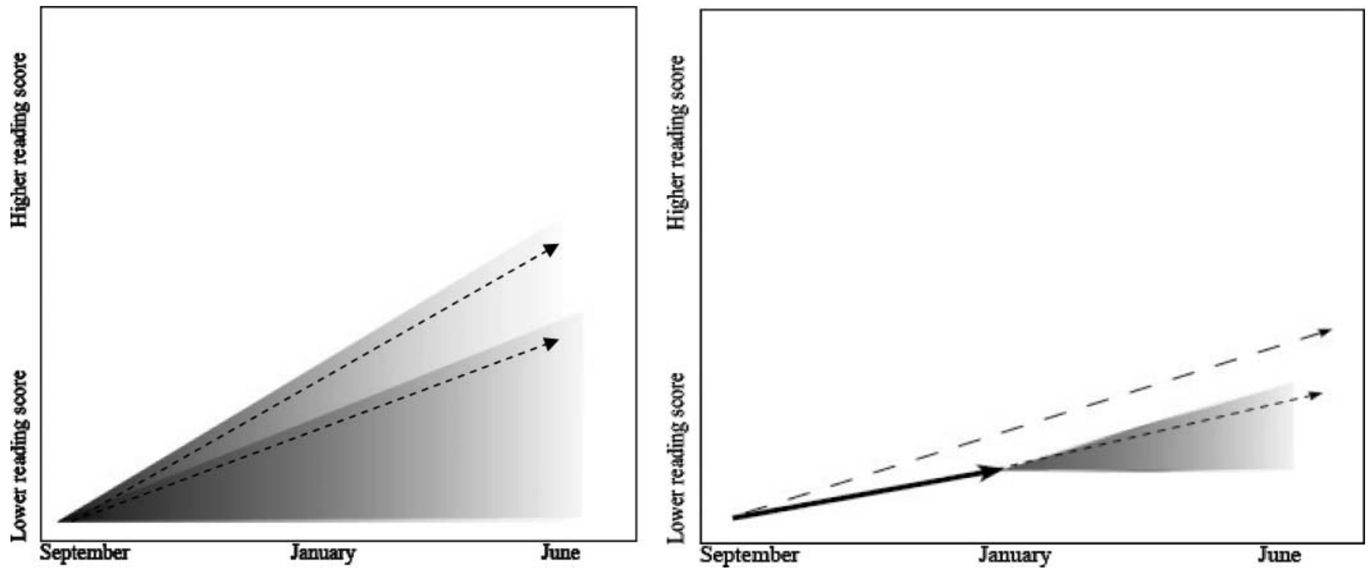

Figure 7. Left: Beginning-of-the-year potential achievement trajectories (cone) for two children where the child with the smaller cone suggests a smaller range of potential trajectories (e.g., dyslexia). The dotted lines represent the target trajectories for each. Right: Midyear achievement potential trajectories. The solid line represents the observed achievement. The dotted line represents the revised target trajectory. *Note.* Less-than-predicted achievement can preclude attainment of original target outcome and reduce achievement potential.

types of reading instruction activities that were most likely to promote their optimal levels of reading achievement. The A2i dynamic systems forecasting models exemplify the theory of reading instruction we employed in this study. These algorithms use information from multiple sources to predict specific amounts and types of reading instruction that should, theoretically, lead to stronger student reading skills.

A key assumption in our models is that students' initial reading and vocabulary skills reflect the influence of measured and unmeasurable sources of influence. This might include the effect of previous instruction received (e.g., during preschool and kindergarten), the influence of the home and community environments, children's health and well-being, potential and identified disabilities, and genetic influences. Thus, the potential growth trajectory for a child with a familial history of dyslexia or an identified disability might be lower than for a child living in an impoverished neighborhood, even if both have similar reading and vocabulary scores at the beginning of first grade (see Figure 7, left). Our results reveal that well-designed and implemented instruction should still be able to deflect the trajectory of achievement upwards for children, including those with disabilities. Moreover, such differentiated instruction should be as effective for children starting first grade with weaker vocabulary and word reading skills as it is for children with stronger skills.

Limitations

There are limitations to this study that should be considered. First, this was designed as an efficacy study. That is, we aimed to compare the ISI intervention implemented as fully as possible with business as usual instruction. Thus, claims that ISI might be as effective without, for example, the PD protocol or the A2i software are not warranted based on the results of this study. Nor can we claim that widely implemented (i.e., scaled up) the ISI intervention effects on teachers' instruction and student outcomes would be realized. For example, all of the teachers used the Open Court reading curriculum. Hence, these results might not generalize to students participating in other core curriculums. It may be that Open Court lent itself to differentiating instruction and replicating the study with different curriculums would have different results. Arguably, the effect might be smaller with a core that focused principally on meaning-focused or whole-class instruction. At the same time, effects might be larger when using a core that is designed to be differentiated and provides a balance of code-focused and meaning-focused activities.

The A2i algorithms themselves certainly can be refined and improved to increase the effectiveness of instruction. We consider them to be a first estimate only, and we are currently testing new algorithms to predict amounts of TCM meaning-focused and CM code-focused instruction. It is possible that the ISI effect was related to time on task rather than the precision with which children received the recommended amounts of instruction. Our results indicate that it is unlikely that TCM code-focused instruction was only time on task, but the results for CM meaning-focused instruction are equivocal. Still, even with these fairly rudimentary models, our findings support the use of more complex, integrated, and dynamic models of the underlying systems involved in reading acquisition and instruction.

We selected a measure of word reading as the focus of the target outcomes in the A2i forecasting intervention models because the previous correlation studies used a word reading outcome. However, it would have been more informative to use additional measures of reading, such as text comprehension, as an outcome and this will be done in future studies.

With only seven schools, this study was underpowered to find potentially important educational effects, particularly for Child $\times$ Treatment interactions. Hence, the ESE $\times$ Treatment and the School SES $\times$ Treatment interactions found to be nonsignificant might still be important especially considering the differences in treatment effect sizes for students who received ESE and those who did not and between the higher poverty and more affluent schools.

Implementing ISI in the Classroom

Coupled with our previous findings (Connor, Morrison, Fishman, et al., 2007; Connor, Piasta, et al., 2009) and others (Al Otaiba et al., 2010; Landry et al., 2010; Mathes et al., 2005), these results support the use of practices that consider, and are responsive to, individual differences among children; that are systematically differentiated in the amounts, content, and level (i.e., difficulty) of instruction provided based on students' vocabulary and reading skills; and that are explicitly informed by active use of valid and reliable assessments throughout the school year. Returning to the five components of the ISI intervention, (a) conceptualizing reading instruction across multiple dimensions, (b) student assessment, (c) A2i software, (d) PD, and (e) implementation during a dedicated literacy block utilizing flexible homogeneous, skill-based small-group instruction, probably all components contributed to the effects of the ISI intervention on students' word reading. Specifically the PD we provided to teachers in the treatment group taught teachers to organize reading instruction activities across the dimensions of grouping, management, and content, how to use assessment to guide instruction with attention to the A2i recommended amounts as well as appropriate skill level, and classroom management. We used PD methods generally found to be effective in the extant literature (Bos et al., 1999; Vaughn & Coleman, 2004) including the use of online resources, videos of master teachers, regular classroom support, and monthly meetings to support communities of practice (Bos et al., 1999). It is not unreasonable to suggest that schools and districts should be able to provide this level of support. It is not clear to what extent the A2i software itself facilitated teacher PD, but it likely contributed to the efficacy of the PD and to teachers' practice. Key is that it explicitly translated assessment results into specific recommendations for small-group instruction. Other instructional decision support systems in education show similar promise (Landry et al., 2010) as do similar systems in the medical field (Garg et al., 2005; Kawamoto, Houlihan, Balas, & Lobach, 2005).

Overall, the students in this school district were receiving effective instruction, as evidenced by their generally strong gains in reading and vocabulary. Teachers in both treatment and control conditions were observed to use classroom-level (e.g., whole class) and student-level (e.g., small group, individual) instruction during a dedicated block of time devoted to literacy. Thus, our control teachers were differentiating instruction to at least some extent. Both groups had access to study and school assessment results, were expected to provide small-group instruction, and every school had a literacy coach.

However, only the treatment teachers had access to the full ISI intervention, including the A2i software and PD, which appears to have changed their practices. Based on teachers' A2i use and classroom observations, teachers implemented ISI with good fidelity overall. For example, on average from fall to spring, Treatment teachers achieved a fitted TCM code-focused DFR of only 6.6 min/day, which falls within the expectation of the protocol that teachers provide ± 5 min/day from the actual A2i recommended amount. Teachers' use of the A2i software was substantially greater than was observed in a prior study (Connor, Morrison, Fishman, et al., 2007). We conjecture that this was the result of (a) improvements to the A2i software and (b) our PD efforts. With regard to the A2i software, based on feedback from teachers in our first study, we added progress monitoring graphs and moved the paper-based PD materials online. This included a discussion board and videotape clips illustrating master teachers' use of A2i, centers, and small-group instruction strategies.

Although the overall amount of PD was not increased, we began classroom-based PD in the fall, rather than in the winter. Again, this was based on teacher feedback. The teachers reported that the classroom-based PD and one-on-one help with the software were most helpful in supporting their efforts to implement the ISI intervention.

There are clear challenges to implementing the ISI regimen in the classroom, in addition to the many well-documented challenges teachers face, such as lack of time and resources (Cohen, Raudenbush, & Ball, 2003). Among the treatment teachers there were varying levels of fidelity, which impacted their students' achievement. Indeed, within the treatment group, individual teachers' mean (fall, winter, spring) TCM code-focused DFR ranged from 0 to 26 min. Implementing ISI requires enhanced responsiveness to students' instructional needs based on assessment results, masterful classroom planning and organization, and the provision of differentiated instruction in line with students' reading and language skills. An examination of the observation tapes, A2i use, and results from previous studies (Connor, Piasta, et al., 2009; Connor et al., in press) suggest that breakdowns in any of these skills are associated with less teacher fidelity and weaker student reading outcomes. For example, when the amount of time teachers used the various components of A2i was analyzed, results showed that the more time teachers spent viewing children's vocabulary and reading scores and the progress monitoring charts, the greater their students' reading gains, but there was variability among the treatment teachers with implications for students' reading gains (Connor, Morrison, Fishman, & Schatschneider, 2010). Hence, access to easily interpretable assessment results appears to be a critical aspect of the ISI intervention.

Teachers in the ISI intervention group generally spent less time in whole-class instruction, spent more time in small-group and individual instruction, and were more likely to provide the A2i recommended amounts than were the control teachers, which suggests stronger classroom organization and management. But again, there was variability. The results of this study suggested that variability in how precisely students receive their recommended amounts of instruction had implications for their reading skills. Even with all the PD support, some treatment teachers were no better than control teachers in providing A2i recommended instruction, using and interpreting assessment information, using small groups, and managing their classroom. Why some teachers take up what is offered in PD, change their practice, and improve student outcomes and other do so to a lesser extent is not well understood but important.

SUMMARY

Given the importance of academic success to children's well-being and ultimate success in life, viewing the process of learning as the complex process it is (Robinson, 1993; Yoshikawa & Hsueh, 2001), impacted by child and environmental factors, not the least of which is instruction itself, will lead to more useful models of instruction and, ultimately, to designing and proactively implementing more effective classroom environments for all children. Results of this study suggest that a dynamic system forecasting intervention model of reading instruction, which conceptualizes instruction across multiple dimensions, may help us better understand how multiple sources of influence interact to affect students' reading achievement. This, in turn, should lead to the design and implementation of more effective reading instruction, which teachers can learn to do. With more data, we can refine dynamic forecasting intervention models and, theoretically, design better instructional regimes and better tools to support teachers' practice and student reading achievement.

REFERENCES

Al Otaiba, S., Connor, C. M., Folsom, J. S., Greulich, L., Meadows, J., & Li, Z. (2010). *Assessment data-informed guidance to individualize kindergarten reading instruction: Findings from a cluster-randomized control field trial*. Manuscript accepted for publication.

Al Otaiba, S., Connor, C. M., Kosanovich, M., Schatschneider, C., Dyrlund, A. K., & Lane, H. (2008). Reading First kindergarten classroom instruction and students' phonological awareness and decoding fluency growth. *Journal of School Psychology*, *48*, 281–314.

Bos, C., Mather, N., Narr, R. F., & Babur, N. (1999). Interactive, collaborative professional development in early literacy instruction: Supporting the balancing act. *Learning Disabilities Research and Practice*, *14*, 227–238.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology: Theoretical models of human development* (6th ed., Vol. 1, pp. 793–828). Hoboken, NJ: Wiley.

Brophy, J. E. (1983). Classroom organization and management. *Elementary School Journal*, *83*, 265–285.

Buell, M. J., & Cassidy, D. J. (2001). The complex and dynamic nature of quality in early care and educational programs: A case for chaos. *Journal of Research in Childhood Education*, *15*, 209–219.

Cameron, C. E., Connor, C. M., Morrison, F. J., & Jewkes, A. M. (2008). Effects of classroom organization on letter-word reading in first grade. *Journal of School Psychology*, *6*, 173–192.

Cicchetti, D. (1990). An historical perspective on the discipline of developmental psychopathology. In A. J. Rolf, D. Masten, Cicchetti, K. Nuechterlein, & S. Weintraub (Eds.), *Risk and protective factors in the development of psychopathology* (pp. 2–28). New York: Cambridge University Press.

Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, *25*, 119–142.

Connor, C. M. (in press). Child by Instruction interactions: Language and literacy connections. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook on early literacy* (3rd ed.). New York: Guilford.

Connor, C. M., Jakobsons, L. J., Crowe, E., & Meadows, J. (2009). Instruction, differentiation, and student engagement in Reading First classrooms. *Elementary School Journal*, *109*, 221–250.

Connor, C. M., Jakobsons, L. J., & Granger, J. (2006, February). *Individual differences in the reading fluency of children and their instructional implications—Instruction and fluency*. Paper presented at the Pacific Coast Research Conference, Coronado, CA.

Connor, C. M., Morrison, F. J., Fishman, B., Ponitz, C. C., Glasney, S., Underwood, P., et al. (2009). The ISI classroom observation system: Examining the literacy instruction provided to individual students. *Educational Researcher*, *38*(2), 85–99.

Connor, C. M., Morrison, F. J., Fishman, B., & Schatschneider, C. (in press). Assessment and instruction connections: The implications of child X instruction Interactions effects on student learning. In J. Sabatini & E. R. Albro (Eds.), *Assessing reading in the 21st century: Aligning and applying advances in the reading and measurement sciences*. Lanham, MD: R& L Education.

Connor, C. M., Morrison, F. J., Fishman, B. J., Schatschneider, C., & Underwood, P. (2007). The early years: Algorithm-guided individualized reading instruction. *Science*, *315*(5811), 464–465.

Connor, C. M., Morrison, F. J., & Katch, E. L. (2004). Beyond the reading wars: The effect of classroom instruction by child interactions on early reading. *Scientific Studies of Reading*, *8*, 305–336.

Connor, C. M., Morrison, F. J., & Petrella, J. N. (2004). Effective reading comprehension instruction: Examining child by instruction interactions. *Journal of Educational Psychology*, *96*, 682–698.

Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's literacy skill growth. *Journal of Educational Psychology*, *98*, 665–689.

Connor, C. M., Morrison, F. J., & Underwood, P. (2006, June). *A second chance in second grade: The independent and cumulative impact of first and second grade reading instruction on students' letter-word reading skill growth*. Paper presented at the Annual Meeting of the Society for the Scientific Study of Reading, Vancouver, BC.

Connor, C. M., Morrison, F. J., & Underwood, P. (2007). A second chance in second grade? The independent and cumulative Impact of first and second grade reading Instruction and students' letter-word reading skill growth. *Scientific Studies of Reading*, *11*, 199–233.

Connor, C. M., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., et al. (2009). Individualizing student instruction precisely: Effects of child by instruction interactions on first graders' literacy development. *Child Development*, *80*(1), 77–100.

Connor, C. M., Ponitz, C. E. C., Phillips, B., Travis, Q. M., Glasney, S., & Morrison, F. J. (in press). First graders' literacy and self-regulation Gains: The effect of individualizing instruction. *Journal of School Psychology*.

Connor, C. M., Son, S.-H., Hindman, A. H., & Morrison, F. J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience: Complex effects on first graders' vocabulary and early reading outcomes. *Journal of School Psychology*, *43*, 343–375.

Cook, T. D., & Payne, M. (2002). Objecting to the objections to using random assignment in educational research. In F. Mosteller & R. Boruch (Eds.), *Evidence matters: Randomized trials in education research* (pp. 150–178). Washington DC: Brookings Institute.

Cronbach, L. J., & Snow, R. E. (1969). *Individual differences in learning ability as a function of instructional variables*. Stanford, CA: Stanford University, School of Education. (ERIC Document Reproduction Service No. ED029001)

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.

Crowe, E., & Connor, C. M. (2007, May). *Examining core reading curriculums: Relations between instructional materials, student level of reading achievement, and oral reading fluency*. Paper presented at the Institute for Educations Sciences, Washington DC.

Crowe, E., Connor, C. M., & Petscher, Y. (2009). Examining the core: Relations among reading curriculums, poverty, and first through third grade reading achievement. *Journal of School Psychology*, *47*, 187–214.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at risk children. *Journal of Educational Psychology*, *90*, 37–55.

Foorman, B. R., Schatschneider, C., Eakin, M. N., Fletcher, J. M., Moats, L. C., & Francis, D. J. (2006). The impact of instructional practices in Grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology*, *31*, 1–29.

Fuchs, L. S., Fuch, D., & Phillips, N. (1994). The relation between teachers' beliefs about the importance of good student work habits, teacher planning, and student achievement. *Elementary School Journal*, *94*, 331–345.

Garg, A. X., Adhikari, N. K. J., McDonald, H., Rosas-Arellano, M. P., Devereaux, P. J., Beyene, J., et al. (2005). Effects of computerized clinical decision support systems on practitioner performance and patient outcomes. *Journal of the American Medical Association*, *293*, 1233–1237.

Graue, E., Clements, M. A., Reynolds, A. J., & Niles, M. D. (2004). More than teacher directed or child initiated: Preschool curriculum type, parent involvement, and children's outcomes in the Child–Parent Centers. *Education Policy Analysis Archives*, *12*(72), 1–40.

Gueron, J. M. (2002). The politics of random assignment. In F. Mosteller & R. Boruch (Eds.), *Evidence matters: Randomized trials in education research* (pp. 15–49). Washington DC: Brookings Institute.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*, 127–160.

Jo, B. (2002). Statistical power in randomized intervention studies with noncompliance. *Psychological Methods*, *7*, 178–193.

Juel, C., & Minden-Cupp, C. (2000). Learning to read words: Linguistic units and instructional strategies. *Reading Research Quarterly*, *35*, 458–492.

Kawamoto, K., Houlihan, C., Balas, E. A., & Lobach, D. F. (2005). Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. *British Journal of Medicine*, *330*, 765.

Kyllonen, P. C., & Lajoie, S. P. (2003). Reassessing aptitude: Introduction to a special issue in honor of Richard E. Snow. *Educational Psychologist*, *38*(2), 79–83

Landry, S. H., Antony, J. L., Swank, P. R., & Monseque-Bailey, P. (2010). Effectiveness of comprehensive professional development for teachers of at-risk preschoolers. *Journal of Educational Psychology*, *101*, 448–465.

Markwardt, F. C. (1998). *Peabody Individual Achievement Test–Revised*. Circle Pines, MN: American Guidance Services.

Masten, A. S., & Coatsworth, J. D. (1998). The development of competence in favorable and unfavorable environments: Lessons from research on successful children. *American Psychologist*, *53*, 205–220.

Mather, N., & Woodcock, R. W. (2001). *Woodcock Johnson III Tests of Achievement: Examiner's manual*. Itasca, IL: Riverside.

Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly*, *40*, 148–182.

Morrison, F. J., Bachman, H. J., & Connor, C. M. (2005). *Improving literacy in America: Guidelines from research*. New Haven, CT: Yale University Press.

Morrison, G. M., & Cosden, M. A. (1997). Risk, resilience, and adjustment of individuals with learning disabilities. *Learning Disability Quarterly*, *20*, 43–60.

NAEP. (2007). *National Assessment of Educational Progress: The nation's report card [Electronic version]*. Retrieved January 2010 from http://nces.ed.gov/nationsreportcard/

Neal, D., & Schanzenbach, D. W. (2007). *Left behind by design: Proficiency counts and test-based accountability*. Retrieved July 2007 from http://www.aei.org/docLib/20070716_NealSchanzenbachPaper.pdf

NICHD-ECCRN. (2002). The relation of global first grade classroom environment to structural classroom features and teacher and student behaviors. *Elementary School Journal*, *102*, 367–387.

NICHD-ECCRN. (2004). Multiple pathways to early academic achievement. *Harvard Educational Review*, *74*(1), 1–29.

Noldus Information Technology. (2001). The observer Video-Pro: Interactive multimedia tutorial (Version 1.0) [Software]. Leesburg, VA: Noldus Information Technology.

NRP. (2000). *National Reading Panel report: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (No. NIH Pub. 00-4769). Washington, DC: U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, National Institute of Child Health and Human Development.

Raudenbush, S. W. (2005). How do we study what happens next? *Annals of the American Academy of Political and Social Science*, *602*(1), 131–144.

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

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, *2*(2), 31–74.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2002). How should reading be taught? *Scientific American*, *286*(3), 84–91.

Reynolds, A. J., & Ou, S.-R. (2004). Alterable predictors of child well-being in the Chicago longitudinal study. *Children and Youth Services Review*, *26*, 1–14.

Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2002). Age 21 cost-benefit analysis of the Title I Chicago child-parent centers. *Educational Evaluation and Policy Analysis*, *24*, 267–303.

Rhome, J. R. (2007). *Technical summary of the National Hurricane Center track and intensity models*. Retrieved October 12, 2008, from http://www.nhc.noaa.gov/modelsummary.shtml

Robinson, R. (1993). Chaos or nonlinear dynamics: Implications for reading research. *Reading Research & Instruction*, *32*(4), 15–23.

Rosenthal, R., & Rosnow, R. L. (1984). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.

Rutter, M., & Maughan, B. (2002). School effectiveness findings 1979–2002. *Journal of School Psycholgoy*, *40*(6).

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97–110). New York: Guilford.

Smith, J. B., Lee, V. E., & Newmann, F. M. (2001). Instruction and achievement in Chicago elementary schools [Electronic version]. *CCSR Publications*. Retrieved December 2009 from http://ccsr.uchicago.edu/publications/p0f01.pdf

Snow, C. E. (2001). *Reading for understanding*. Santa Monica, CA: RAND Education and the Science and Technology Policy Institute.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Taylor, B. M., Pearson, D. P., Clark, K., & Walpole, S. (2000). Effective schools and accomplished teachers: lessons about primary-grade reading instruction in low-income schools. *Elementary School Journal*, *101*, 121–165.

Taylor, B. M., Pearson, P. D., & Rodriquez, M. C. (2003). Reading growth in high-poverty classrooms: The influences of teacher practices that encourage cognitive engagement in literacy learning. *Elementary School Journal*, *104*, 3–28.

Torgesen, J. K. (2000). Individual differences in response to early intervention in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice*, *15*, 55–64.

Vaughn, S., & Coleman, M. (2004). The role of mentoring in promoting use of research-based practices in reading. *Remedial and Special Education*, *25*(1), 25–38.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., et al. (1996). Cognitive profiles of difficult to remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, *88*, 601–638.

Venter, A., Maxwell, S. E., & Bolig, E. (2002). Power in randomized group comparisons: The value of adding a single intermediate time point to a traditional pretest-posttest design. *Psychological Methods*, *7*(2), 194–209.

Wharton-McDonald, R., Pressley, M., & Hampston, J. M. (1998). Literacy instruction in nine first-grade classrooms: Teacher characteristics and student achievement. *Elementary School Journal*, *99*, 101–128.

Xue, Y., & Meisels, S. J. (2004). Early literacy instruction and learning in kindergarten: Evidence from the early childhood longitudinal study—kindergarten class of 1998–1999. *American Educational Research Journal*, *41*(1), 191–229.

Yoshikawa, H., & Hsueh, J. (2001). Child development and public policy: Toward a dynamic systems perspective. *Child Development*, *72*, 1887–1903.

APPENDIX

Excerpt from the ISI Coding Manual for Word Identification/Encoding (TCM code-focused activity)

7.1.4 Word Identification/Encoding (Behavior) we

Word ID/Decoding should be coded when the intent of the activity is to allow students to practice their spelling skills. Whereas Word Identification/Decoding involves going from print to pronunciation or letters to sounds, Word Identification/Encoding involves going from pronunciation to print or sounds to letters (or oral spellings using letter names). Also included in this category is the atypical example of students copying single words from the blackboard. Furthermore, these activities should involve spelling single words (e.g., word lists, a mini-lesson on a single word on which a student is struggling during writing connected text) and methods pertaining to the spelling of single words. Lessons involving the writing of connected text should be coded under Writing. If students are spelling a word while looking at print, see Decoding>Letter Naming/Recognition.

7.1.4.1 Missing (Modifier)

This is a default modifier automatically created by Noldus. It will not be used for our coding purposes.

7.1.4.2 Copying (Modifier) cop

Word ID/Encoding>Copying should be coded for lessons in which single words are copied with the intent of learning their spellings. *If the activity is also intended to improve handwriting, it should still be coded as Word ID/Encoding>Copying but with a note in the comment field indicating that the activity also involved handwriting.* Activities involving the copying of connected text should be coded under Writing>Copying, and activities involving only handwriting practice (without the intent of learning words' spellings) should be coded as Writing>Handwriting Instruction/Practice.

7.1.4.3 Spelling Written (Modifier) sp

Word ID/Encoding>Spelling Written should be coded for those activities which involve hearing a word and writing its spelling (e.g., written spelling test/dictation, spelling words to be written on the blackboard, using magnet letters to spell words). Word ID/Encoding>Spelling Written activities are usually Teacher/Child Managed, especially during dictation tests.

7.1.4.4 Letter Naming and Spelling (Modifier) ln

Word ID/Encoding>Letter Naming and Spelling should be coded for those activities which involve hearing a word and giving its spelling in an oral form (e.g., after a dictation test the teacher asks the students to spell words from their tests.).

7.1.4.5 Sounding Out (Modifier) sou

Word ID/Encoding>Sounding Out should be coded when students are involved in a lesson focused on the explicit use of grapheme-phoneme correspondences for spelling single words (e.g., hear the word /hat/, break it into /h/ /a/ /t/, then determine which letters are associated with these sounds). The teacher explains or models the approach, or the children model the approach by producing each sound before blending them together and forming the word. Whereas Phoneme Awareness>Elision activities start with words which must be split apart into sounds, Word ID/Encoding>Sounding Out requires the children to then associate these sounds with letters.

7.1.4.6 Syllables (Modifier) syl

Word ID/Encoding>Syllables should be coded when students are involved in a lesson focused on the use of syllables for spelling. The teacher explicitly explains or models the approach, or the children model the approach by spelling each syllable individually before combining them together and forming the word (e.g., spelling "mit" and then spelling "ten" to arrive at the spelling of *mitten*). Note that the students are spelling each syllable individually first; they are not attempting to spell the entire word all at once. The lesson is not focused on the use of syllable types to determine the vowel sounds within words, which is coded under Word ID/Encoding>Syllable Types.

7.1.4.7 Syllable Types (Modifier) slt

Word ID/Encoding>Syllable Types should be coded for activities intended to familiarize students with the use of the six syllable types (open, closed, consonant-L-E, magic E, R-controlled vowel, vowel team) to determine spellings for various vowel pronunciations within words. The teacher explicitly explains or models the approach, or the children model the approach by determining the vowel sound and the matching syllable type(s), and then using the form of the syllable type to aid in spelling the word (e.g., hear the short /a/ sound, which means a consonant must follow the "a" in the word). Note that the difference between this code and other Word ID/Encoding codes is the explicit attention to the syllable type and the vowel sound within the word. See Appendix A Syllable Types.

7.1.4.8 Analogy/Word Families (Modifier) ana

Word ID/Encoding>Analogy/Word Families should be coded for activities in which students use their knowledge of analogy or word families for spelling (e.g., the word /mat/ sounds like /hat/ but with the /m/ at the beginning, and I know that /at/ sound is spelled "at" so /mat/ must be spelled "mat"). The teacher explicitly explains or models the approach, or the children model the approach by isolating the rime of the word, identifying another word that follows that pattern (i.e., is in the same family), and then spelling the new word by spelling the known rime and new onset.

7.1.4.9 Word Recognition and Reading (Modifier) wo

Word ID/Encoding>Word Recognition and Reading should be coded for activities in which individual students and the students are spelling words aloud *without any of the particular strategies mentioned above* (e.g., Teacher says I can spell "cat" and says the letters "c" "a" "t").

7.1.4.10 Context Cues (Modifier) cnc

Word ID/Encoding>Context Cues should be coded when the teacher explicitly explains, models, or prompts children to use context clues (possibly along with the initial sound of the word) to aid in determining the spelling of a word (e.g., using pictures, using surrounding words, predicting).

7.1.4.11 Morphological/Structural Analysis (Modifier) str

Word ID/Encoding>Morphology should be coded for those activities which focus on spelling words through knowledge of root words, suffixes, and/or prefixes. The teacher explicitly explains or models the approach, or the children model the approach by isolating the morphemes within words, the spellings of which they are familiar, and spelling each morpheme separately before combining them to form the spelling of the whole word. This approach, unlike Word ID/Encoding>Syllables, requires that words be split into meaningful pieces. For example students are trying to spell "preview," first they spell "pre" and then they spell "view" and put both parts together to form the whole word. Another example of Morphological/Structure Analysis is if the student is asked to find all the words that go with the root word "natural."

7.1.4.12 Sight Words (Modifier) si

Word ID/Encoding>Frequent Irregular Sight Words should be coded for those activities when the student is practicing learning their sight words (e.g., *though*), this is frequently done with flash cards. This exercise is meant to practice accuracy not fluency, if fluency is the objective the activity is better coded under Fluency>Words.

7.1.4.13 Rules (Modifier) rul

Word ID/Encoding>Rules should be coded for lessons focused on an explanation or use of a particular phonics or spelling rule (other than the typical grapheme-phoneme correspondences; e.g., using "ck" after a short vowel sound, doubling letters, "i" before "e" except after "c"). Rules regarding the formation of plurals (or returning to a singular form from a plural form) should be coded under Morpheme Awareness>Blending/Plural or Morpheme Awareness>Elision/Plural, respectively.

7.1.4.14CF-TBD (Modifier) cft

The Word ID/Encoding>CF-TBD code should be used only when (a) none of the other Word ID/Encoding modifier codes are appropriate for a given activity and (b) the activity fits the Word ID/Encoding description. *A brief description of the activity should be noted in the comment field*. Note that, by definition, these activities should be code focused.