Individualizing Student Instruction Precisely: Effects of Child by Instruction Interactions on First Graders' Literacy Development

Carol McDonald Connor, **Shayne B. Piasta**, **Barry Fishman**, **Stephanie Glasney**, **Christopher Schatschneider**, **Elizabeth Crowe**, **Phyllis Underwood**, and **Frederick J. Morrison**

Abstract

Recent findings demonstrate that the most effective reading instruction may vary with children's language and literacy skills. These child X instruction interactions imply that individualizing instruction would be a potent strategy for improving students' literacy. A cluster-randomized control field trial, conducted in 10 high-moderate poverty schools, examined effects of individualizing literacy instruction. The instruction each first grader received (n=461 in 47 classrooms, mean age = 6.7 years), fall, winter and spring, was recorded. Comparing intervention-recommended amounts of instruction with observed amounts revealed that intervention teachers individualized instruction more precisely than did comparison teachers. Importantly, the more precisely children received recommended amounts of instruction, the stronger was their literacy skill growth. Results provide strong evidence of child X instruction interaction effects on literacy outcomes.

Too many children in America fail to achieve proficient reading skills and the rate is particularly troubling, close to 60%, for children living in poverty and who belong to underrepresented minorities (NAEP, 2005). Whereas multiple factors can affect children's literacy development including home, parenting, parent educational levels, preschool, community resources, as well as formal schooling (Connor, Son, Hindman, & Morrison, 2005; NICHD-ECCRN, 2004), classroom instruction is one of the most important sources of influence. Further, one reason children fail to achieve proficient reading skills is because they do not receive appropriate amounts of particular types of literacy instruction during the primary grades. Early literacy instruction that is balanced between phonics and more meaningful reading experiences has been shown to be more effective than instruction that focuses on one to the exclusion of the other (Mathes, Denton, Fletcher, Anthony, Francis, & Schatschneider, 2005; Xue & Meisels, 2004).

Moreover, the impact of any particular instructional strategy may depend on the language and literacy skills children bring to the classroom. In other words, there are child characteristic-by-instruction interactions (child X instruction interactions; Connor, Morrison, & Katch, 2004a; Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Juel & Minden-Cupp, 2000). By implication, these findings point to the potential importance of individualizing (or personalizing or differentiating) instruction based on the child's entering skill levels. To date, although child X instruction interactions have emerged across grades and outcomes (Al Otaiba, Connor, Kosanovich, Schatschneider, Dyrlund, & Lane, 2008; Connor, Jakobsons, Crowe, & Meadows, in press; Foorman, Schatschneider, Eakin, Fletcher, Moats, & Francis, 2006), these studies have been predominantly descriptive and correlational.

Contact: Carol McDonald Connor, cconnor@fsu.edu.

To evaluate the causal implications of child X instruction interactions, we conducted a randomized control field trial. Results showed a significant treatment effect for an intervention that individualized students' instruction based on students' vocabulary and reading skills, utilizing homogeneous student groupings and computer software support (Connor, Morrison, Fishman, Schatschneider, & Underwood, 2007a). The multi-faceted nature of the intervention, however, left open questions regarding the specific role of individualizing instruction on students' growth. Thus, the present study aims to disentangle the effect of individualizing instruction from other sources of influence (e.g., teacher qualifications) by examining the precision with which teachers provided recommended amounts and types of instruction and the effect of this instruction on first graders' literacy skill growth.

Aptitude-by-Treatment Interactions

The idea that the effect of instruction may depend on students' abilities is not new. During the late 1950s Cronbach introduced the theory of aptitude-by-treatment interactions (ATI) (Cronbach, 1957). Chall (1967) and Bond and Dykstra (1967) examined whether students' abilities were related to how they responded to various types of reading instruction. Researchers of the era generally identified five aptitude factors for which there was some indication of interaction effects (Bond & Dykstra, 1967; Lo, 1973). These included intelligence, perceptual speed, spatial perception, listening skills, and learning rate. In Barr's (1984) Reading Handbook chapter on primary reading, ATIs were carefully described and supporting evidence cited. While promising, research examining the construct of ATIs produced a paucity of sound experimental evidence (Bracht, 1970; Cronbach & Webb, 1975; R. E. Snow, 1991). This was despite decades of dedicated research on individualization. In an extensive review and meta-analysis of the education literature, Cronbach and Snow (1977) argued that there was only suggestive evidence of ATIs. Moreover, they questioned the studies' methodologies and, when they could not replicate the findings using more rigorous analytic strategies, noted:

> …well-substantiated findings regarding ATI are scarce. Few investigations have been replicated. Many reports (of both positive and negative results) must be discounted because of poor procedure (p. 6).

As a field, however, education has never fully abandoned the promise of child X instruction interactions. Predictable differential responses to reading instruction strategies have the potential, theoretically, of explaining why some children respond well to schooling while others fail and, most importantly, how we might support all children's literacy development. Indeed, this notion is at the heart of "responsiveness to intervention" protocols (D. Fuchs, Mock, Morgan, & Young, 2003 2003). Now, with a better understanding of the specific cognitive and social mechanisms underlying the skill of reading (Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001; C. E. Snow, Burns, & Griffin, 1998), the advent of more sophisticated analytic strategies, such as hierarchical linear modeling (Foorman et al., 2006; Raudenbush & Bryk, 2002; Raudenbush, Hong, & Rowan, 2002), and multidimensional conceptualizations of instruction (Connor, Morrison, & Slominski, 2006a; Morrison, Bachman, & Connor, 2005 2005), correlational evidence on the nature of child X instruction interactions is converging.

Child-by-instruction Interactions and Multiple Dimensions of Instruction

In an earlier study (Connor et al., 2004a), which provided the underlying models for the current study, the specific type of instruction first graders received interacted with their vocabulary and word-reading skills. Instruction was conceptualized across multiple dimensions, hypothesizing that within a school day, teachers used many types of literacy instruction even if their overall curriculum was whole language or more focused on basic

reading skills (Marilyn Jager Adams, 1990; Chall, 1967). Indeed, this multiple dimensional approach to investigating classroom literacy instruction represented important methodological progress towards identifying child X instruction interactions.

Three dimensions of instruction, which operated simultaneously to define a given literacy activity (see Table 1) were included: (1) *code versus meaning-focused*, where *code-focused instruction* focused explicitly on the task of decoding, such as relating phonemes to letters, learning to blend them to decode words, and included alphabet activities, phonological awareness, phonics, and letter and word fluency. *Meaning-focus activities* supported students' efforts to actively extract and construct meaning from text (C. E. Snow, 2001) and included reading aloud, reading independently, writing, language, vocabulary, and comprehension strategies.

(2) The dimension of *teacher/child- versus child-managed instruction* captures who is responsible for focusing the child's attention on the learning activity at hand, the teacher and students jointly, or the child. Thus an activity where the teacher is actively interacting with children is considered teacher/child managed. Activities where children are working with peers or independently are considered child managed.

(3) The dimension of *change-over-time* or *slope* accepts that particular instructional strategies may be more important for students at certain times of the year. For example, Juel and Minden-Cupp (2000) found that when one teacher provided greater amounts of teacher/ child-managed code-focused instruction in the fall, and decreased the amount over the school year, children with weaker reading skills made greater progress in her classroom than did children with similar skills in a classroom where substantial amounts of code-focused instruction were provided all year long. When instruction is observed at least three times during the school year, the slope of change in minutes per month can be computed.

These dimensions operate simultaneously (see Table 1). For example, a teacher discussing a book with students would be a teacher/child-managed meaning-focused activity. If the teacher then asks the children to read the rest of the story in the library corner or with a buddy, the activity would become a child-managed meaning-focused activity. If the teacher is conducting phonological awareness instruction, this would be considered teacher/child-managed code-focused instruction (in contrast to the previous teacher/child-managed activity, which was meaning focused). If the children are then given a phonics worksheet to complete at their desks, that activity is child-managed code-focused. The month of the school year contributes to the decision about how much time is spent in specific types of activities for a specific child at a particular point in time.

The advantage of this multi-dimensional conceptualization is that the dynamic and complex interactions among teachers and students surrounding reading instruction can be captured in more detail than can be accomplished with global curriculum-level conceptualizations. Additional dimensions can be used to define literacy activities more precisely. For example, the current study also considers whether the activity is presented to the entire class, to a smaller group of children, or individually. Utilizing more dimensions provides greater precision in defining literacy activities. At the same time, however, adding dimensions creates more literacy instruction variables to include in statistical models.

Connor, Morrison and Katch (2004a) found significant child X instruction interactions that predicted first graders' word recognition skill growth (i.e., residualized change) noting that total amounts of teacher/child-managed meaning-focused and child-managed code-focused instruction did not significantly predict students' word reading skills. It is these specific child X instruction interactions that we sought to investigate in this study. First, children with weaker fall word recognition skills made greater progress when they were in

classrooms where teachers provided substantial amounts of teacher/child-managed code-focused instruction than did children with similar skills in classrooms where very little teacher/child-managed code-focused instruction was observed (in the 2004 study, different terminology was used). In contrast, children with strong fall word recognition skills made less progress in classrooms where substantial amounts of teacher/child-managed code-focused instruction were observed while students with similar skills made greater progress in classrooms where less teacher/child-managed code-focused instruction was observed.

Additionally, children who started first grade with weaker vocabulary skills demonstrated greater word recognition skill growth in classrooms where they received small amounts of child-managed meaning-focused instruction in the fall with steady increases in amount until spring (i.e., change over time or slope). This is in contrast to children with similar skills who made less progress when they were in classrooms that provided greater amounts of child-managed meaning-focused instruction consistently all year long. Yet, in the latter classrooms, children with stronger fall vocabulary scores showed greater word recognition skill growth than they did in classrooms with small amounts of child-managed meaning-focused instruction. Teacher/child-managed meaning-focused and child-managed code-focused instruction was not systematically associated with students' word reading skill growth. Because these results were correlational, a cluster-randomized control field trial was conducted (Connor et al., 2007a).

Assessing the Efficacy of Individualizing Student Instruction

For the cluster-randomized control field trial, schools were matched by percentage of students qualifying for Free or Reduced Priced Lunch, Reading First status, and third grade reading scores on the state mandated achievement test (see Table 2). Then one of each matched school pair was randomly assigned to the Individualizing Student Instruction (ISI) intervention condition with the remaining school assigned to a wait-list control comparison condition. Teachers in the comparison group received the intervention training the following school year.

The intervention group received training and professional development on how to individualize literacy instruction in the classroom using the recommendations and planning strategies provided by Assessment-to-Instruction (A2i) web-based software. A2i software incorporated algorithms that recommended amounts and types of instruction based on the interactions observed in earlier studies (Connor et al., in press; Connor et al., 2004a; Connor, Morrison, & Underwood, 2007b 2007; Morrison et al., 2005) and relied on the multidimensional instructional framework. In this manuscript, we use a shorthand to refer to the instructional approaches: teacher/child-managed, code-focused instruction is TCM-CF and in formulae simply TCMCF; child-managed, code-focused instruction is CM-CF or CMCF, and so on.

The hierarchical linear models were able to predict children's word reading skill growth when the month of the school year (from the slope in the models), amounts of instruction, and initial word reading and vocabulary skills were known (Morrison et al., 2005, 2005). It is these equations that provided the foundation for the A2i algorithms. The mixed-model equation, using grade equivalent (GE) as the metric rather than raw scores, is as follows, where Y_{ij} is the predicted spring reading outcome for child i in classroom j:

$Y_{ij} = .05 + 1.1*$fall reading $+ .01*$fall Vocabulary $+ .05*$TCMCF-A $- .82*$TCMCF-S $-.15*$CMMF-A $+ .7*$CMMF-S $-.04*$fall reading$*$TCMCF $+ .01*$fall Vocabulary$*$CMMF-A $- .01*$fall Vocabulary$*$CMMF-S $+ u_{0j} + r_{ij}$

Essentially, the A2i algorithms use a pre-set Y_{ij}, which is the end of year target outcome, the child's assessed reading and vocabulary scores, and then solve for recommended amounts of TCMCF and CMMF. TCMCF-A and CMMF-A are the amounts of TCMCF and CMMF in the fall. TCMCF-S and CMMF-S represent the slope or change over time, which is change in amount of TCMCF or CMMF per month. TCMMF and CMCF were set based on mean amounts observed across studies. The algorithms were then refined using evidence from other studies (Al Otaiba et al., 2008, in press; Connor et al., in press; Connor, Morrison, & Petrella, 2004b 2004; Connor et al., 2006a; Connor et al., 2007b) and tested in this study.

The A2i algorithms used in this investigation solve for TCM-CF and CM-MF recommended amounts in a series of steps. Additionally, because the classroom teachers in the Connor et al., 2004 study provided relatively small amounts of TCM-CF, a minimum of 13 minutes of TCM-CF was added to the equation based on research with other samples (Connor et al., in press; Connor, Morrison, & Underwood, 2006b; Foorman et al., 2006).

The two steps to solve for amount of TMCF are:

1. TCMCF-A = ((Target Outcome or Yij − (.2*reading grade equivalent))/(.05 + (.05*reading grade equivalent))) + 13

2. TCMCF-Recommended = TCMCF-A − (.82*Month)

Thus, to compute recommended amounts of teacher/child-managed code-focused instruction, first the target outcome (the Y_{ij} of the original model) is set at either grade level or a school year's gain in skill using grade equivalent (GE) as the metric. For example, if the child is reading at a GE level of 1.3 at the beginning of first grade, the target outcome is 2.2 (1.3 + .9). If the child begins first grade with skills below grade level expectations (e.g., .5), the target is set at grade level (i.e., 1.9). Then the assessed reading score, in GEs, is entered and the equations are solved. Month is centered at August, thus September is equal to 1, October is equal to 2 and so on. The value −.82 represents the slope. Therefore, the recommended amount of TCM-CF decreases each month by .82 minutes. The relation between reading score, month, and recommended amount of TCM-CF is non-linear. This means that the recommended amount of TCM-CF goes up exponentially as children fall farther behind grade level expectations.

The algorithms for CM-MF are provided in Appendix A. The relation between vocabulary AE and recommended CM-MF over time resembles a fan. The open end of the fan shows the amounts of CM-MF for children with lower vocabulary scores (age equivalent [AE] = 3.0 years). These amounts begin with small amounts per day and increase monthly (about 8 minutes/month) until they reach about 40 minutes in the spring. The hinge of the fan represents the children with very high vocabulary scores for whom the algorithms recommend about 40 minutes of CM-MF per day all year long. Children whose vocabulary scores fall within more typical ranges start out the school year with relatively less CM-MF recommended and the amounts increase until they too have a recommendation of about 40 minutes/day of CM-MF in the spring.

Importantly, the recommended amounts fall on a continuum varying incrementally as students' reading and vocabulary scores vary. For both TCM-CF and CM-MF, the patterns of recommendations reflect the child X instruction interactions described in Connor, Morrison, and Katch (2004) and it is these interactions that this study was designed to test.

A2i grouping algorithms provided recommended homogeneous groupings of students based on their word reading skills. The rationale for using small groups was based on the effective schools literature (Taylor, Pearson, Clark, & Walpole, 2000 2000; Wharton-McDonald, Pressley, & Hampston, 1998 1998) and on preschool research (Connor et al., 2006a), which

suggested that (1) children in schools that used small homogeneous skill-based groups had generally stronger reading skills compared to children in schools that did not, (2) there were main effects for preschool instruction provided in small groups, and (3) that the effect of small group instruction on students' reading skill gains was greater than similar instruction provided to the whole class.

A2i also incorporated planning software and indexed the schools' core reading curriculum to the multidimensional types of instruction. By indexing the core reading curriculum, we aimed to support teachers' ability to differentiate the content of the instruction as well as the amount. The A2i software recommended only minutes per day for each type of instruction. We relied on teachers' knowledge and decision-making, curriculum scope and sequence, and our professional development to help teachers differentiate the specific content of the instruction. For example, although teaching letter-sound associations is a code-focused activity, it would not be an appropriate activity for a child who already knew how to read fluently. However, it would be an appropriate activity for a child who was just learning the alphabetic principle (Marilyn J. Adams, 2001).

Additionally, teachers received intensive training on how to use A2i and how to individualize instruction in their classrooms based on the A2i recommendations. The Individualizing Student Instruction (ISI) intervention professional development used a mentoring or reading coach model that is employed widely, particularly in Reading First schools (Gersten, Chard, & Baker, 2000; Showers, Joyce, & Bennett, 1987; US DOE, 2004; Vaughn & Coleman, 2004). In this model, teachers formed a grade-level community of learners, guided by a highly trained reading specialist (the research partner), who also provided classroom-based support to teachers. Research suggests that reflection as a member of a learning community or study group may have particular power (Lave & Wenger, 1991; Ramanathan, 2002). In addition, a genuine partnership between teachers and mentors (e.g., reading coaches, researchers), including respect for teachers' expertise and responsiveness to their thoughts and concerns, was used to foster receptiveness to the study goals as well as to enhance the quality of the professional development itself (DuFour & Eaker, 1998; Freeman & King, 2003; Helterbran & Fennimore, 2004).

Individualizing Student Instruction Intervention—The ISI intervention was provided by the classroom teachers to all of the children in their classroom from September through the end of the school year in May. A description of the fully implemented intervention is provided in Appendix B. The two-hour length of the literacy block was mandated by the district for both intervention and comparison classrooms so the manipulation was the ISI intervention rather than the dedicated literacy block. Children were assessed in the fall, winter and spring on a battery of language and literacy skills.

Results of the analyses using HLM (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2004, 2004) revealed a significant effect of treatment (Connor et al., 2007a). Children in the ISI intervention classrooms demonstrated greater growth in reading comprehension (i.e., residualized changed) than did children in the control classrooms, on average. This difference represented about a two month difference in grade equivalent scores (effect size, $d = .25$). Plus, the more teachers used A2i software, the greater was their students' reading comprehension skill growth and this effect was greater for children who began the school year with weaker vocabulary skills.

Still, these results beg the question, what was the specific effect of individualizing instruction on students' literacy growth? Other sources of influence may be responsible for the observed treatment effect. For example, might the treatment effect have been the result of the professional development teachers received rather than the different amounts and

types of instruction children received? More knowledgeable teachers may have been more likely to use A2i or were more effective generally (Moats, 1994), which may have been unrelated to individualizing instruction. There is some evidence that teachers' qualifications contribute to their effectiveness in the classroom, as measured by a number of practice and student outcomes (Darling-Hammond, 2000; Goldhaber & Brewer, 2000; NICHD-ECCRN, 2002; Pianta, Howes, Burchinal, Bryant, Clifford, Early, & Barbarin, 2005), although the effects are frequently small. Teachers' knowledge about reading and language concepts has been shown to relate to student outcomes as well (Moats, 1994; Moats & Foorman, 2003).

At the same time, control teachers may have been providing algorithm recommended amounts (drift of the intervention), which would have tended to weaken documented treatment effects. Additionally, Reading First funds were available to five of the schools and the training and extra services provided may have contributed to the treatment effect. Although we matched schools on Reading First status, we included that variable in our models as well.

In the present study, we begin to disentangle the relative contribution of the instruction children received and other sources of influence on student achievement. By observing the amounts and types of literacy instruction provided in both the intervention and the comparison classrooms during the fall, winter and spring and then comparing total amounts provided in the classroom with the amounts recommended by the A2i algorithms, we can examine the specific effects of instruction. Finding that amounts and types of instruction that more precisely match A2i recommended amounts predict students' literacy skill growth would provide strong evidence that child X instruction interactions contribute to individual differences in students' literacy achievement.

Research Questions

In summary, there is accumulating evidence that child X instruction interactions may be causally related to the observed variability in student achievement within and across classrooms. Most of these studies, however, are descriptive and correlational, examine instruction at the classroom or curriculum level, or do not examine the instruction of both intervention and comparison group teachers. The purpose of this study is to investigate the implementation of the ISI intervention in the classroom and the precision with which teachers taught the A2i recommended amounts for each type of instruction for individual children in their classroom. We hypothesize that it was the instruction that children received, and not other sources of influence, that should predict their reading outcomes if indeed child X instruction interactions are predicting student outcomes. Thus we asked the following research questions:

1. Generally, what kinds of literacy instruction were provided in these first grade classrooms, did amounts and types of instruction differ for children who shared the same classroom, and were there differences in instructional patterns between the intervention and comparison classrooms? We hypothesized that, compared to teachers in the comparison group, teachers in the intervention group would tend to provide more instruction in small groups or individually rather than with the whole class. We also predicted that children who shared the same classroom would receive differing amounts and types of instruction.

2. Did instruction that more precisely matched the A2i recommended amounts of teacher/child-managed code-focused and child-managed meaning-focused instruction predict stronger student outcomes? First, we hypothesized that teachers in the intervention group would provide the A2i recommended amounts more precisely than would the comparison group teachers. Next, because predictable child X instruction interactions affect the efficacy of particular instruction strategies

for specific children, we predicted that total amounts of instruction would not predict student outcomes. However, we predicted that the precision with which teachers provided the A2i recommended amounts for each student would predict growth in student literacy skills, even when controlling for teacher, school and student characteristics that might be related to the efficacy of the ISI intervention.

Methods

Participants

Four-hundred sixty-four students in 47 classrooms in 10 schools participated in this study. Schools, located in one school district in Florida, were highly diverse ethnically, and were located in neighborhoods that varied in SES. School SES was based on the percentage of children qualifying for the free or reduced priced lunch program. Half of the schools were participating in the Florida Reading First program, which is a federally funded program run by school districts and designed to improve instruction at historically low income and underperforming schools. Descriptive information for each school is provided in Table 2. The district administration nominated 10 schools, which were matched based on percentage of children qualifying for free or reduced price lunch, the mandated Florida Comprehensive Achievement Tests (FCAT) 3rd grade reading scores, and Reading First status. Then, one member of each school pair was randomly assigned to the ISI intervention group. With this design, teachers at the same school stayed together to improve professional development efforts, as well as to prevent drift of the intervention to the comparison group.

All of the students assigned to participating teachers were invited to join the study, including children for whom English was a second language and who qualified for special services. Parental consent was obtained for 76% of the students (n = 616). Sixty percent of the children were girls. Fifty-nine percent of children qualified for free or reduced price lunch. Fifty-four percent were African American, 37% were European American, and the remaining children belonged to other ethnic groups.

From this group, students were randomly selected as target children for the classroom observations. For each classroom, students were rank ordered by their fall *Woodcock-Johnson-III Tests of Achievement* (Mather & Woodcock, 2001) letter-word identification W score (description below) and divided into three groups of equal size. In this way, each classroom had a group of students who had weak, average, and strong reading skills according to the norms of their classroom. Four students from each group were randomly selected. If there were not 12 participating students in a particular classroom, then all of the children were selected. This method was utilized anticipating poorer attendance and more attrition among the students with weaker reading scores, which was the case. In all, 464 students were selected, observed and their instruction coded in their classroom at least once. Thirty-three children were observed only once, 87 were observed twice, and 344 were observed in the fall, winter, and spring. Using a Bonferroni Correction for multiple analyses (alpha was divided by the number of models [6] so $\alpha = .05/6 = .008$), HLM revealed no significant differences in fall or spring scores when sub-samples and full samples were compared. HLM analyses revealed no significant differences in fall letter-word reading and passage comprehension W scores between groups. Children in the comparison classrooms had significantly higher fall vocabulary scores compared to children in the intervention classrooms [coefficient = −5.26, t (45) = −3.887, p < .001].

Forty-nine teachers began the study, 23 in the treatment group and 27 in the control. In one intervention school, children changed classrooms during the language arts block and, thus, one of the teachers taught second graders during the winter and spring observations. Therefore, his classroom was only observed in the fall shortly before he discontinued his

participation in the study. We did code that classroom and these data are included in the analyses. Also, two control teachers team-taught one classroom. Although both teachers were videotaped, one teacher was selected at random and we coded her instruction. Those data were used in the analyses. Another control teacher left the study because of health reasons before we could observe her classroom. A special education control classroom was also observed. However, children were also observed in their regular classroom, and so the classroom data were included in the analyses. This left a total of 47 teachers and classrooms observed at least once. Eighty-seven percent of the teachers were female, 64% were European American and the remaining 36% were African American.

Generally, teachers' qualifications and characteristics were similar across the two groups. Means for intervention and comparison groups with confidence intervals are available in Table 3. Although control teachers started out the year with higher scores on our test of knowledge about language and literacy concepts, teachers in the intervention group demonstrated greater gains on the assessment from fall to spring. Teachers in the intervention group were more likely to teach at a school with a higher percentage of children eligible for free or reduced price lunch compared to teachers in the comparison group.

Procedures

Utilizing a wait-list control design, the teachers in the intervention classrooms received A2i and professional development in the 2005–2006 year and the teachers in the comparison classrooms received training and A2i technology during the 2006–2007 school year. Only the 2005–2006 results are included in this study. All schools in the district were required to provide a 120 minute block of uninterrupted language arts instruction, of which 45 minutes were supposed to include small group instruction. For this reason many of the control teachers also used small groups during the dedicated language arts block. Thus the principal experimental manipulation was providing the A2i recommended amounts of instruction for each child.

In all, six research partners provided professional development to the teachers in the intervention group. Research partners were assigned to a specific school (one school had two research partners) and no systematic differences among schools or research partners were observed for teacher ISI implementation or student outcomes.

Professional development focused on (1) teaching teachers how to plan instruction using A2i and (2) how to implement the recommended amounts by individualizing instruction. This included training on: (a) classroom management and organization, (b) differentiating content and delivery of instruction based on students' assessed skills, and (c) using high quality research-based literacy activities, especially during child-managed activities. Research partners met with individual teachers at the schools every other week during their literacy block to provide classroom-based support and met with grade-level teams at their respective schools monthly following the established protocol. Teachers in the intervention group also participated in a half-day workshop (3 hours) in spring 2005 and a full day workshop (6 hours) in fall 2005. The amount of time spent with teachers was carefully monitored to be as consistent as possible among teachers. Each teacher received approximately 35 hours of professional development at their school in addition to the 9 hours of workshop.

Assessments

Student Assessments—Students' language and literacy skills were assessed in fall 2005, and again in winter and spring 2006 using a battery of language and literacy assessments, including tests from the *Woodcock Johnson Tests of Achievement-III* (Mather

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

We assessed students' word reading skills using the *Letter-Word Identification test*, which asks children to recognize and name increasingly unfamiliar letters and words out of context. We assessed children's reading comprehension skills using the *Passage Comprehension test,* which utilizes a cloze procedure where children are asked to read a sentence or brief passage and supply the missing word. To assess expressive vocabulary, we used the *Picture Vocabulary test*, which asks children to name pictures of increasingly unfamiliar objects.

The letter-word identification grade equivalent (GE) and the picture vocabulary age equivalent (AE) scores were entered into the A2i software. These scores were used to compute recommended amounts of each instruction type for children in the intervention and comparison classrooms. Mid-year, children were re-administered the alternate form of the Letter-word identification and Picture Vocabulary tests. Recommended amounts in A2i were recomputed using the new scores. The intervention group teachers first gained access to algorithm recommendations and assessment information provided by A2i software in September 2005 and used the software continuously through May 2006. The comparison group teachers were provided written reports of the assessments results for their students in September 2005 and February 2006.

Teacher Assessments—Teachers completed a survey in the fall designed to obtain information on their experience and education. Additionally, teachers' knowledge of English language and literacy was assessed using the Teacher Knowledge Assessment: Language and Print (TKA:LP) in the fall and (Piasta, Connor, Fishman, & Morrison, 2008). The TKA:LP was designed for this study to assess teachers' understanding of English phonology, orthography, and morphology, as well as important concepts of literacy acquisition and instruction with an emphasis on metalinguistic and phonics concepts. For example, teachers were asked to select the word containing a short vowel sound (choices: *treat, start, slip, paw, father*), count the number of syllables in the word *walked* and the number of morphemes in the word *unbelievable*, and identify the following phonological task: "I am going to say a word and then I want you to break the word apart. Tell me each of the sounds in the word *dog*" (choices: *blending, rhyming, segmentation, deletion*). Questions on the TKA:LP were adapted from earlier work (Bos, Mather, Dickson, Podhajski, & Chard, 2001; Mather, Bos, & Babur, 2001; Moats, 1994; Moats & Foorman, 2003). The TKA:LP consisted of 34 multiple choice items and 11 short answer items and had a reliability of alpha = .87. Out of a possible 40 points, fall scores on the TKA:LP ranged from 9 to 36 (*M* = 23.45, *SD* = 7.27).

Classroom observation—All participating teachers were videotaped at their convenience three times throughout the year during the fall, winter, and spring. Observations were rescheduled if a teacher was absent or requested that we observe them on another day. Across all classrooms, observations were conducted during the entire literacy block. We videotaped all day during the winter; however, full day analysis was beyond the scope of this study. In two schools (one intervention & one control), children changed classrooms during the literacy block to achieve more homogeneous skill-based groupings. We report observation data for the child during the literacy block when they were with their reading teachers and not for the time they spent with their homeroom teachers. This decision was based on our observation that teachers were more likely to use small groups during the time students were in reading skills-based classes.

Research assistants used two digital video cameras with wide-angle lenses for each observation so that we would be better able to capture as much of the classroom instruction as possible. Child descriptions were recorded for all participating students that were present, however, only the target students' activities and instruction were coded from the video. Because video cameras were set up to capture the widest range of activities in each classroom, detailed field notes were also written throughout the observation period, recording specific information that might not be interpretable from the videotape alone. Notes included times when children entered and left the classroom, explanations of events or activities that took place outside of the camera's view, descriptions of worksheets and other activities (e.g. from field notes, "students are completing a vocabulary worksheet where they must match words and pictures."). We did not observe or code activities that took place outside of the classroom nor did we follow students who left their classroom during the observation period.

Classroom video observations were coded in the laboratory using the Noldus Observer Pro version 5.0 software package (Noldus Information Technology, 2001). The coding system was expanded and adapted from the Pathways to Literacy Coding scheme (Connor et al., 2006a) and was designed to capture the amount of time, in minutes: seconds, that target students spent in various classroom activities. Videos were coded by trained research assistants, some of whom also did the video observations. To the extent possible, coders were blind to condition. All activities lasting at least 15 seconds were coded, including both academic (e.g., literacy instruction) and non-academic (e.g., attendance) activities. Good inter-rater reliability (Kappa = .80) was achieved on training videos among all coders before formal coding commenced. Additionally, approximately 5% of the completed observations were randomly selected and 30 minutes were re-coded by another research assistant. Reliability among coders was moderate to almost perfect with a Cohen's kappa mean of .76 and a range from .50 to .92 (Landis & Koch, 1977).

In accordance with the study's research questions and hypotheses, classroom activities were coded with respect to three dimensions: management, grouping unit, and content. A sample transcript is provided in Appendix C. Again, the management dimension (see Table 1) considered who was focusing the child's attention on the learning opportunity, the child (CM), or the teacher and child together (TCM). Activities were coded as TCM when the teacher was actively interacting with students. CM activities were coded when the child completed activities without the support of the teacher. These included activities in which the child worked independently or with peers (e.g., silent reading, buddy reading, completing worksheets).

The second coding dimension considered the unit in which students were grouped for the activity: Whole Class, Small Groups, Pairs, or Individual (see Table 1). Activities in which the entire class participated were designated as Whole Class. Small Group and Pair activities were coded when students worked together in smaller units to complete worksheets, read texts, write and edit stories, etc. Small Group activities involved groups of three or more students while Pair activities involved groups of two students. Individual activities were coded when children worked independently, including individual seatwork such as worksheets, journal writing, and silent reading.

The third dimension of the coding scheme captured the content of the classroom activities (see Table 1). Literacy, other academic (e.g., science) and non-academic (e.g., transitions) activities were coded. Literacy activities were first coded as to the broad content area targeted by instruction (i.e., Phoneme Awareness, Syllable Awareness, Morpheme Awareness, Onset/Rime Awareness, Word Identification/Decoding, Word Identification/ Encoding, Grapheme-Phoneme Correspondence, Fluency, Print Concepts, Oral Language,

Print Vocabulary, Comprehension, Text Reading, and Writing). Subcategories within each of the broad content areas were used to code activities more specifically (the complete coding manual is available upon request from the corresponding author). For example, a comprehension activity may have focused on a particular comprehension strategy, such as previewing text, making predictions or inferences, activating prior knowledge, etc. The coding scheme allowed all dimensions (management, grouping unit, content) to be coded for each classroom activity for each individual child (see Appendix C).

Results

Child Outcomes

In general, students made good gains in reading and vocabulary skills over the school year but there was substantial variability (see Table 4). For example, standard scores (mean = 100, SD = 15) on the letter-word identification test improved from 104 in August to 110 by May. Plus, there was a fair to high degree of stability (see Table 5) with fall scores moderately to highly correlated with spring scores. W scores, similar to Rasch scores (Mather & Woodcock, 2001), are on an interval scale and were used in all analyses.

Nature and Variation in Classroom Instruction

Descriptions of Classroom Activities—Comparing total amounts of language and literacy instruction provided during the language arts block across the school year (fall, winter, and spring) revealed fairly consistent amounts falling between 70 and 80 minutes per observation, averaged across all children and classrooms. About 10 minutes were spent, on average, in organization, which includes explaining how to do activities, instructing children on classroom routines, etc. This amount was consistent across the school year. Amounts of time students spent in non-instructional activities (e.g., waiting for instruction to begin, standing in line, etc.) increased from fall to winter and remained consistent through spring and comprised approximately 25% of the time observed during the language arts block. The remaining time was spent on content area instruction including science, social studies and mathematics (about 5% of the time observed).

Examining types of literacy instruction more closely (see Table 6) revealed substantial variability among classrooms. Generally, over the school year, most of the time spent in literacy instruction was spent reading text either aloud, in pairs, or individually (17 minutes, on average). Word identification encoding was the next most frequently observed type of literacy instruction observed (11 minutes on average), followed by writing (10 minutes). On average, small amounts of time were spent on phonological awareness (2 minutes) and grapheme-phoneme correspondence activities (4 minutes) and these amounts decreased from fall to spring (8 to 2 minutes respectively). Generally, more time in basic decoding skills instruction (e.g., alphabet activities, grapheme-phoneme correspondence) was observed in the fall compared to spring. More time in advanced meaning-focused activities (e.g., comprehension strategies) was observed in the spring compared to the fall. When we examined instruction amounts observed between children in the intervention and comparison classroom, no clear pattern of difference emerged at this level of analyses (i.e., management and grouping conflated).

Computing total amounts of instruction using multiple dimensions—Types of literacy instruction (see Table 1) were combined into either code- or meaning-focused variables and categorized by grouping (e.g., small group, individual) and management (e.g., teacher/child managed or child managed). For TCM instruction, whole class instruction was not included in the total observed amounts to be compared to the A2i recommended amounts. This provided 12 variables in all (see Table 1). Instruction provided by any adult in

the classroom (teacher, specialists, aides, etc.) was considered teacher/child-managed. Child-managed instruction included peer-managed instruction. Although important, examining peer interactions was beyond the scope of this study. Overall across all classrooms, 9 minutes of TCM-code focused small group instruction was observed in the fall, 6 minutes in the winter and 4 minutes in the spring. Approximately 12 minutes of CM code-focused instruction was observed in the fall, 8 minutes in the winter and spring. A fairly steady amount of TCM-meaning focused small group instruction was observed across the school year; 12 minutes in the fall, 13 minutes in winter and 11 minutes in spring. Generally increasing amounts of CM meaning focused instruction were observed during the school year; 16 minutes in fall, 20 minutes in winter, and 22 minutes in spring.

Computing Distance from Recommendation (DFR)—Using the A2i algorithms translated into SPSS syntax (version 14.0.1), recommended amounts, in minutes, for all four types of instruction were computed for each of the students (both intervention and comparison groups) using his or her fall scores for the fall and winter observations and his or her January scores for the spring observations.

The A2i recommended amount was subtracted from the actually observed amount of instruction for each child and histograms were examined for fall, winter and spring observations. Generally, teachers failed to provide recommended amounts and provided both more and less than the recommended amounts of teacher/child-managed code-focused (TCM-CF) and child-managed meaning-focused (CM-MF) instruction during observations. Children who shared the same classroom received very different amounts and types of instruction. For example, in one classroom, one child received more than 40 minutes of TCM-code focused instruction whereas another received less than 20. Moreover, only some of these amounts approximated the amounts recommended by A2i.

To compute mean amounts of instruction for each child across the fall, winter, and spring, while taking into account the nested nature of the observations (observations nested within teachers) and to impute missing observation data, we created two-level HLM models with repeated observations at level 1 and children at level 2. Empirical Bayes residuals for each child provided a mean amount and mean slope (minutes change per month) for each instruction type. Descriptive information is provided in Table 4.

To compute the distance from the A2i recommendation (DFR), we used the absolute value of the observed minus the recommended amounts. Considering children nested in classrooms, teachers were generally becoming more precise teaching the recommended amounts of TCM-CF instruction from fall to spring (smaller DFR) but were becoming less precise teaching the recommended amounts of CM-MF instruction. Children who received more precise amounts of TCM-CF instruction also tended to receive more precise amounts of CM-MF instruction ($r = .179$, p < .001).

Student Outcomes and Classroom Instruction

To examine the impact of classroom instruction on students' letter-word identification and passage comprehension skills, we built HLM models for each outcome. HLM was preferred because of the nested nature of our data, students nested in classrooms, nested in schools (Raudenbush & Bryk, 2002). We first examined 3-level models for each outcome (W scores) with students at level 1, classrooms at level 2, and schools at level 3. There was little variability between schools for either outcome (letter-word reading intraclass correlation [ICC] = .05; passage comprehension ICC = .08), which was fully explained once school level SES and Reading First status were added. Therefore, we used more parsimonious two level models with children nested in classrooms and school level variables, school SES and Reading First status added at the classroom level.

We built the two-level models systematically starting with an unconditional model (letter-word reading ICC = .17; passage comprehension ICC = .18), then added the classroom level variables, and then students' fall scores and the number of times they were observed (nj) to level 1. In the passage comprehension model, there was significant between-classroom variability in fall passage comprehension W scores so this variable was allowed to remain random at level 2. All other variables were tested and fixed at level 2. Although students' fall vocabulary did not significantly predict letter-word reading W score, we left it in the model because the intervention groups varied significantly on the measure. Using these models as our base, we then investigated the impact of total amounts of instruction types and then the impact of DFRs for each instruction type. Descriptive statistics for the models are provided in Table 4.

Total Amounts of Instruction—Entering total amounts of TCM-CF, CM-MF, TCM-MF and CM-CF instruction into our model at level 1 (remember these are child-level variables) revealed that, with one exception, TCM-MF, total amounts of instruction did not predict either outcome (see Table 7). Most notably, TCM-CF and CM-MF total amounts were not significantly associated with students' letter-word reading and passage comprehension score growth (i.e., residualized change).

In contrast, the more time students spent in TCM-MF instruction, the greater was their passage comprehension skill growth. The magnitude of this effect depended on the number of minutes of instruction above or below the mean (12 minutes). Thus, for example, according to these modeled results, children who received 30 minutes (18 minutes more than the mean amount) would be predicted to achieve passage comprehension scores about 5.6 points higher than the intercept of 467, which is a small to moderate effect size (d) of .37. There was a similar effect for letter-word reading which was marginally significant [coefficient = .36, t(488) = 1.882, p=.06].

Distance from Recommended Amount (DFR) Predicting Outcomes—As we hypothesized, the HLM results revealed that students in the intervention group had smaller TCM-CF and CM-MF DFRs, controlling for fall vocabulary and reading scores, than did students in the comparison classroom. [TCM-CF amount, intervention group = 1, coefficient =−1.337, t(45) = −3.189, p = .003; CM-MF amount, intervention group = 1, coefficient = −2.724, t(45) = −2.865, p = .007]. Thus, generally, children in the treatment classrooms were more likely to receive the A2i recommended amounts of instruction than were children in the control classrooms.

As anticipated, children who received instruction that more precisely matched the amounts recommended by the A2i algorithms (smaller DFR) achieved significantly stronger spring letter-word identification and passage comprehension scores, controlling for fall status, than did children who received less precise amounts (Table 8, Figure 4). Overall, the more precisely teachers provided the A2i algorithm recommended amounts of TCM-CF, controlling for slope, and CM-MF (i.e., smaller DFR) the stronger were their students' letter-word reading and passage comprehension score growth.

Additionally, TCM-CF slope predicted outcomes. Increasing precision over the school year (negative DFR slope) was associated with stronger letter-word reading and passage comprehension score growth.

Because the DFR variables were continuous, the effect size (d) changed as the variables' values changed. The magnitude of the effects may be gauged by examining Figure 4, which presents the results graphically using standard scores (normative sample mean = 100, SD = 15) as the outcome metric. Of note, there were no significant intervention/comparison group

by DFR interactions [TCM-CF amount coefficient = −.28, t(446) = −.177; TCM-CF slope coefficient = 2.59, t(466)=.435; CM-MF coefficient = .25, t(446) = 1.098], indicating that DFR predicted student outcomes regardless of whether their teacher was in the intervention or comparison group.

Students with stronger fall vocabulary scores were less likely to receive A2i algorithm recommended amounts of CM-MF than were students with weaker scores (see Figure 5) and this held across intervention and comparison groups. There were no significant interactions with either fall letter-word or passage comprehension score.

Neither teacher qualifications nor school characteristics predicted DFR and so were trimmed from the models. Specifically, the following variables did not predict TCM-CF or CM-MF DFR: teacher knowledge [TCM-CF amount coefficient = −.026, t(43)=−.853; CM-MF amount coefficient = .096, t(43) = 1.040] or experience [TCM-CF coefficient = −.040, t(43) = −.104, p = ; CM-MF coefficient = .513, t(43) = .649], school SES [TCM-CF coefficient = .009, t(43) = .665; CM-MF coefficient = −.031, t(43) = −.892], or Reading First status [TCM-CF coefficient = −.715, t(43) = −.972; CM-MF coefficient = 1.27, t(43) = .777]

Discussion

In this study, we investigated the classroom instruction provided to students in ISI intervention and comparison classrooms, asking whether child X instruction interactions were systematically related to students' reading skill outcomes. Specifically, we asked whether receiving A2i recommended amounts and types of instruction, informed by child X instruction algorithms, would lead to greater student reading achievement. Overall, students demonstrated grade appropriate growth in word reading and reading comprehension skills. There was, however, substantial variability.

The total amounts of teacher/child-managed code-focused, child-managed meaning-focused and child-managed code-focused instruction students received did not predict either letter-word reading or passage comprehension skill growth (see Table 7). When, for each child, we computed distance from recommendation (DFR), which was the absolute value of the observed amount minus the A2i recommended amount, we discovered that precision mattered. As hypothesized, teachers in the intervention group provided A2i recommended amounts more precisely than did the teachers in the comparison group. That is, children in the intervention group tended to have smaller DFRs than did children in comparison classrooms. We found that teachers in the intervention group were more likely to use small groups and to provide more child-managed meaning-focused instruction than were teachers in the comparison group. Both behaviors were encouraged through our professional development and scaffolded in the A2i software, and most likely contributed to the smaller DFRs observed in the intervention classrooms.

Importantly, the smaller the DFR for teacher/child-managed code-focused and child-managed meaning-focused instruction, the greater was the students' word reading and reading comprehension skill growth (see Table 8 & Figure 4). The effects were additive and were large when the most precise and least precise instruction was considered.

Taken together, these results support our hypothesis that it was receiving the precise amounts of individualized instruction that predicted students' reading outcomes and contributed to the significant overall treatment effect of the ISI intervention (Connor et al., 2007a). When we tested other sources of influence, they did not predict outcomes. Specifically, teachers' qualifications, teacher professional development, and school characteristics did not predict student outcomes. These findings provide strong evidence that

child X instruction interactions are causally implicated in the widely varying levels of student achievement observed between and within many classrooms.

Implications of Child X Instruction Interactions

As we conceptualize designing effective individualized instruction, there are several findings worth highlighting. First, our results indicated that the dimension *change over time* influenced students' literacy growth. That is, precisely providing recommended amounts of instruction, which change systematically over the course of the school year, as well as in response to students' changing skills, is associated with stronger student reading outcomes. Our results suggest that amounts that were appropriate at the beginning of the year for certain children were not by the end of the year.

Second, we found that going over the recommended amounts did not predict stronger achievement. It was not the case that more instruction was better.

Third, we verified that the relation between the amounts of instruction recommended and students' reading skills is non-linear. For example, examining the relation between students' assessed word reading skills and the recommended amount of teacher/child-managed code-focused instruction revealed that children whose skills were above grade level expectations needed fairly similar and modest amounts of teacher/child-managed code-focused instruction, especially toward the end of the school year, whereas children whose skills fell below grade level required exponentially more as they fell farther and farther behind grade level expectations. A simple linear relation would tend to recommend too much code-focused instruction for grade level readers and not enough for children who were performing well below grade level.

Taking these results together indicates that predicting appropriate amounts and types of instruction is not as straightforward as has been previously suggested (Fielding, Kerr, & Rosier, 2007). Much of reading research has focused on struggling readers (NRP, 2000; Rayner et al., 2001). Our results suggest that findings for struggling readers do not generalize seamlessly to more typical and above average readers. Such students appear to require substantially less time on the alphabetic principle to achieve proficient reading skills compared to children who begin first grade with weaker vocabulary and letter-word reading skills. Moreover, the complexity of the associations and the multidimensional nature of instruction may explain why so many teachers find it difficult to translate assessment results into effective practice (Roehrig, Duggar, Moats, Glover, & Mincey, in press in press).

Classroom Instruction

A compelling feature of observed classroom instruction was that, for the most part, teachers provided both meaning- and code-focused instruction. Indeed, across intervention and comparison classrooms, teachers provided substantial amounts of meaning-focused instruction using core reading curricula that are, by reputation and intent, principally code-focused (Foorman et al., 1998).

Notably, the more teacher/child-managed meaning-focused instruction children received, the stronger was their reading comprehension skill growth. This supports previous research findings (NRP, 2000; Pressley & Wharton-McDonald, 1997; C. E. Snow, 2001). Additionally, there were no child X instruction interactions; the effect of teacher/child-managed meaning focused instruction was a main effect. This indicates that meaning-focused instructional activities, such as reading with children and explicitly teaching comprehension strategies, had a positive effect on students' reading skill gains regardless of whether they had strong or weak reading skills. Although other studies have found evidence of child X instruction interactions for teacher/child-managed meaning-focused instruction

(Connor et al., in press; Connor et al., 2004a), these studies combined whole class and small group. It may be that the small group nature of the meaning-focused instruction contributed to its efficacy. Altogether, the results underscore the importance of meaning-focused instruction for all children.

We also found that children with the highest vocabulary scores in the fall were less likely to receive the recommended amounts of child-managed meaning-focused instruction when compared to children with lower vocabulary scores. Recent studies indicate that this finding may not be unique (Neal & Schanzenbach, 2007; Viadero, 2007). Many teachers tend to be less effective with children whose skills fall at the extremes compared with children whose skills are more typical of the classroom. Another reason may be that children with stronger vocabulary skills would be more likely to be reading above grade level. Especially in high poverty schools, teachers and administrators may assume that they are "doing just fine" and not provide them small group time and time to practice established skills. Although we might assume that it would be relatively easy to hand these children a book to read or encourage them to write in their journals, these actions are relatively counterintuitive and unlikely to happen, especially if the teacher provides substantial amounts of whole class instruction, which is almost necessarily directed at the children most typical of his or her classroom. Another consideration is that the core curricula focused on basic code-focused skills, especially at the beginning of the year. Thus independent reading and writing may have received less attention overall.

Research that includes careful observation of classrooms is expensive. However, these results and others (Pianta, 2006; Pianta et al., 2005; Pressley, Wharton-McDonald, Allington, Block, Morrow, Tracey, Baker, Brooks, & Crenin, 2001; Taylor et al., 2000) demonstrate that we cannot afford to treat the classroom like a black box. More distal sources of influence, while important and easier to measure (e.g., parental education, teacher qualifications), do not illuminate the mechanisms underlying student learning. Moreover, the results of this study illustrate that children who share the same classroom receive, by design or not, very different amounts and types of instruction. This may be the case, as well, for the social and emotional support they receive, which, research shows, also contribute to students' school success (Hamre & Pianta, 2005). Studying this within-classroom variability may offer us a better understanding of the complex and dynamic classroom environment and how we can optimize it to support all children.

Assessment

Using assessment to guide instruction is at the heart of individualizing instruction but can be difficult for many teachers (Roehrig et al., in press). A2i was developed with the specific purpose of helping teachers translate assessment results into viable recommendations for instruction, which could then be planned and implemented. Results illustrate that child X instruction interactions are really about precisely tailoring instruction to individual students' skills. Instruction that simply provides more time in reading without attending to children's specific reading strengths and weaknesses will be less effective overall.

A limitation of the A2i software used in this investigation was that although the algorithms provided recommended amounts, they did not suggest appropriate content or activities. More powerful algorithms are currently being developed that will provide teachers with guidance on appropriate content and activities, as well as recommended amounts of particular types of instruction. We are calling this the *Smart Lesson Plan*.

Spring DFRs predicted more of the variability in students' reading outcomes when computed using the January reassessment results rather than fall scores. This suggests that more frequent monitoring of students' skills should lead to more precise recommendations

and, hence, more effective patterns of instruction. Raudenbush (Raudenbush) describes such assessment informed instruction as a regime, where assessment and instruction form an iterative cycle with one informing the other (Cohen, Raudenbush, & Ball, 2003 Ball, 2003).

Complexity in the Classroom

This study demonstrates that effective instruction is highly complex and that optimal patterns of instruction differ for each child across a continuum. At the same time, the classroom system of instruction is comprised of relatively simple, albeit nonlinear functions (e.g., child X instruction interactions). These work together, yielding a complex system that appears to be chaotic but is, in fact, at least somewhat predictable in the presence of adequate feedback (assessment of children's skills). It is also dynamic because children are learning and as they learn their optimal patterns of instruction change.

This complex and dynamical systems application to instruction (Ford & Lerner, 1992; Op't Eynde & Turner, 2006; Sameroff & MacKenzie, 2003) suggests that classroom instruction that is (1) responsive to each student's changing skills and needs (cognitive, behavioral, social-emotional), (2) is carefully planned but appropriately modified in the moment-by-moment context of instructional opportunities, and (3) that relies upon a solid bed of rigorous research and specialized knowledge, may be more effective than instruction that considers the classroom environment more globally, more intuitively, less dynamically, and less diagnostically. A fundamental challenge that follows this work is that, once valid child X instruction interactions have been identified, teachers will need specialized support to be able to enact instructional interventions with precision. The A2i software is one attempt to support more precise enactment, and professional development represents a range of support systems. As we elucidate the multiple sources of influence that affect children's development and how children themselves influence the classroom learning environment, while continuing to identify effective strategies to improve the efficacy of instruction (and teachers' ability to provide such instruction), more children will achieve the foundational literacy skills that are crucial for their ongoing academic success

Acknowledgments

References

A2i. *A teacher version of A2i will be available after August 1, 2007. More information about A2i is available at* http://know.soe.umich.edu/A2I/login.asp. *The login is A2idemo and the password is ISI06!. The software and protocol are currently freely available to other researchers upon request.*

Adams, MJ. Beginning to Read: Thinking and Learning about Print. Cambridge, MA: The MIT Press; 1990.

Adams, MJ. Alphabetic anxiety and explicit, systematic phonics instruction: A cognitive science perspective. In: Neuman, SB.; Dickinson, DK., editors. Handbook of early literacy research. Vol. I. New York: Guilford Press; 2001. p. 66-80.

Al Otaiba S, Connor CM, Kosanovich M, Schatschneider C, Dyrlund AK, Lane H. Reading First kindergarten classroom instruction and students' phonological awareness and decoding fluency growth. Journal of School Psychology. 2008; 48:281–314. [PubMed: 19083361]

Barr, R. Beginning reading instruction: From debate to reformation. In: Pearson, PD.; Barr, R.; Kamil, ML.; Mosenthal, P., editors. Handbook of reading research. Vol. 1. New York: Longman; 1984. p. 545-581.

Bond GL, Dykstra R. The cooperative research program in first-grade reading instruction. Reading Research Quarterly. 1967; 2:5–142.

Borko, H.; Niles, J. Descriptions of teacher planning: Ideas for teachers and research. In: Richardson-Koehler, V., editor. Educators' handbook: A research perspective. New York: Longman; 1987. p. 167-187.

Bos C, Mather N, Dickson S, Podhajski B, Chard D. Perceptions and knowledge of preservice and inservice educators about early reading instruction. Annals of Dyslexia. 2001; 51:97–120.

Bracht GH. Experimental factors related to aptitude-treatment interactions. Review of Educational Research. 1970; 40:627–645.

Brophy JE. Teacher behavior and its effects. Journal of Educational Psychology. 1979; 71:733–750.

Brophy, JE.; Good, TL. Teacher behavior and student achievement. In: Wittrock, MC., editor. Handbook of research on teaching. 3. New York: Macmillan; 1986. p. 328-375.

Cameron CE, Connor CM, Morrison FJ. Effects of variation in teacher organization on classroom functioning. Journal of School Psychology. 2005; 43:61–85.

Chall, JS. Learning to read: The great debate. New York: McGraw-Hill Book Co; 1967.

Cohen DK, Raudenbush SW, Ball DL. Resources, instruction, and research. Educational Evaluation and Policy Analysis. 2003; 25:119–142.

Connor CM, Jakobsons LJ, Crowe E, Meadows J. Instruction, differentiation, and student engagement in Reading First classrooms. Elementary School Journal. in press.

Connor CM, Morrison FJ, Fishman BJ, Schatschneider C, Underwood P. THE EARLY YEARS: Algorithm-guided individualized reading instruction. Science. 2007a; 315:464–465. [PubMed: 17255498]

Connor CM, Morrison FJ, Katch EL. Beyond the reading wars: The effect of classroom instruction by child interactions on early reading. Scientific Studies of Reading. 2004a; 8:305–336.

Connor CM, Morrison FJ, Petrella JN. Effective reading comprehension instruction: Examining child by instruction interactions. Journal of Educational Psychology. 2004b; 96:682–698.

Connor CM, Morrison FJ, Slominski L. Preschool instruction and children's literacy skill growth. Journal of Educational Psychology. 2006a; 98:665–689.

Connor, CM.; Morrison, FJ.; Underwood, P. 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; 2006b, June; Paper presented at the Annual Meeting of the Society for the Scientific Study of Reading. Vancouver, BC.

Connor CM, Morrison FJ, Underwood P. 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. 2007b; 11:199–233.

Connor CM, Son S, Hindman A, Morrison FJ. Teacher qualifications, classroom practices, family characteristics and preschool experience: Complex effects on first graders' vocabulary and early reading outcomes. Journal of School Psychology. 2005; 43:343–375.

Cronbach LJ. The two disciplines of scientific psychology. American Psychologist. 1957; 12:671–684.

Cronbach, LJ.; Snow, RE. Aptitudes and instructional methods: A handbook for research on interactions. New York: Irvington; 1977.

Cronbach LJ, Webb N. Between-class and within-class effects in a reported aptitude X treatment interaction: Reanalysis of a study by G. L. Anderson. Journal of Educational Psychology. 1975; 67:717–724.

Darling-Hammond, L. Teacher quality and student achievement: A review of state policy evidence [Electronic Version]. Educational Policy Analysis and Archives. 2000. from http://epaa.asu.edu/epaa/v8n1/

DuFour, R.; Eaker, R. Professional learning communities at work: best practices for enhancing student achievement. Alexandria, VA: Association for Supervision and Curriculum Development; 1998.

Fielding, L.; Kerr, N.; Rosier, P. Annual growth for all students, catch-up growth for those who are behind. Kennewick, WA: The New Foundation Press; 2007.

Foorman BR, Francis DJ, Fletcher JM, Schatschneider C, Mehta P. The role of instruction in learning to read: Preventing reading failure in at risk children. Journal of Educational Psychology. 1998; 90:37–55.

Foorman BR, Schatschneider C, Eakin MN, Fletcher JM, Moats LC, Francis DJ. The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty schools. Contemporary Educational Psychology. 2006; 31:1–29.

Ford, DH.; Lerner, RM. Developmental systems theory: An integrative approach. Thousand Oaks, CA: Sage; 1992.

Freeman GD, King JL. A partnership for school readiness. Educational Leadership. 2003; 60:76–79.

Fuchs D, Mock D, Morgan PL, Young CL. Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. Learning Disabilities Research & Practice. 2003; 18:157–171.

Fuchs LS, Fuch D, Phillips N. The relation between teachers' beliefs about the importance of good student work habits, teacher planning, and student achievement. Elementary School Journal. 1994; 94:331–345.

Gersten R, Chard D, Baker S. Factors enhancing sustained use of research-based practices. Journal of Learning Disabilities. 2000; 33:445–457. [PubMed: 15495547]

Goldhaber DD, Brewer DJ. Does teacher certification matter? High school teacher certification status and student achievement. Educational Evaluation and Policy Analysis. 2000; 22:129–146.

Hamre BK, Pianta RC. Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? Child Development. 2005; 76:949–967. [PubMed: 16149994]

Helterbran VR, Fennimore BS. Collaborative early childhood professional development: building from a base of teacher investigation. Early Childhood Education Journal. 2004; 31:267–271.

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

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]

Lave, J.; Wenger, E. Situate learning legitimate peripheral participation. Cambridge: Cambridge University Press; 1991.

Lo, MY. Statistical analysis of interpretation and its application to data from the cooperative research program in primary reading instruction Unpublished Dissertation. State University of New York; Buffalo, Buffalo: 1973.

Mather N, Bos C, Babur N. Perceptions and knowledge of preservice and inservice teachers about early literacy instruction. Journal of Learning Disabilities. 2001; 34:472–482. [PubMed: 15503595]

Mather, N.; Woodcock, RW. Woodcock Johnson III tests of achievement: Examiner's manual. Itasca, IL: Riverside; 2001.

Mathes PG, Denton CA, Fletcher JM, Anthony JL, Francis DJ, Schatschneider C. The effects of theoretically different instruction and student characteristics on the skills of struggling readers. Reading Research Quarterly. 2005; 40:148–182.

Moats LC. The missing foundation in teacher education: Knowledge of the structure of spoken and written language. Annals of Dyslexia. 1994; 44:81–102.

Moats LC, Foorman BR. Measuring teachers' content knowledge of language and reading. Annals of Dyslexia. 2003; 53:23–45.

Morrison, FJ.; Bachman, HJ.; Connor, CM. Improving literacy in America: Guidelines from research. New Haven, CT: Yale University Press; 2005.

NAEP. The nation's report card. Washington DC: National Center for Education Statisticso. Document Number); 2005.

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

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

NICHD-ECCRN. Multiple pathways to early academic achievement. Harvard Educational Review. 2004; 74:1–29.

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

NRP. 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); 2000.

Op't Eynde P, Turner JE. Focusing on the complexity of emotion issues in academic learning: A dynamical component systems approach. Educational Psychology Review. 2006; 18:361–376.

Pianta, RC. Standardized observation and professional development. In: Martinez-Beck, MJZI., editor. Critical issues in early childhood professional development. Baltimore: Brookes Publishing; 2006. p. 231-254.

Pianta RC, Howes C, Burchinal M, Bryant D, Clifford R, Early D, Barbarin O. Features of Pre-Kindergarten Programs, Classrooms, and Teachers: Do They Predict Observed Classroom Quality and Child-Teacher Interactions? Applied Developmental Science. 2005; 9:144–159.

Piasta SB, Connor CM, Fishman B, Morrison FJ. Teachers' knowledge of literacy, classroom practices, and student reading growth. Manuscript in press, Scientific Studies of Reading. 2008

Pressley M, Wharton-McDonald R. Skilled comprehension and its development through instruction. School Psychology Review. 1997; 26:448–466.

Pressley M, Wharton-McDonald R, Allington R, Block CC, Morrow L, Tracey D, Baker K, Brooks G, Crenin J. A study of effective first-grade literacy instruction. Scientific Studies of Reading. 2001; 5:35–58.

Ramanathan, V. Enhancing the critical edge of (L2) teacher-education: Some issues in advanced literacy. In: Colombi, MJSMC., editor. Developing advanced literacy in first and second languages: Meaning with power. Mahwah, NJ: Lawrence Erlbaum Associates; 2002. p. 187-207.

Raudenbush, SW. Distinguished contributions to education research award (2006) lecture: How shall we study the causes and effects of classroom teaching?; 2007, April 12; Paper presented at the Annual Meeting of the American Education Research Association; Chicago.

Raudenbush, SW.; Bryk, A.; Cheong, YF.; Congdon, R.; du Toit, M. HLM6: Hierarchical linear and nonlinear modeling. Lincolnwood, IL: Scientific Software International; 2004.

Raudenbush, SW.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods. 2. Thousand Oaks, CA: Sage; 2002.

Raudenbush, SW.; Hong, G.; Rowan, B. Studying the causal effects of instruction with application to primary-school mathematics. Washington DC: Office of Educational Research and Improvement; 2002.

Rayner K, Foorman BR, Perfetti CA, Pesetsky D, Seidenberg MS. How psychological science informs the teaching of reading. Psychological Science in the Public Interest. 2001; 2:31–74.

Roehrig AD, Duggar SW, Moats LC, Glover M, Mincey B. When teachers work to use progress monitoring data to inform literacy instruction: Identifying potential supports and challenges. Remedial and Special Education. in press.

Sameroff AJ, MacKenzie MJ. Research strategies for capturing transactional models of development: The limits of the possible. Development and Psychopathology. 2003; 15:613–640. [PubMed: 14582934]

Showers B, Joyce B, Bennett B. Synthesis of research on staff development: A framework for future study and a state-of-the-art analysis. Educational Leadership. 1987:77–87.

Snow, CE. Reading for understanding. Santa Monica, CA: RAND Education and the Science and Technology Policy Institute; 2001.

Snow, CE.; Burns, MS.; Griffin, P., editors. Preventing reading difficulties in young children. Washington, DC: National Academy Press; 1998.

Snow RE. Aptitude-treatment interaction as a framework for research on individual differences in psychotherapy. Journal of Consulting and Clinical Psychology. 1991; 59:205–216. [PubMed: 2030178]

Taylor BM, Pearson DP, Clark K, Walpole S. Effective schools and accomplished teachers: lessons about primary-grade reading instruction in low-income schools. Elementary School Journal. 2000; 101:121–165.

US DOE. No child left behind: A toolkit for teachers. Washington DC: US Department of Education, Office of the Deputy Secretary; 2004.

Vaughn S, Coleman M. The role of mentoring in promoting use of research-based practices in reading. Remedial and Special Education. 2004; 25:25–38.

Viadero, D. NCLB Seen as Curbing Low, High Achievers' Gains [Electronic Version]; Educational Week. 2007. p. 7Retrieved July 2007, from http://lnk.edweek.org/edweek/index.html?url=/ew/articles/2007/07/16/43nclb_web.h26.html&tkn=HN0YoGR4vSrb8IGv%2FEqP44BxjGf7b%2BBy

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

Xue Y, Meisels SJ. Early literacy instruction and learning in kindergarten: Evidence from the early childhood longitudinal study -- kindergarten class of 1998–1999. American Educational Research Journal. 2004; 41:191–229.

Appendix A

Algorithms for Recommended Amounts of Child-managed Meaning-focused (CMMF) instruction

The A2i algorithm for CMMF works in the same way as for TCMCF using the following steps. The assessment metric is Age Equivalent (AE) in years. CMMF-A is the amount, CMMF-S is the slope (minutes/month), and CMMF-Recommended is the amount that is provided by the A2i software.

1. CMMF-A = ((3.76 − TO + 1.4*Vocabulary AE))/(.30)) − 14

2. CMMF-S = 10 − (.24*CMMF-A)

3. CMMF-Recommended = CMMF-A + .5*(CMMF-S*Month)

Thus the algorithms recommend that children with stronger vocabulary skills receive fairly high amounts of CM-MF throughout the school year in contrast to children with weaker skills. For these children, small amounts in the fall with increases each month through the end of the school year are recommended. The fall amount and magnitude of the increase depends on the child's vocabulary skills.

Appendix B

The ISI Intervention in the Classroom

There was considerable variability across the treatment group classrooms in how fully the ISI intervention was implemented (Connor et al., 2007a). Thus, it is useful to consider what comprised full implementation, which was the goal of the professional development provided. Teachers used the schools' core reading curriculum supplemented with other research-supported literacy activities during the schools' two hour literacy block. Children in the same classroom were at different levels or lessons within the curriculum depending on their assessed reading skills. Explicit teacher/child-managed instruction, both-code and

meaning-focused activities, were provided for the A2i recommended amounts of time in smaller groups. Group size varied from as many as 8 children to as few as one or two. Homogeneous groups based on learning needs were used consistently (Taylor et al., 2000; Wharton-McDonald et al., 1998) and there was a clear plan of instruction (Borko & Niles, 1987; L. S. Fuchs, Fuch, & Phillips, 1994), using either the A2i lesson plan or other district approved lesson plan.

Children followed a daily routine with a clearly displayed organizational chart showing them where to go and what to do (J. E. Brophy, 1979; Cameron, Connor, & Morrison, 2005 2005). During child-managed activities, students worked individually, in small groups, pairs or accomplished seat work. These activities were designed to provide practice for skills that had been learned during teacher/child-managed instruction. Center activities and materials developed by the Florida Center for Reading Research (http://www.fcrr.org/curriculum/studentCenterActivities.htm) were provided to all intervention teachers and were available to teachers in the comparison group. Classroom management was important (J. E. Brophy, 1979; J.E. Brophy & Good, 1986), transitions were quick and well rehearsed (Wharton-McDonald et al., 1998), and all of the children were engaged in meaningful literacy activities for virtually the entire literacy block (Pressley et al., 2001; Taylor et al., 2000; Wharton-McDonald et al., 1998).
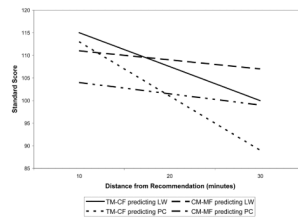
Figure 1.
Comparing the effects of Distance from Recommendation (DFR, the absolute value of the difference between the recommended and observed amounts of instruction) for instruction types on standard scores for Letter-word identification (LW) and Passage Comprehension (PC). Standard scores have a mean of 100 and a standard deviation of 15. Effect sizes increase with decreasing DFR. DFR range on the x-axis represents the 90th to the 10th percentile of the sample. All other variables, including TMCF DFR slope, were held constant at their sample mean.

Table 1

Putting the Dimensions of Instruction Together

	Teacher/child managed (TCM)		Peer & child managed (CM)	
	Code-focused (CF)	**Meaning-focused (MF)**	**Code-focused (CF)**	**Meaning-focused (MF)**
Whole Class	(TCM-CF) The teacher writes 'run' on the board and asks students to break the word into /r/ /u/ /n/ and then blend the sounds together to form /run/.	(TCM-MF) The teacher reads a book aloud to the class. Every so often he stops to ask the children to predict what is going to happen next.	(CM-CF) All students complete a workbook page on word families (e.g., cat, bat, sat) while the teacher sits at her desk and reviews assessment results.	(CM-MF) All students write in their journals while the teacher writes in her journal.
Small Group & Pair	(TCM-CF) The teacher reads a list of words aloud and the small group or pair of students put their thumbs up if they hear the long 'o' sound and thumbs down if they do not hear the sound.	(TCM-MF) While reading a book to a small group of children (or pair), the teacher asks students to make predictions about what will happen next.	(CM-CF) Two students take turns testing each other on reading sight words on flash cards.	(CM-MF) A group of students work together at a center using flash cards to make compound words, which they then define and use in a sentence.
Individual	(TCM-CF) The teacher works with an individual student and is timing how long it takes him to read a list of sounds. She then provides feedback on word attack and sight word strategies	(TCM-MF) During a shared reading activity, the teacher assists a student individually on using comprehension strategies to enhance understanding	(CM-CF) A student completes a worksheet where he must color the pictures for which each name includes the long 'a' sound.	(CM-MF) After listening to a book on tape, a student fills out a worksheet that asks her to answer questions about the characters and to provide a summary of the story.

Table 2

School Descriptives

School	Percentage of students eligible for Free and Reduced Lunch	Reading First	Mean Third Grade FCAT Reading Score	Core Curriculum	Total # of first grade classrooms	Intervention Group
1	93	yes	278	Reading Mastery	3	no
2	96	yes	286	Open Court	6	yes
3	88	yes	286	Open Court	6	no
4	82	yes	296	Reading Mastery	5	yes
5	69	no	294	Open Court	4	yes
6	67	no	316	Open Court	5	yes
7	57	yes	297	Open Court	5	no
8	37	no	326	Open Court	7	no
9	29	no	349	Open Court	5	yes
10	24	no	340	Open Court	6	no

Note. Intervention Group: yes = treatment, no = control; Reading First: yes = Reading First, no = Not Reading First

Table 3

Teacher and School Descriptives

Dependent Variable	Group	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
School Percentage of Children Eligible for Free or Reduced Price Lunch (FRL, (0=not eligible, 1=eligible)	Comparison	51.083	4.356	42.305	59.862
	Intervention	76.318	4.549	67.149	85.487
Reading First (1= Reading First, 0= Not Reading First so mean represents proportion)	Comparison	.458	.104	.248	.668
	Intervention	.500	.109	.281	.719
Teacher Knowledge Survey total correct fall[1]	Comparison	27.500	1.217	25.048	29.952
	Intervention	17.636	1.271	15.075	20.197
Teacher Knowledge Survey gain from fall to spring	Comparison	.083	1.656	−3.254	3.420
	Intervention	5.364	1.729	1.878	8.849
Higher Degree (1= Higher degree attained, 0= No higher degree)	Comparison	.208	.104	.000	.418
	Intervention	.455	.109	.236	.673
Years teaching experience	Comparison	9.958	2.032	5.863	14.053
	Intervention	11.489	2.122	7.212	15.766
Years teaching Grade 1	Comparison	5.708	1.122	3.448	7.969
	Intervention	4.239	1.172	1.877	6.600

[1] The highest possible score on the Teacher Knowledge Survey is 40 points.

Table 4

HLM Descriptives for Student- and Classroom-level Variables

Level 1 Student Descriptive Statistics				
	Mean	**SD**	**Minimum**	**Maximum**
Number of Times Observed	2.67	0.61	1.00	3.00
Fall WJ Letter-Word W Score	410.38	31.14	324.00	519.00
Fall WJ Letter-Word Standard Score	104.01	16.71	50.00	194.00
Fall WJ Picture Vocabulary W Score	479.02	10.91	442.00	513.00
Fall WJ Picture Vocabulary Standard Score	102.31	11.15	71.00	139.00
Fall WJ Passage Comprehension W Score	449.60	21.09	377.00	497.00
Fall WJ Passage Comprehension Standard Score	100.17	15.46	44.00	148.00
Spring WJ Letter-Word W Score	454.61	25.51	367.00	525.00
Spring WJ Letter-Word Standard Score	109.72	14.28	60.00	149.00
Spring WJ Passage Comprehension W Score	466.27	15.79	404.00	506.00
Spring WJ Passage Comprehension Standard Score	102.09	12.86	64.00	136.00
Spring WJ Picture Vocabulary W Score	483.48	10.72	456.00	528.00
Spring WJ Picture Vocabulary Standard Score	102.72	10.88	77.00	149.00
Teacher/Child-Managed Code-Focused Amount Total	6.34	4.23	1.37	30.05
Child-Managed Meaning-Focused Amount Total	21.01	6.85	7.85	48.42
Teacher/Child-Managed Meaning Focused Amount Total	11.63	5.07	3.79	26.39
Child-Managed Code-Focused Amount Total	9.83	4.15	4.57	25.61
Teacher/Child-Managed Code-Focused Distance from Recommendation Amount	19.62	2.29	12.95	27.84
Teacher/Child-Managed Code-Focused Distance from Recommendation Slope	−0.71	0.60	−3.15	1.00
Child-Managed Meaning-Focused Distance from Recommendation Amount	21.04	5.07	10.07	33.09
Child-Managed Meaning-Focused Distance from Recommendation Slope	0.57	0.41	−0.63	1.57

Level 2 – Classroom Descriptive Statistics	Mean	SD	Minimum	Maximum
		24.8		
Percentage of Students Eligible for FRL	62.43	8	24.00	96.00
School Participates in Reading First (=1)	0.47	0.50	0.00	1.00
Teacher in Intervention Group (=1)	0.49	0.51	0.00	1.00
Teacher Knowledge Survey Fall Score	22.77	7.63	9.00	36.00
Teacher Knowledge Survey Gain (Fall to Spring)	2.94	8.66	−19.00	20.00
Years of Education beyond Bachelor	0.32	0.52	0.00	2.00

Table 5

Correlations

	Fall WJ Picture Vocabulary W Score	Fall WJ Passage Comprehension W Score	Spring WJ Letter-word W Score	Spring WJ Passage Comprehension W Score	Spring WJ Picture Vocabulary W Score
Fall WJ Letter-word W Score	.367(**)	.795(**)	.769(**)	.677(**)	.446(**)
Fall WJ Picture Vocabulary W Score	-	.368(**)	.302(**)	.403(**)	.735(**)
Fall WJ Passage Comprehension W Score		-	.733(**)	.692(**)	.404(**)
Spring WJ Letter-word W Score			-	.845(**)	.398(**)
Spring WJ Passage Comprehension W Score				-	.492(**)

Note. Correlation is significant at the 0.01 level (2-tailed).

Table 6

Mean Amounts of Literacy Instruction for all Observations by Content Summed across Grouping and Management

Literacy Activity	Mean (minutes)	SD	Minimum	Maximum
Text reading total	17.48	19.86	.00	196.08
Word identification encoding total	10.83	15.23	.00	118.47
Writing total	10.22	16.88	.00	154.30
Comprehension total	9.22	11.55	.00	94.19
Print vocabulary total	8.50	9.84	.00	112.18
Word identification decoding total	7.98	10.19	.00	105.27
Grapheme-phoneme correspondence total	4.48	7.75	.00	63.45
Phonological awareness combined total	1.99	4.79	.00	64.71
Print concepts total	1.73	5.70	.00	59.49
Oral language total	1.63	3.47	.00	26.49
Sentence and text fluency total	1.07	3.37	.00	40.61
Phonological awareness total	1.05	3.03	.00	38.17
Letter and word fluency total	1.02	4.01	.00	67.36
Onset-rime total	.69	3.51	.00	63.53
Morphological awareness total	.63	2.50	.00	31.41
Oral vocabulary total	.43	1.22	.00	8.04
Syllable awareness total	.25	.94	.00	8.54

Table 7

Effects of Total Amounts of Instruction for Spring Letter-Word Reading (top) and Passage Comprehension (bottom)

Spring WJ Letter-Word Outcome Score as a Function of Total Amounts of Instruction in Minutes

Fixed Effect	Coefficient	Standard Error	t-ratio	Approx. df	p- value
Spring WJ Letter-Word W Intercept	454.770	1.751	259.732	44	<0.001
Child Level Variables					
Number of Times Observed	1.032	1.363	0.757	448	0.449
Fall WJ Letter-Word W score	0.411	0.042	9.818	448	0.000
Fall WJ Picture Vocabulary W score	−0.009	0.066	−0.130	448	0.897
Fall WJ Passage Comprehension W score	0.380	0.067	5.640	448	<0.001
Teacher/Child-Managed Code-Focused Amount	0.174	0.248	0.702	448	0.483
Child-Managed Meaning-Focused Amount	0.066	0.135	0.486	448	0.627
Teacher/Child-Managed Meaning-Focused Amount	0.361	0.192	1.882	448	0.060
Child-Managed Code-Focused Amount	−0.032	0.246	−0.130	448	0.897
Classroom Level Variables					
Percent of Students Eligible for FRL	−0.046	0.058	−0.792	44	0.433
Reading First	−0.215	3.172	−0.068	44	0.947

Random Effect	Standard Deviation	Variance Component	df	Chi- Square	p-value
Between Classroom Residual	4.012	16.093	44	80.149	0.001
Within Classroom Residual	14.611	213.494			
Deviance	3788.330				

Spring WJ Passage Comprehension Score as a Function of Total Amounts of Instruction in Minutes

Fixed Effect	Coefficient	Standard Error	t-ratio	Approx. d.f.	p-value
Spring WJ Passage Comprehension W Score Intercept	466.839	1.139	409.949	44	<0.001
Child Level Variables					
Number of Times Observed	1.307	0.861	1.519	448	0.129
WJ Fall Letter-Word W score	0.147	0.027	5.446	448	<0.001
WJ Fall Picture Vocabulary W score	0.195	0.048	4.079	448	<0.001
WJ Fall Passage Comprehension W score	0.262	0.043	6.063	448	<0.001

Spring WJ Letter-Word Outcome Score as a Function of Total Amounts of Instruction in Minutes

Fixed Effect	Coefficient	Standard Error	t-ratio	Approx. df	p- value
Teacher/Child-Managed Code-Focused Amount	−0.022	0.16	−0.135	448	0.893
Child-Managed Meaning-Focused Amount	0.071	0.11	0.644	448	0.520
Teacher/Child-Managed Meaning-Focused Amount	0.314	0.133	2.366	448	0.018
Child-Managed Code-Focused Amount	−0.002	0.125	−0.018	448	0.986
Classroom Level Variables					
Free and Reduced Lunch	−0.052	0.03	−1.729	44	0.090
Reading First	−1.294	1.888	−0.686	44	0.496

Random Effect	Standard Deviation	Variance Component	df	Chi- Square	p-value
Between Classroom Residual	2.257	5.094	44	67.157	0.014
WJ Fall Passage Comprehension W score	0.108	0.012	46	64.804	0.035
Within Classroom Residual	9.943	98.856			

Deviance	3440.469	

Table 8

Effects of Distance from Recommended Amounts of Instruction (DFR) for Spring Letter-Word Reading (top) and Passage Comprehension (bottom). DFR is the absolute value of the difference between the recommended and observed amounts of instruction.

Spring WJ Letter-Word Score as a Function of Distance from Recommended Amounts

Fixed Effect	Coefficient	Standard Error	t-ratio	Approx. df	p-value
Spring WJ Letter-Word W score Intercept	533.731	15.994	33.371	44	<0.001
Child Level Variables					
Number of Times Observed	0.850	1.378	0.617	449	0.537
Fall WJ Letter-Word W score	0.427	0.040	10.707	449	<0.001
Fall WJ Picture Vocabulary W score	0.019	0.073	0.258	449	0.797
Fall Passage Comprehension W score	0.364	0.066	5.525	449	<0.001
Teacher/Child-Managed Code-Focused Distance from Recommended Amount	−4.170	0.970	−4.298	449	<0.001
Teacher/Child-Managed Code-Focused Distance from Recommended Slope	−14.407	3.980	−3.619	449	0.001
Child-Managed Meaning-Focused Distance from Recommended Amount	−0.339	0.144	−2.352	449	0.019
Classroom Level Variables					
Percentage of Students Eligible for FRL	−0.032	0.065	−0.495	44	0.623
Reading First	−0.934	3.549	−0.263	44	0.794

Random Effect	Standard Deviation	Variance Component	d.f.	Chi-Square	p-value
Between Classroom Residual	3.993	15.941	44	80.721	0.001
Within Classroom Residual	14.454	208.914			

Deviance					
3778.609					

Spring WJ Passage Comprehension Score Outcome as a Function of Distance from Recommended Amounts in Minutes

Fixed Effect	Coefficient	Standard Error	t-ratio	Approx. d.f.	p-value
Spring WJ Passage Comprehension W score Intercept	522.622	13.826	37.801	44	<0.001
Child Level Variables					
Number of Times Observed	1.143	0.884	1.293	449	0.197
Fall WJ Letter-Word W score	0.164	0.026	6.391	449	<0.001
WJ Fall Picture Vocabulary W score	0.216	0.054	3.997	449	<0.001
WJ Fall Passage Comprehension W score	0.248	0.043	5.699	46	<0.001

Spring WJ Letter-Word Score as a Function of Distance from Recommended Amounts

Fixed Effect	Coefficient	Standard Error	t-ratio	Approx. df	p-value
Teacher/Child-Managed Code-Focused DFR Amount	−2.938	0.807	−3.640	449	0.001
Teacher/Child-Managed Code-Focused DFR Slope (minutes/month)	−10.541	3.423	−3.080	449	0.003
Child-Managed Meaning-Focused Distance from Recommended amount	−0.253	0.099	−2.561	449	0.011
Classroom Level Variables					
Free and Reduced Lunch	−0.035	0.034	−1.031	44	0.309
Reading First	−2.017	1.982	−1.017	44	0.315

Random Effect	Standard Deviation	Variance Componet	df	Chi- Square	p-value
Between Classroom Residual	2.014	4.055	44	64.221	0.025
Fall WJ Passage Comprehension W score	0.108	0.012	46	67.747	0.020
Within Classroom Residual	9.890	97.813			

Deviance	3433.500	

Appendix C

The following is a sample transcript of a coded observation followed by selected code definitions. The complete coding manual is available upon request from the corresponding author. Start time and end time are in seconds, timemin represents the total time in the activity in minutes. Child-managed instruction is coded as CM and teacher/child-managed instruction is coded TCM. Note that activities are coded for each student (see childid).

Start time	End time	Timemin	childid	Grouping	Management	Content	Modifier 1	Modifier 2
1517.56	2306.33	13.14617	A	Individual	CM	Noninstructional	Transition/Act	
2306.33	2500	3.227833	A	Wholeclass	TCM	WordID/encoding	Spelling/Written	Paper/Pencil
252.9	492.33	3.9905	B	Individual	CM	Noninstructional	Transition/Act	
492.33	684.63	3.205	B	Individual	CM	WordID/encoding	Copying	Paper/Pencil
684.63	2010.83	22.10333	B	Individual	CM	WordID/encoding	Copying	Paper/Pencil
2010.83	2306.33	4.925	B	Individual	CM	Noninstructional	Drawing/coloring	
2306.33	2500	3.227833	B	Wholeclass	TCM	WordID/encoding	Spelling/Written	Paper/Pencil
633.9	684.6	0.845	C	Individual	CM	Noninstructional	Transition/Act	
684.6	770.96	1.439333	C	Individual	CM	Noninstructional	Transition/Act	
770.96	1477.63	11.77783	C	Individual	CM	WordID/encoding	Copying	Paper/Pencil
1477.63	2306.33	13.81167	C	Individual	CM	Textreading	SSR	Text/Tradebook
2306.33	2500	3.227833	C	Wholeclass	TCM	WordID/encoding	Spelling/Written	Paper/Pencil
198.96	301.4	1.707333	D	Individual	CM	Noninstructional	Transition/Act	
301.4	684.63	6.387167	D	Individual	CM	WordID/encoding	Copying	Paper/Pencil
684.63	2020.13	22.25833	D	Individual	CM	WordID/encoding	Copying	Paper/Pencil
2020.13	2306.33	4.77	D	Individual	CM	Textreading	SSR	Text/Tradebook
2306.33	2500	3.227833	D	Wholeclass	TCM	WordID/encoding	Spelling/Written	Paper/Pencil
641.26	684.63	0.722833	E	Individual	CM	Noninstructional	Transition/Act	
684.63	986.03	5.023333	E	Individual	CM	Noninstructional	Transition/Act	
986.03	1745.43	12.65667	E	Individual	CM	WordID/encoding	Copying	Paper/Pencil
1745.43	2306.33	9.348333	E	Individual	CM	Textreading	SSR	Text/Tradebook
2306.33	2500	3.227833	E	Wholeclass	TCM	WordID/encoding	Spelling/Written	Paper/Pencil
105.5	684.63	9.652167	F	Individual	ComputerM	MF-TBD	vocabulary	games
684.63	1095.1	6.841167	F	Individual	ComputerM	MF-TBD		
1095.1	1988.1	14.88333	F	Individual	CM	WordID/encoding	Copying	Paper/Pencil
1988.1	2306.33	5.303833	F	Individual	CM	Textreading	SSR	Text/Tradebook
2306.33	2500	3.227833	F	Wholeclass	TCM	WordID/encoding	Spelling/Written	Paper/Pencil

Start time	End time	Timemin	childid	Grouping	Management	Content	Modifier 1	Modifier 2
316.23	641.26	5.417167	G	Individual	CM	Noninstructional	Transition/Act	
641.26	684.63	0.722833	G	Individual	CM	WordID/encoding	Copying	Paper/Pencil
684.63	1491.5	13.44783	G	Individual	CM	WordID/encoding	Copying	Paper/Pencil
1491.5	1910.3	6.98	G	Individual	CM	Textreading	SSR	Text/Tradebook
1910.3	2211.33	5.017167	G	Individual	ComputerM	CF-TBD	sound&spelling_activities	
2211.33	2306.33	1.583333	G	Individual	CM	WordID/encoding	Copying	Paper/Pencil
2306.33	2500	3.227833	G	Wholeclass	TCM	WordID/encoding	Spelling/Written	Paper/Pencil
7.6	87.6	1.333333	H	Individual	CM	Noninstructional	Transition/Act	
87.6	316.23	3.8105	H	Individual	CM	Noninstructional	Transition/Act	
316.23	684.63	6.14	H	Individual	CM	WordID/encoding	Copying	Paper/Pencil
684.63	1492.56	13.4655	H	Individual	CM	WordID/encoding	Copying	Paper/Pencil
1492.56	2306.33	13.56283	H	Individual	CM	Noninstructional	Drawing/coloring	
2306.33	2500	3.227833	H	Wholeclass	TCM	WordID/encoding	Spelling/Written	Paper/Pencil
105.5	301.4	3.265	J	Individual	CM	Noninstructional	Transition/Act	
301.4	684.63	6.387167	J	Individual	CM	WordID/encoding	Copying	Paper/Pencil
684.63	1021.16	5.608833	J	Individual	CM	WordID/encoding	Copying	Paper/Pencil
1021.16	2306.33	21.4195	J	Individual	CM	Text	reading	SSR
2306.33	2500	3.227833	J	Wholeclass	TCM	WordID/encoding	Spelling/Written	Paper/Pencil
87.6	198.93	1.8555	K	Individual	CM	Noninstructional	Transition/Act	
198.93	684.63	8.095	K	Individual	CM	WordID/encoding	Copying	Paper/Pencil
684.63	986.06	5.023833	K	Individual	CM	WordID/encoding	Copying	Paper/Pencil
986.06	2306.33	22.0045	K	Individual	CM	Text	reading	SSR
2306.33	2500	3.227833	K	Wholeclass	TCM	WordID/encoding	Spelling/Written	Paper/Pencil
622	684.63	1.043833	L	Individual	CM	Noninstructional	Transition/Act	
684.63	822.26	2.293833	L	Individual	CM	Noninstructional	Transition/Act	
822.26	2299.36	24.61833	L	Individual	CM	WordID/encoding	Copying	Paper/Pencil
2299.36	2500	3.344	L	Individual	ComputerM	Assessment	Accelerated Reader test	

Selected Code Definitions:

Word Identification/Encoding -- Subcategories: Sounding Out, Syllables, Syllable Types, Analogy/Word Families, Morphology, Spelling (Written, Oral), Rules, Copying

Word Identification/Encoding was coded for those activities allowing students to develop and practice their spelling skills. These activities focused on the spelling of single words and moved from pronunciation to print. Also included in the Word Identification/Encoding category were activities in which students copied single words from the blackboard (e.g., copying their spelling words for the week) and activities in which students gave oral spellings of words.

Text Reading -- Subcategories: Reading Aloud (Teacher, Student, Student Choral, Teacher/Student Choral, Technology), Silent Sustained Reading

Text Reading was coded for activities in which connected text was read without an explicit focus on building fluency. Text Reading activities necessarily involved the use of text.

Noninstructional – Subcategories: Disruption, Discipline (Individual, Class), Waiting, Transition (Activity, Clean Up, Line Up, Bathroom), Lunch, Snack, Nap, Game, Free Time, Song, Switch Mode, Off-Task

This includes all non-academic activities, defined as time spent that is not directly involved in instructional activities. Activities include transitions between activities (Transition/Act), waiting for instruction to begin, lining up for recess, and disruptions (e.g., announcements over the loud speaker).