Psychology in the Schools, Vol. 49(6), 2012
View this article online at wileyonlinelibrary.com/journal/pits

DOI: 10.1002/pits.21618

PREVENTING RETENTION: FIRST GRADE CLASSROOM INSTRUCTION AND STUDENT CHARACTERISTICS

JENNIFER LUCAS DOMBEK AND CAROL MCDONALD CONNOR

Florida State University and the Florida Center for Reading Research

Retention is a frequently used strategy to support children who are struggling academically. However, the strategy is costly, and research findings with regard to positive outcomes are mixed. This study examined whether efficacious reading instruction might reduce rates of retention in first grade. We also evaluated the reading instruction the students received compared with children with similar reading and vocabulary skills who were not retained. Additionally, we examined the impact students' self-regulation may have on grade retention. Findings reveal that it was significantly less likely for students to be retained at the end of first grade if their teacher was implementing more efficacious reading instruction. Moreover, there were substantial differences in the literacy instruction provided for children who were and were not retained. Finally, students in the efficacious reading condition who were retained exhibited significantly weaker self-regulation skills than did their matched-promoted peers. © 2012 Wiley Periodicals, Inc.

Grade retention refers to requiring a student who has completed a grade level to repeat that grade for an additional year. The premise of retention is that if students who are not reaching grade level expectations are given an extra year in the same grade, they will develop the academic skills they were initially unable to demonstrate, hence, limiting the likelihood of future school failure (Silberglitt, Appleton, Burns, & Jimerson, 2006; Westbury, 1994). One in four students will be considered for retention (Beebe-Frankenberger, Bocian, MacMillan, & Gresham, 2004) each year, whereas about 10% are actually retained; half of the retentions occur in the elementary years (National Center for Education Statistics, U.S. Department of Education, 2006). Moreover, recent cost–benefit analyses indicate that, unlike other pre-kindergarten through third-grade interventions, retention has a negative cost–benefit ratio, that is, the resources saved with regard to social and monetary capital are less than the costs to implement (Reutzel, Smith, & Fawson, 2005; Reynolds, Magnuson, & Ou, 2010). Common reasons for retention are failure to meet grade level expectations on high-stakes assessments, inability to make adequate progress in one or more content areas or poor performance in one or more content areas, and immaturity or age. Indeed, many studies have reported the short-term and long-term effects of retention on students in the areas of growth, learning, social adjustment, self-competence, and classroom behavior; a substantial number have shown that retention can have a negative effect on one or more of the aforementioned areas for the student (Beebe-Frankenberger et al., 2004; Griffith, Lloyd, Lane, & Tankersley, 2010; Holmes & Matthews, 1984; Hong & Raudenbush, 2005; Jimerson, 2001; Jimerson & Ferguson, 2007; Jimerson, Carlson, Rotert, Egeland, & Sroufe, 1997; Jimerson et al., 2006; McCoy & Reynolds, 1999). The purpose of this quasi-experimental study is to examine whether efficacious reading instruction provided to first-grade students might reduce rates of retention and to identify child characteristics that may contribute to grade retention.

We thank the ISI Project team members for their hard work collecting data. We would also like to recognize Phyllis S. Underwood for her support and contribution of ideas as we investigated grade retention. Additionally, we thank the children, parents, teachers, and school administrators without whom this research would not have been possible. This study was funded by grants R305H04013 and R305B070074, "Child by Instruction Interactions: Effects of Individualizing Instruction," from the U.S. Department of Education, Institute for Education Sciences, and by grant R01HD48539 from the National Institute for Child Health and Human Development. The opinions expressed are ours and do not represent views of the funding agencies.

Correspondance to: Jennifer Lucas Dombek, Florida State University–Florida Center for Reading Research, 2010 Levy Avenue, Suite 100, Tallahassee FL 32310. E-mail: jdombek@fcrr.org

Retention and Student Achievement

When comparing groups of students who were retained to a matched but continuously promoted group of students (hereafter referred to as matched-promoted), students who were retained generally achieved academically at lower or similar levels (Beebe-Frankenberger et al., 2004; Griffith et al., 2010; Holmes and Matthews, 1984; Hong & Raudenbush, 2005; Jimerson et al., 1997; Jimerson & Ferguson, 2007; McCoy & Reynolds, 1999). A meta-analysis of 44 studies indicated that students who were retained were performing about half a standard deviation unit below a matched-promoted group of students (Holmes &Matthews, 1984). Jimerson (2001) extended these findings with a meta-analysis of 20 studies addressing the efficacy of grade retention, yielding 175 analyses on academic achievement outcomes. When comparing students who were retained with matched-promoted students, 48% of the analyses showed no differences in academic achievement between the groups, and 47% favored matched-promoted students academically, whereas only 5% favored retained students academically, with students who were retained performing the poorest in reading, at .54 standard deviation units lower than the matched-promoted group. Highlighting long-term negative effects, McCoy & Reynolds (1999) found that when students were retained once or twice between kindergarten and seventh grade, they performed at lower levels in reading and mathematics by age 14 compared with same-age students. Similar results were found by Griffith et al. (2010) for reading achievement, with differences worsening over 10th and 12th grades. Additionally, Hong and Raudenbush (2005) found that kindergarten retention, on average, resulted in a loss of two thirds of a standard deviation of the outcome for academic growth in reading and math.

Evidence also suggests that retaining a student produces academic outcomes similar to matched-promoted students, indicating that the students who were retained did not benefit academically from the retained year (Jimerson et al., 1997; Jimerson & Ferguson, 2007; Silberglitt et al., 2006; Westbury, 1994). Silberglitt et al. (2006) found no difference in growth rate slopes when comparing the first and second year in the same grade for students retained once in elementary school or when comparing with a matched-promoted group at the end of sixth grade on a curriculum-based measure of reading fluency.

At the same time, some results have favored students who were retained over their matched-promoted peers, although gains tended to taper off in subsequent years (Mantzicopoulos & Morrison, 1992; Peterson, DeGracie, & Ayabe, 1987; Roderick & Nagaoka, 2005); again, findings are mixed (Wu, West, & Hughes, 2008). Peterson et al. (1987) found that when retention occurred in first or second grade, those retained outperformed matched-promoted peers in reading and mathematics by the end of the repeat year; however, in almost all cases, students lost this advantage within 2 to 3 years postretention. Similar findings for retention in kindergarten (Mantzicopoulos & Morrison, 1992) and third grade (Roderick & Nagaoka, 2005) support this notion of loss in gains over time. Conversely, Wu et al. (2008) found that although students who were retained in first grade had slower growth rates in reading and mathematics in the short term, faster growth rates in reading and mathematics were evident in the long term (4 years later). In general, it is not likely that students will benefit from retention (rather, retention will at best maintain and at worst hinder their academic achievement), that the cost–benefit ratio is negative, and that initial positive effects tend to diminish as students continue their school career.

Retention and Student Characteristics

When compared with matched-promoted peers, students who were retained have been described as more likely to exhibit poor self-regulation and adjustment, including aggressive behavior in the classroom (Jimerson & Ferguson, 2007). Self-regulation, a construct measured by tasks requiring students to utilize multiple cognitive, behavioral, and social–emotional skills, including,

attention, working memory, and inhibitory control, has been found to be higher in students who achieve at higher levels academically (Connor et al., 2010; Ponitz et al., 2008; Ponitz, McClelland, Matthews, & Morrison, 2009;). This finding suggests a relationship between self-regulation and academic achievements that is not typically assessed in students who are performing at lower levels academically but may impact their ability to learn in the same ways as their typically performing or matched-promoted peers.

In addition, these students have been described as more likely to exhibit lower levels of personal adjustment, self-concept, and attitude toward school (Holmes & Matthews, 1984); lower levels of perceived self-competence and higher rates of delinquency by age 14 (McCoy & Reynolds, 1999); and poorer emotional health, peer acceptance, and behavior in the classroom (Jimerson et al., 1997).

Alternative findings indicate that students who were retained had a short-term increase in ratings of school belongingness and peer-rated liking, compared with continuously promoted students, but this increase did not hold over time (Wu, West, & Hughes, 2010). Retained students showed improvements in peer acceptance, teacher perceived and peer perceived academic competence in the retained year (Gleason, Kwok, & Hughes, 2007), hyperactivity and engagement, peer-rated sadness and withdrawal (Wu et al., 2010), or no differences (Beebe-Frankenberger et al., 2004) compared with high-risk promoted students. However, the direction of these associations is unclear. It may be that particular student characteristics, such as self-regulation, make it more likely that a student will or will not be retained rather than that such behavior is caused by retention.

Intervention, Instruction, and Progress Monitoring

The No Child Left Behind Act initiated an emphasis on evidence-based reading instruction, including explicit instruction in phonological awareness, phonics, fluency, oral language, and comprehension (National Institute of Child Health and Human Development [NICHD], 2000). In addition, there is accumulating evidence that providing individualized instruction to meet students' specific learning needs is generally more effective than more global whole-class approaches (O'Connor, Bocian, Beebe-Frankenberger, & Linklater, 2010) which may help in reducing rates of retention. Unfortunately, there is little research reporting findings for classroom instruction, programs, strategies, interventions, curriculum, or support provided for students prior to retention decisions being made (Karweit, 1991, 1999; Stone & Engel, 2007), but some research has been conducted. Students who were retained have been found to receive fewer academic services during the following school year (the retained year) and to experience a larger decrease in services from the first to second year, indicating that retention, rather than other services, was considered to be the intervention (Schnurr, Kundert, & Nickerson, 2009; Peterson and Hughes, 2011).

There is evidence indicating that for at-risk students or those who have been retained, teachers are providing additional support in the classroom. When surveyed on frequency of strategy use to meet the needs of retained or at-risk students, teachers reported most frequently employing cooperative learning strategies, group work, small-group instruction, instructional support from paraprofessionals or volunteers, and one-to-one tutoring to meet the needs these students (Picklo & Christenson, 2005). A study by Abbott et al. (2010) employed small-group intervention, in addition to the regular classroom literacy instruction, during the repeat year and the following school year (2 years total) for students who had been retained in kindergarten or first grade and for their matched same-grade peers. Kindergarteners did not benefit from retention, even after the intervention was employed. However, first-grade students who were retained eventually outperformed their matched same-grade peers and were able to reach average range for literacy outcomes when provided with the intervention. Although both studies actively engaged students in learning by employing evidence-based practices, lack of a control group makes it difficult to determine whether using such strategies might reduce retention rates. The importance of a study investigating retention and the instruction

students receive is necessary to further understand the practice of grade retention (Abbott et al., 2010; Hong & Raudenbush, 2005; Jimerson et al., 2006; Stone & Engel, 2007). If we can find a way to sufficiently meet the specific learning needs of each student, we will be better able to promote academic growth for all students (Beebe-Frankenberger et al., 2004). Investigating the instruction provided to students who were retained and their matched-promoted peers is an important aim of the current study.

Efficacious Reading Instruction: Individualized Student Instruction

As noted previously, accumulating evidence reveals that research-informed literacy instruction can promote stronger gains in students' reading skills (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; NICHD, 2000; Torgesen, 2002) and that using assessment to inform instruction to differentiate reading activities can also improve the effectiveness of reading instruction (Taylor, Pearson, Clark, & Walpole, 2000; Torgesen, 2000; Wharton-McDonald, Pressley, & Hampston, 1998). One intervention that explicitly trains teachers to use assessment to differentiate reading instruction and to implement evidence-based reading practices is the Individualizing Student Instruction (ISI) intervention (Connor, Morrison, Fishman, Schatschneider, & Underwood, 2007) and is the focus of this study. The intervention provides teachers with professional development, utilizing a coaching model (Bos, Mather, Narr, & Babur, 1999).

As part of the ISI training, teachers are taught to use Assessment to Instruction (A2i) software, which provides salient links between assessment results and recommendations for each student with regard to amounts and types of literacy instruction (Connor, Morrison, Fishman, & Schatschneider, 2011). A2i recommends the amount and type of instruction needed for each student based on assessment results. These recommendations were developed based on child-by-instruction interactions as discussed in previous research by Connor, Morrison & Katch (2004). The algorithms used by A2i produce recommendations based on initial student assessments and sets an end-of-year goal for 1 year's worth of growth or grade-level achievement for each student, whichever is greater. Recommendations from A2i are updated with each set of assessments administered to students to be sure instructional recommendations change as the students' skills develop. These recommendations suggest specific amounts (minutes per day) and types (code-focused or meaning-focused and teacher-managed or child-managed) of instruction for each student. Research has shown that students' reading skill growth in first grade is improved when receiving explicit and systematic instruction in phonological decoding (code-focused) instruction (National Reading Panel, 2000; Snow, Burns, & Griffin, 1998). Code-focused instruction would include the alphabetic principle, phonological awareness, sight word reading, and phonics activities. Activities that incorporate reading aloud, teaching comprehension strategies, repeated reading, and word meaning, all considered meaning-focused activities, have been found to predict first-grade outcomes as well (Connor, Morrison, & Petrella, 2004; National Reading Panel, 2000; Snow, 2001). A2i also provides teachers with planning and professional development resources.

Overall, students in the A2i classrooms achieved greater gains in passage comprehension and word reading compared with students in the control group (Connor, Morrison, Fishman et al., 2007). In addition, the more time teachers spent using A2i, the stronger their students' end-of-year reading skills were. As well, how precisely teachers provided the A2i recommended amounts of instruction or the distance from recommendation (DFR) was predictive of students' reading outcomes, whereas the total amounts of teacher/child-managed, code-focused (TCM-CF); child-managed, code-focused (CM-CF); teacher/child-managed, meaning-focused (TCM-MF); and child-managed, meaning-focused (CM-FM) instruction did not predict student reading gains (Connor, Piasta, et al., 2009). Students in ISI classrooms had smaller DFR than did students in control classrooms and DFR was predictive of student outcomes whether or not the student was in an ISI

intervention or control classroom. These results have been replicated in first and third grade (Connor, Morrison, Fishman, Guilianai, et al., 2011; Connor, Morrison, Fishman, & Schatschneider, in press).

Rationale and Study Aims

Altogether, the extant literature indicates that retention is a frequently used but not highly effective intervention for students who are struggling with reading. It has a negative cost–benefit ratio, and early gains tend to be attenuated as students progress through school. There is some evidence that focused interventions can improve the outcomes for students who have already been retained but not on whether such instruction might actually reduce the rates of retention. We hypothesized that increasing the efficacy of first-grade instruction through individualizing and, thereby, improving students' reading skill gains, might potentially reduce rates of retention and avoid its largely deleterious effects. Using a randomized control study design, the aims of this study are threefold: (1) to examine whether differentiated instruction designed to meet students' individual literacy learning needs, compared with business-as-usual practices, impacts rates of first graders' retention; (2) to examine whether there are differences in the literacy instruction provided to students who were retained compared with their matched-promoted peers; and (3) to identify other potential student characteristics, specifically self-regulation, that might predict retention. With regard to Aim 3, it is possible that students who begin the school year with weaker self-regulation and then continue to demonstrate weak self-regulation in the minds of their teachers are more likely to be retained than are students whose self-regulation is generally stronger at the beginning of the year. In this case, self-regulation would more likely contribute to decisions to retain rather than represent a result of retention.

Method

Research Design

This study used both experimental and quasi-experimental designs. For Aim 1, we examined rates of retention for students whose schools were matched and randomly assigned to implement the ISI intervention or to wait 1 year—a delayed treatment group. Each school was matched on percentage of children qualifying for free and reduced-priced lunch (FARL), the mandated Florida Comprehensive Achievement Test (FCAT) third-grade reading scores, and Reading First status. Then, one school from each pair was randomly assigned to the ISI treatment; the other school in each pair was assigned to the delayed treatment control condition. The teachers in the delayed treatment control group implemented the ISI intervention during the following school year and represented a business-as-usual control for the purposes of this study. For Aims 2 and 3, we identified students who had been promoted but had similar word reading and vocabulary skills and, when possible, attended the same classroom. This provided a matched-promoted group. By observing classrooms and assessing students' academic and self-regulation skills at the beginning and at the end of the year, we could investigate whether instruction was associated with whether or not a student was retained and began to disentangle whether certain student characteristics might predict decisions to retain or actually be the result of retention.

Participants

Participants were part of the ISI study, which is an ongoing cluster-randomized control field trial and longitudinal study of student achievement. The participating schools were located in one demographically diverse Florida school district. Half of the schools were participating in the Florida Reading First program, a federally funded program that provides support for low-income and

Table 1
Group Descriptions

	Retained			Promoted		
	ISI	Control	Total	ISI	Control	Total
N	21	42	63	16	35	51
Female	10	20	30	9	16	25
Male	11	22	33	7	19	26
African American	90.48%	52.38%	65.08%	87.50%	51.43%	62.75%
White	4.76%	45.24%	31.75%	6.25%	42.86%	31.37%
Other	4.76%	2.38%	3.17%	0	5.72%	5.88%
FARL	92%	50%	60%	75%	63%	66%
Fall LW	91.05 (11.1)	93.29 (12.7)	92.53 (12.3)	93.63 (11.0)	91.71 (12.6)	92.31 (12.0)
Fall PV	93.48 (8.3)	99.85 (9.5)	97.69 (9.6)	93.44 (8.4)	100.83 (11.9)	98.51 (11.4)
Fall PC	83.93 (13.10)	85.10 (12.9)	84.72 (12.8)	95.43 (10.2)	92.26 (10.0)	93.24 (10.1)
Fall HTKS	22.08 (11.7)	26.5 (8.8)	25.13 (9.9)	31.07 (5.9)	31.14(5.9)	31.12 (5.8)
Winter SSRS PB	107.67 (13.0)	108.87 (15.9)	108.55 (15)	102.20 (12.2)	108.42 (15.4)	106.69 (14.7)
Winter SSRS AC	76.78 (7.4)	78.38 (9.5)	77.94 (8.9)	85.20 (12.5)	87.85 (10.4)	87.11 (10.9)
Winter SSRS SSK	84.44 (9.98)	86.46 (1.9)	85.91 (13.6)	96.90 (14.5)	81.31 (16.8)	92.86 (16.2)
Spring LW	100.69 (13.2)	98.41 (13.7)	99.00 (13.5)	107.79 (10.5)	102.35 (13.7)	103.94 (13.0)
Spring PV	96.85 (7.1)	98.95 (9.1)	98.40 (8.6)	96.50 (8.2)	98.15 (10.6)	97.67 (9.9)
Spring PC	92.46 (12.5)	91.43 (12.6)	91.70 (12.5)	98.57 (10.3)	95.44 (10.1)	96.35 (11.6)

Note. FARL is the percentage of children who qualify for free and reduced-priced lunch. Reading, vocabulary, and SSRS scores are presented as standard scores (SD), where 100 (15) represents mean achievement. LW = letter-word identification; PV = picture vocabulary; PC = passage comprehension; SSRS = Social Skills Rating System; PB = problem behaviors; AC = academic competence; SSK = social skills.

poorly performing schools to improve instruction. In all, 11 schools, 53 teachers, and 815 students participated in the larger ISI study during the 2005-2006 school year.

Of the 815 first-grade students included in the ISI efficacy study, 63 (7.73%) from the larger study were retained. Thus, participants in this study included 114 students who entered first grade (63 retained and 51 matched-promoted students) and their 43 teachers from 10 different schools (see Table 1). The students in this study were divided into four groups according to retention status and assignment to the ISI intervention (treatment) or control condition: treatment-retained (TR)—21 students who were in the treatment condition and retained at the end of first grade; and treatment-promoted (TP)—16 students who were in the treatment condition and promoted to grade; control-retained (CR)—42 students who were in the control condition and retained at the end of first grade; and control-promoted (CP)—35 students who were in the control condition and promoted to second grade. Tables 1, 2, and 3 provide descriptive information for these groups.

The decision to retain a student was based on the district promotion policy for first-grade students, to which teachers and schools had to adhere. To be promoted, a student must have scored at or above the 25th national percentile on the SAT-10 and mastered all of the Core Benchmarks at 70% or better for Reading and Math. A third assessment was utilized if the students did not meet the Reading and/or Math requirements. In addition, the student had to master all of the core benchmarks at 70% or better for Science and Writing and achieve a 3.5 or better (out of 6) on the district-wide Writes Upon Request (a measure similar to the fourth-grade FCAT writing assessment that is administered to students in the third and fourth nine weeks of the school year). The ISI study personnel did not participate in deciding whether a student would be retained.

Table 2
Teachers and Students per School

			Teachers			Number of Students	
	School	Condition	In Larger Study	In Retention Study	Who Retained Students	Retained	Matched
	B	Control	6	6	6	7	6
	D	Control	3	2	2	5	5
	E	Control	7	6	4	7	6
	G	Control	4	4	3	5	4
	I	Control	6	5	5	18	14
Total	5		26	23	20	42	35
	A	ISI	6	4	3	4	4
	C	ISI	6	5	4	4	4
	F	ISI	4	3	2	2	1
	H	ISI	6	6	5	10	6
	J	ISI	3	2	1	1	1
	K	ISI	2	0	0	0	0
Total	6		27	20	15	21	16

Table 3
Spring Letter-Word Identification, Passage Comprehension, and Picture Vocabulary Means Across Groups

				95% Confidence Interval	
Dependent Variable	Retention Study Group	Mean	Std. Error	Lower Bound	Upper Bound
Spring Letter-Word	Control Retained	420.35	3.44	413.51	427.19
Identification W Score	Control Promoted	441.97	3.59	434.84	449.10
	Treatment Retained	419.00	5.81	407.47	430.54
	Treatment Promoted	439.43	5.60	428.31	450.54
Spring Passage	Control Retained	444.78	2.47	439.88	449.69
Comprehension W Score	Control Promoted	458.59	2.58	453.47	463.71
	Treatment Retained	442.69	4.17	434.41	450.97
	Treatment Promoted	454.79	4.20	446.81	462.77
Spring Picture	Control Retained	478.84	1.51	475.84	481.83
Vocabulary W Score	Control Promoted	479.59	1.57	476.47	482.71
	Treatment Retained	475.85	2.54	470.80	480.90
	Treatment Promoted	475.64	2.45	470.78	480.51

Note. Means are presented as W scores.

The retention of students was confirmed by research staff after reviewing school records. However, information regarding the specific areas in which the students were not able to meet grade-level expectations was not available to study researchers. Interviews with teachers revealed that the most common reason for retention was failure to meet promotion expectations in reading.

The retained and promoted groups were matched on school, gender, and fall Letter-Word Identification and Picture Vocabulary W scores. In four cases, gender was not matched to ensure that students who were retained were matched to students with a similar W score at the same school. In

six cases, a match was shared with another student who was retained to ensure the closest possible matches on achievement at the same school. There were also 3 TR and 4 CR students for whom a match was not available because there was not a promoted student at the same school with a comparable fall Letter-Word Identification score.

Forty-three teachers in 10 schools were included in this retention study; five schools were in the treatment condition, and five schools were in the control condition, with a total of 20 teachers in the treatment condition and 23 teachers in the control condition (see Table 2). Among the teachers, 92.9% were women, 66.7% were Caucasian, and 33.33% were African American. Therefore, this study only included 10 schools and 43 teachers of the original 53 ISI study teachers and 11 schools because only teachers who had classrooms containing students who were retained or those selected as matched-promoted peers were included.

Measures

Students' language and literacy skills were assessed in the fall, winter, and spring of each school year on a battery of reading, language, and cognitive tasks, including self-regulation. The assessments were administered individually to children by a trained researcher in a quiet location in the school.

Reading. Word reading skills were assessed using the Letter-Word Identification subtest from the Woodcock-Johnson Tests of Achievement-III (WJ; Mather and Woodcock, 2001). The WJ was chosen as the means to assess students because of its frequent use in schools and in educational research. This task required students to identify letters and words of increasing difficulty out of context, with reliabilities of .88 and above. Reading comprehension was assessed using the WJ Passage Comprehension subtest, which required students to complete a cloze procedure; students were asked to read a short passage and had to provide missing words. The Passage Comprehension test has reliabilities of .73 and above.

Vocabulary. Expressive vocabulary was assessed using the WJ Picture Vocabulary subtest (reliabilities of .70 and above for this age group), which required the students to look at pictures of increasingly unfamiliar objects and appropriately name what they saw.

The reading and vocabulary scores were used by the A2i software to compute recommended amounts of TCM-CF and CM-MF amounts (minutes/day). The software used specific end-of-year outcome targets, which were computed by adding .9 to students' fall Letter-Word Identification grade equivalent (GE) score (which represents 1 school year increase). The minimum target GE was 2.1, based on district expectations for end-of-year achievement. Intent-to-treat models for Years 1 and 2 revealed significant effects of treatment for Letter-Word Identification and Passage Comprehension, with effects sizes (d) of .27 and .50, respectively (Connor, Morrison, Fishman et al., 2007; Connor, Morrison, Schatschneider et al., 2011).

Self-Regulation. Self-regulation was evaluated using two assessments. One assessed students directly and the other indirectly using teacher report. We consider self-regulation to be a multi-component construct that taps students' attention, task inhibition, and working memory, and is represented across disciplines as cognitive, behavioral, and social–emotional regulation. To measure self-regulation directly in the fall and spring, we used the Head-Toes-Knees-Shoulders (HTKS) task, which is an expanded version of the Head-to-Toes task utilized in prior research with preschool and kindergarten students (Ponitz et al., 2008). This expanded task, which has been utilized with kindergarten and first-grade students, requires students to listen to the examiner's directions, remember rules for actions the student is to produce, and control gross motor impulse responses in order to follow the rules of the task (Connor et al., 2010; Ponitz et al., 2009). Fall performance on this task has been found to predict spring academic performance, when controlling for background factors

and fall academic performance (Ponitz et al., 2009). To measure teachers' perception of students' self-regulation, we used the Social Skills Rating System (SSRS; Gresham & Elliot, 1990), which is a widely used teacher survey designed to assess students' social skills, behavior problems, and academic competence.

Classroom Observational Data and Variables

Videotaped classroom observations were conducted in the fall, winter, and spring to compare instruction in treatment and control classrooms. We coded classroom observations using the ISI Coding System (Connor et al., 2009), which quantifies the amounts and types of instruction and non-instruction for randomly selected individual students in the classroom. Research assistants who coded the videos were blinded to study condition to the extent possible. Each observation lasted the duration of the literacy block, which was usually 90 minutes.

Observations were scheduled with the teacher and were conducted by trained research assistants. Each classroom was filmed with two video cameras, equipped with wide-angle lenses to be sure all aspects of classroom instruction would be captured. In addition, research assistants recorded detailed field notes, including descriptions of the students and activities that may be difficult for the camera to capture (i.e., content on student worksheet, materials being used at a student or teacher center, etc.). Teachers were informed that the purpose of conducting videotaped observations was to document the types of academic experiences students were receiving in the classroom to help researchers better understand the relationship between students' academic abilities and their school experiences. A common concern about conducting videotaped observations is the extent to which the observations might contribute to changes in teacher or student behavior. We believe that the presence of a videographer in the classroom minimally influenced teacher and student behaviors. Teachers were asked to conduct class as usual, given that those in the treatment condition were providing instruction aligned with the goals of the study. As for student behavior, it is unlikely that videographers were seen as strangers in the classroom, as students had been interacting with project staff throughout the course of the school year during individual student assessments.

Classroom observation videos were downloaded and coded in the ISI study laboratory using Noldus Observer software (Noldus Information Technology, 2010). To capture the multiple dimensions of instruction identified by Connor et al. (Connor, Morrison, & Katch, 2004; Connor, Morrison, & Petrella, 2004; Connor, Morrison, Fishman, et al., 2007), each activity was coded according to three dimensions including a content code (word decoding, print awareness, etc.), a grouping code (whole class, small group, etc.), and a management code (teacher/child managed [TCM], and child-managed [CM]). The complete protocol and coding manual are available on request. For reliability purposes, all coders had to achieve reliabilities of at least .70 (Cohen's kappa) with the lead coder before they were allowed to code independently. Reliability was checked for approximately 10% of the coded observations and ranged from .94 to .98 (kappa) for duration ($\pm$ 10 seconds) and activity. Thus, reliability on all coded data remained high.

We were interested in examining the extent to which teachers provided students differing amounts and types of instruction; therefore, we coded video at the student level. As a result, Child A might be reading independently at the library center (child-managed text-reading, which is considered meaning-focused, CM-MF), whereas Child B might be working with a small group at the teacher table on a phonological awareness activity (TCM phonological awareness, which is considered a code-focused activity, TCM-CF). Any activity lasting 15 seconds or more was coded. Non-instructional activities (transitions, off-task behavior) were also coded to clearly differentiate between time spent in instruction and time spent in non-instructional activities. In this study, we examined the amounts of TCM-CF and CM-MF small-group instruction provided to students for

Table 4
Examples of Classroom Literacy Activities Categorized by Instruction Variables

	Teacher/Child Managed	Child Managed
Code-Focused	• Alphabet activity • Letter sight-sound • Initial consonant stripping • Word segmentation • Phonics activities • Phonological awareness activities • Sight word reading	• Spelling • Independent repeated reading of words • Phonics activities • Phonological awareness activities • Non-word reading activities • Computer activities code-focused
Meaning-Focused	• Vocabulary • Teacher read aloud • Student read aloud, choral • Group writing, writing • Instruction, model writing • Listening comprehension • Discussion • Repeated reading of text • Timed reading of connected text • Reading comprehension strategies use	• Student read aloud, individual • Buddy reading • Sustained silent reading • Reading comprehension worksheets • Student individual writing

Note. From Connor, C. M., Morrison, F. J., & Underwood, P. (2007). A second chance in second grade: The independent and cumulative impact of first and second grade reading instruction and students' letter-word reading skill growth. *Scientific Studies of Reading*, 11(3), 199–233. Reprinted with permission of the author.

a subset of the retained and promoted students in this study ($n = 61$; 29 retained; 32 matched-promoted) who had been randomly selected for coding.

In this study, we used three specific instructional variables from these data because they predicted gains in students' reading skills (Connor et al., 2009). They were small-group TCM-CF, CM-MF, and TCM-MF (see Table 4 for examples of activities that fall under these instructional variables). ISI intervention teachers were specifically taught to differentiate instruction during centers using small groups. Means for amounts provided to each group (TR, TP, CR, and CP) are provided in Table 5. A2i provided specific recommended amounts of TCM-CF and CM-MF, so we computed DFR for them. DFR amounts (min) represent the absolute value of the difference between the A2i recommended amount and the observed amounts of instruction. Smaller DFRs are consistently associated with greater reading skill gains (Connor et al., 2009).

We did not compute DFR for TCM-MF because there were no A2i recommendations for this kind of instruction, so the recommended amount was set at the same level for all students. However, results showed that the more time students spent in small-group TCM-MF, the greater were their reading skill gains (Connor et al., 2009). Thus, our third variable was total amount of time in TCM-MF.

Results

In general, the students in this study, both those who were retained and their matched-promoted peers, were performing below-grade-level expectations at the beginning of first grade (see Table 1). For example, students in the retained group (both ISI and control) generally achieved standard scores on the Letter-Word Identification assessment of about 92, which is 8 points below a typical standard score of 100. Promoted students (both ISI and control groups) also achieved a standard score of

Table 5
Total Mean Amounts of Teacher/Child Managed Meaning-Focused, Teacher/Child Managed Code-Focused, Child-Managed Meaning-Focused, and Child-Managed Code-Focused Instruction (Minutes) Across Groups

					95% Confidence Interval for Mean			
		Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Min	Max
TCM-MF	Control Retained	9.70	3.93	0.86	7.91	11.49	4.02	20.763
	Control Promoted	11.26	5.39	1.12	8.92	13.59	4.58	24.65
	Treatment Retained	13.21	5.44	1.92	8.66	17.75	7.87	22.89
	Treatment Promoted	12.64	2.68	0.89	10.58	14.69	8.98	16.49
	Total	11.18	4.68	0.60	9.98	12.38	4.02	24.65
TCM-CF	Control Retained	4.66	1.82	0.40	3.83	5.49	2.48	10.43
	Control Promoted	4.63	2.24	0.47	3.66	5.60	2.47	12.75
	Treatment Retained	8.43	4.86	1.72	4.37	12.50	2.63	18.19
	Treatment Promoted	9.71	3.62	1.21	6.92	12.49	2.82	15.09
	Total	5.89	3.41	0.44	5.02	6.76	2.47	18.19
CM-MF	Control Retained	18.60	6.07	1.32	15.83	21.35	9.24	30.02
	Control Promoted	20.82	7.44	1.55	17.60	24.04	9.97	34.26
	Treatment Retained	22.99	6.51	2.30	17.54	28.42	14.09	30.00
	Treatment Promoted	23.14	3.06	1.02	20.78	25.50	17.80	29.12
	Total	20.68	6.48	0.83	19.02	22.34	9.24	34.26
CM-CF	Control Retained	9.50	3.47	0.76	7.92	11.08	5.96	17.19
	Control Promoted	8.68	2.49	0.52	7.61	9.76	5.12	14.78
	Treatment Retained	9.50	3.37	1.19	6.68	12.31	5.66	15.20
	Treatment Promoted	12.91	5.38	1.79	8.77	17.04	6.43	25.31
	Total	9.70	3.67	.47	8.75	10.63	5.12	25.31

Note. TCM-MF = teacher/child-managed meaning-focused; TCM-CF = teacher/child-managed code-focused; CM-MF = child-managed meaning-focused; CM-CF = child-managed code-focused.

about 92, on average. Indeed, the only significant differences among the groups' fall scores was on the Passage Comprehension assessment, where the students in the retained group had significantly lower scores than did the matched-promoted group (mean difference = 13.8, $p < .05$). Scores on the reading and vocabulary assessments correlated moderately to highly with correlation coefficients (r), ranging from .10 (fall Picture Vocabulary and fall Passage Comprehension) to .84 (spring Letter-Word Identification and spring Passage Comprehension). Correlations are provided in Table 6. End of year Letter-Word Identification and Passage Comprehension were the outcome variables with Picture Vocabulary included as a descriptive covariate.

Based on Hierarchical Linear Model (HLM) results, on average, children who were retained demonstrated significantly weaker word reading and passage comprehension skill gains (i.e., residualized change) than did their matched-promoted peers. However, there were no significant differences in their vocabulary gains (see Table 7). The unconditional models revealed low interclass correlations (ICC) of .005 for Letter-Word Identification, .06 for Passage Comprehension, and .0006 for Picture Vocabulary. Fall score, group mean centered, and retention (1 = retained, 0 = matched-promoted peer), not centered, were added to each model at the level of the child. School percentage of FARL was added, grand mean centered, at the classroom level.

Table 6
Correlations for Letter-Word Identification and Picture Vocabulary W Score

	1	2	3	4	5	6
Fall Letter-Word Identification	–					
Fall Picture Vocabulary	.345**	–				
Fall Passage Comprehension	.470**	.100	–			
Spring Letter-Word Identification	.482**	.133	.559**	–		
Spring Picture Vocabulary	.249*	.715**	.107	.191	–	
Spring Passage Comprehension	.510**	.241*	.632**	.844**	.233*	–

*$p < .05$, **$p < .01$.

Table 7
HLM Results Comparing Children Who Were Retained With Matched-Promoted Peers on Spring Letter-Word Identification, Passage Comprehension and Picture Vocabulary W Scores

Fixed Effects	Letter-Word Identification Coefficient (SE)	Passage Comprehension Coefficient (SE)	Picture Vocabulary Coefficient (SE)
Intercept	440.39 (3.05)***	454.55 (2.60)***	477.52 (1.19)***
Fall Score	.61 (.20)**	.44 (.21)***	.69 (.09)***
Retained	−20.85 (4.10)***	−7.12 (3.57)*	1.14 (1.49)
School % FARL	−3.55 (2.24)	−2.86 (1.98)	−1.42 (.91)
Random Effects			
Classroom-Level Variance	27.43	56.51**	9.89*
Student-Level Variance	367.84	155.42	45.91

Note. FARL = free and reduced-price lunch. Variable retained is 1 and 0 for the matched-promoted peer;degrees of freedom are 33 for classroom-level variables and 89 for student-level variables.
$p < .05$. **$p \leq .01$. ***$p < .001$.

We then compared spring Letter-Word Identification and Picture Vocabulary outcomes for the four groups, TR, TP, CR, and CP by leaving the TP group as the fixed referent and adding dummy variables for the other groups (e.g., TR = 1, all others = 0; CR = 1, all others = 0) at the level of the child. Because the retained and matched-promoted groups differed in their fall Passage Comprehension, for this model, fall Passage Comprehension score was group mean centered. For all three models, school percentage of FARL was grand mean centered. We used the hypothesis testing function of HLM to test differences between the TR and CR groups. Results are provided in Table 8 and indicate that for Letter-Word Identification, the students who were retained in the TR and CR groups showed significantly weaker spring scores compared with students who were promoted in the TP and CP groups. Hypothesis testing revealed that the TR group had significantly lower scores than did the CR group, χ^2 (2) = 8.59, $p = .013$. However, there were no significant differences between the TP and CP groups. For Passage Comprehension and Picture Vocabulary, there were no significant differences in spring scores among the groups, controlling for fall score.

The research question informing Aim 1 was: Are children more or less likely to be retained if they participated in the ISI intervention compared with children in control classrooms? Our hypothesis was supported. When we examined differences in rates of retention for the ISI intervention and control group students who were retained using a chi-square analysis, results revealed that there was a significant difference between the groups, $\chi^2(1) = 7.42$, $p = .006$. Children were significantly

Table 8
HLM Results for Spring Outcomes Across Treatment-Retained, Control-Promoted, and Control-Retained Groups, Where Treatment-Promoted is the Fixed Reference Group

Fixed Effects	Letter-Word Identification Coefficient (SE)	Passage Comprehension Coefficient (SE)	Picture Vocabulary Coefficient (SE)
Intercept	438.09 (6.00)***	450.30 (4.42)***	474.54 (2.47)***
Fall Score		.43 (.12)**	
Treatment Retained	−23.18 (8.63)**	−3.63 (6.16)	1.48 (3.86)
Control Promoted	4.18 (7.00)	6.27 (5.35)	5.14 (3.11)
Control Retained	−17.60 (6.93)*	−2.34 (5.57)	4.38 (3.07)
School % FARL	−3.50 (2.24)	−2.81 (2.02)	−1.27 (.98)
Comparing TR and CR, χ^2 (2)	8.59*	2.38	2.37
Random Effects	Letter-Word Identification Coefficient (SE)	Passage Comprehension Coefficient (SE)	
Classroom-Level Variance	7.36	63.34**	
Student-Level Variance	422.87	153.17	

Note. FARL = free and reduced-price lunch; TR = treatment retained; CR = control retained; degrees of freedom are 33 for classroom-level variables and 89 for student-level variables.
* $p < .05$. ** $p \leq .01$. *** $p < .001$.

Table 9
Cross-Tabulation for Control Retained Versus Treatment Retained Groups

Study Condition	Retained	Promoted	Total
Control			
Count	42	367	409
Expected Count	31.6	377.4	409
Treatment			
Count	21	385	406
Expected Count	31.4	374.6	406
Total ISI Study			
Count	63	752	815

less likely to be retained if they participated in an ISI classroom than were children in control classrooms (see Table 9).

Our second aim was to examine whether there were differences in the literacy instruction provided to students who were retained compared with their matched-promoted peers. Again, we examined differences in small-group TCM-CF and CM-MF DFR minutes and amounts (min) of TCM-MF. Our hypothesis was that students who were retained would be less likely to receive the A2i recommended amounts and, hence, their TCM-CF and CM-MF DFR amount would be larger compared who students who were promoted. We also anticipated that students who were retained would spend less time in small-group TCM-MF than their peers who were promoted. We also hypothesized that this might vary between treatment and control conditions.

We used HLM to compare, across groups, TCM-CF and CM-MF DFR and total amount of TCM-MF instruction students received on average in the fall, winter, and spring. The ICC was

Table 10
HLM Model Comparing Distance From Recommendation Across Groups, for Teacher/Child-Managed Code-Focused and Child-Managed Meaning-Focused Instruction, Where Treatment Promoted Is the Fixed Reference Group

Fixed Effects	TCM-CF DFR	CM-MF DFR	TCM-MF Amount
Intercept	17.90 (.77)***	17.27 (1.50)***	12.50 (.93)***
Treatment Retained	1.04 (.96)	5.35 (1.79)**	−.14 (1.21)
Control Promoted	2.99 (.92)**	4.22 (1.82)*	−2.51 (1.47)∼
Control Retained	2.79 (.92)**	4.28 (1.81)*	−3.30 (1.20)**
School % FARL	.03 (.32)	−.77 (1.51)	.32 (.61)
Comparing TR and CR, χ^2 (2)	9.27**	10.54**	7.53*
Random Effects	TCM-CF DFR	CM-MF DFR	
Classroom-Level Variance	1.96**	9.05***	
Student-Level Variance	2.79	9.49	

Note. DFR = distance from recommendation in minutes; TCM-CF = teacher/child-managed code-focused; CM-MF = child-managed meaning-focused; TCM-MF = teacher/child-managed meaning-focused; FARL = free and reduced-price lunch; TR = treatment retained; CR = control retained. For TCM-MF, total amount is shown. Degrees of freedom are 27 for classroom level variables and 53 for student level variables.
* $p < .05$. ** $p \leq .01$. *** $p < .001$.

.40 for TCM-CF DFR, .43 for CM-MF DFR, and .21 for TCM-MF amount. In general, students who were promoted (TP and CP) tended to have smaller DFR amounts than did students who were retained (TR and CR) but this differed by type of instruction; therefore, our hypothesis was only partially supported (see Table 10 and Figure 1). Overall, for both TCM-CF and CM-MF DFR, students in the ISI intervention classrooms, whether retained or promoted, tended to have smaller DFR scores than did students in the control groups (CR and CP). In addition, DFR did not significantly differ for CR and CP. When we compared the TR and TP groups, they did not have significantly different TCM-CF DFR scores. However, TP students had significantly smaller CM-MF DFR scores than did the students in the TR group. That is, the TP students were more likely to receive the recommended amounts of CM-MF than were the TR students. Thus, variability within the ISI intervention implementation predicted whether or not a student might be retained.

When we compared the total amount of small-group TCM-MF instruction across groups using HLM, our hypothesis was, again, partially supported. Students who were retained generally spent less time in small-group TCM-MF instruction than did their promoted peers, but this varied by whether they were in the treatment or control group. Specifically, students in the TP group were more likely to spend greater amounts of time in small-group TCM-MF instruction than were students who were in the CR group. There was no difference in amounts of TCM-MF for students in either promoted group (CP and TP, $p = .09$). Comparing children in the TR and CR groups revealed that children in the TR group were observed to participate for significantly more time in TCM-MF instruction than were children in the CR group. Finally, there were no significant differences in amounts of TCM-MF students in the TR and TP groups received.

Our third aim was to identify other potential student characteristics, specifically self-regulation, that might predict retention. Our hypothesis was that students who began the year with weaker self-regulation or whose teachers perceived them to have weaker self-regulation would be more likely to be retained, regardless of whether they were in a treatment or control classroom. We used multivariate general linear models with fall HTKS and the social, behavioral (higher scores indicate greater behavior problems), and academic competence scales of the SSRS as the outcome and whether students were retained or not as the fixed factor. In a second analysis, we used spring HTKS

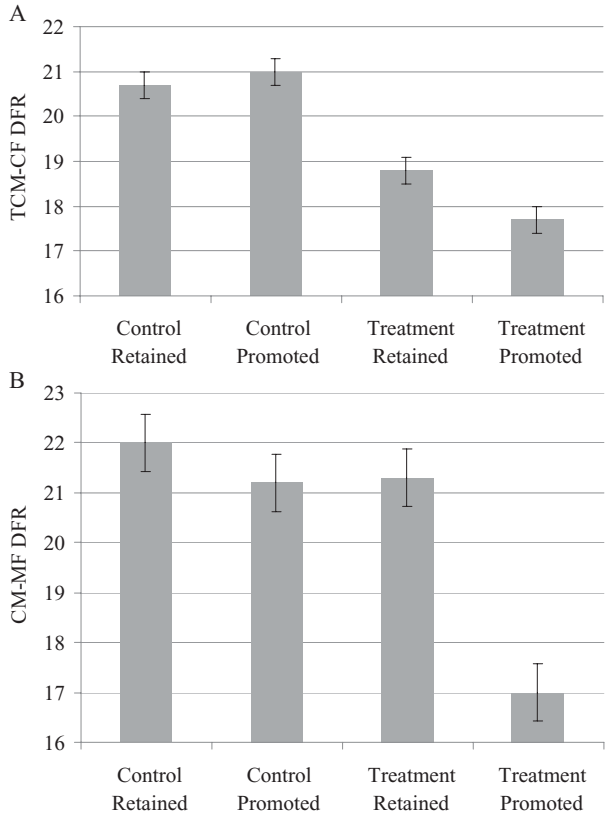

FIGURE 1. (A) Year 1 mean of teacher/child-managed code-focused distance from recommendation (TCM-CF DFR). Distance from recommendation is the absolute value of the difference between the amount (min) recommended by the A2i software and the observed amount. (B) Year 1 mean of child-managed meaning-focused distance from recommendation (CM-MF DFR). Distance from recommendation is the absolute value of the difference between the amount (min) recommended by the A2i software and the observed amount.

as the outcome and the four groups (TR, TP, CR, CP) as the fixed factor. Results from the first analysis revealed that students who were retained had significantly weaker self-regulation overall than their promoted peers, Wilks' Lambda $(4, 64) = .784$, $p = .003$. Post-hoc analyses revealed that fall HTKS scores were significantly lower for students who were ultimately retained. With regard to teachers' perceptions in the winter, students who were retained were judged to be less academically competent, $F(1, 67) = 14.51$, $p < .001$, and there was a trend that they were perceived to have weaker social skills, $F(1,67) = 3.69$, $p = .058$. There were no significant differences perceived with regard to behavior problems ($p = .60$; Table 1).

When we examined differences in spring HTKS for students who were retained or promoted, we found no significant differences between retained or promoted students, $t(70) = -1.43$, $p = .16$, or among groups (TR, TP, CR, CP), $F(3, 68) = .867$, $p = .462$. To double check that it was not teacher perception but directly assessed fall HTKS that predicted retention, we repeated the t test for fall HTKS. We found that students who were ultimately retained began the school year with weaker self-regulation compared with students who were promoted, $t(78) = -3.33$, $p = .001$. When we compared fall HTKS among groups, there were significant differences as well, $F(3, 76) = 4.59$,

$p = .005$, with post hoc analyses (Tukey HSD) revealing TR students to have significantly lower scores compared with TP and CP students ($M = 22.1$ vs. 31.07 and 31.14, respectively). However, CR students had scores ($M = 26.6$) that were not significantly lower than the other groups. Thus, it would seem that fall but not spring self-regulation was predictive of whether or not a student would be retained in first grade.

Discussion

The purpose of the current study was to examine whether providing reading instruction that was individualized based on students' reading and vocabulary scores and with evidence of efficacy might reduce rates of first-grade retention. It also aimed to examine the instruction children received, comparing children who were retained with matched-promoted peers, and to investigate whether students' self-regulation might predict retention. Our hypotheses were supported for the most part. First, we found that students receiving the ISI intervention were less likely to be retained than students who were in the control classrooms. In general, students in the control condition who were promoted, compared with those who were retained, were more likely to receive the amounts of small-group instruction that have been associated with stronger outcomes overall (Connor, 2011); that is, promoted students generally received more TCM-MF instruction than did students who were retained, and they were more likely to receive the A2i recommended amounts (even though they were in control classrooms) than their peers who were retained in first grade. Within the treatment group, the only instructional difference was in the amounts of CM-MF instruction. In general, A2i recommends that students with weak vocabulary and reading skills in the fall receive less CM-MF instruction in the fall, with generally increasing amounts in the spring (Connor et al., 2009). Promoted students were more likely to be provided with these increasing opportunities to practice reading independently and with peers. With regard to the third aim, we found that students' self-regulation at the beginning of the school year, but not at the end of the school year, predicted whether or not they would be retained. Considering previous research by Jimerson and Ferguson (2007), it is not surprising that self-regulation differences exist between retained and promoted students. However, why there were no differences in the spring has not been explained in previous research. It may be that it took the course of the school year for students who initially demonstrated weaker self-regulation to embody the skills necessary to self-regulate. In addition, students who were retained were perceived to be less academically competent than their peers who were promoted, even though their reading and vocabulary scores were matched at the beginning of the school year.

The ISI Intervention and Retention

Why did we find that students were less likely to be retained if they attended ISI intervention schools rather than control schools? There are a number of possible reasons, and they might all have contributed. First, it might have been that ISI teachers were making better instructional decisions for their students, which led to stronger outcomes; hence, there were fewer children left who were performing poorly enough to trigger a retention decision. We have seen in previous research that teachers are making evidence-based instructional decisions for at-risk or retained students (Picklo & Christenson, 2005). However, it may be that because those participating in the ISI intervention were provided with information on how to deliver instruction (small group vs. whole class) as well as the types of activities to promote specific literacy skills, which may have resulted in fewer retained students in the treatment condition. Indeed, examining the difference between the A2i-recommended amounts of instruction and the instruction actually delivered, DFR revealed that students who were in the treatment classrooms and were promoted, even though they began the school year with matched Letter-Word Identification and Picture Vocabulary scores, were more likely to receive the

A2i-recommended amounts of CM-MF (i.e., small DFR). Moreover, their DFRs were smaller than the DFRs for children in control classrooms, particularly the CR group. In addition, children in the treatment classrooms were more likely to spend greater amounts of time in TCM-MF instruction than were children in control classrooms. This suggests that the intervention guided instruction might have supported students' reading achievement, leaving fewer children with weaker reading skills and who were, therefore, retained.

Second, teachers in the ISI condition had easy access to assessment and progress information through the A2i software, which might have helped to inform their decisions. For example, the A2i software provided teachers with updated assessment information in the form of age equivalents, grade equivalents, standard scores, and progress monitoring graphs. The A2i graphs compared targeted growth with observed growth based on periodic assessments of students' reading and language skills. Although decisions regarding retention were made using district-administered assessments and followed specified protocols, the assessment information available to teachers through A2i provided similar information.

Next, although retained students and their matched-promoted peers started the school year with similar Letter-Word Identification and Picture Vocabulary scores, retained students finished the school year with significantly weaker Letter-Word Identification and Passage Comprehension scores. Based on our interviews with teachers, this was to be expected because they overwhelmingly reported that their students were retained in first grade because their reading skills were too weak to meet promotion requirements. What is intriguing is that there were no significant differences in retained versus promoted students' vocabulary scores by the end of the school year. In addition, retained students had generally weaker Passage Comprehension scores (not a matching variable) in the fall. Although Picture Vocabulary scores were matched, this did not translate into equivalent Passage Comprehension scores. Moreover, retained students achieved gains in Passage Comprehension that were similar to the matched-promoted students, although this still left them with generally weaker Passage Comprehension scores by spring. The relationship between deficits in reading skills and retention has been documented in previous research (Holmes &Matthews, 1984; Jimerson, 2001). Consider also that students in treatment classrooms who were promoted were more likely to receive A2i-recommended amounts of CM-MF instruction than were their peers who were retained. This might explain the differences noted in Passage Comprehension gains as well. Although conjecture, our findings point to the potential importance of students' comprehension skills as they enter first grade. It may be that simply assessing students' word reading (i.e., reading lists of words) is not an ideal screen for risk of reading difficulties (and retention). Rather, assessing basic reading comprehension (e.g., matching sentences to a picture or identifying the missing word of a sentence) or connected language might be more informative.

In addition to fall differences in passage comprehension, we also found significant differences in fall self-regulation between students who were retained and their matched-promoted peers. Of note, retained and promoted students did not differ in their spring self-regulation. Moreover, within the treatment group, students who were ultimately retained had significantly lower self-regulation skills than did their peers who were promoted, and the difference was large: almost 9 points on an assessment, with a range of scores from 0 to 38 (where 40 is a perfect score) and a standard deviation of 11.7. Ponitz et al. (2009) found that fall self-regulation predicted spring academic performance, when controlling for background factors and fall academic performance in kindergarten students. Therefore, it is not surprising that our findings indicate students who would ultimately be retained exhibited weaker self-regulation in the fall. The relationship between self-regulation and retention has been found when comparing students after retention has occurred (Jimerson & Ferguson, 2007). However, the current findings suggest that assessing self-regulation before students are retained may be helpful in identifying those students who may be at risk. It may be that students with stronger

self-regulation are able to more quickly master early literacy skills. It may also be that in the primary classroom, when centers are a common means of organizing the classroom, those with weaker self-regulation are less able to adjust to new tasks and work independently.

Study Limitations

There are several limitations in the current study that should be considered when interpreting our results. First, the information explaining reasons for each retention decision was not available. Therefore, we do not know exactly why each student was retained, although teacher interviews suggest that the principal reasons included reading deficits. Next, a word reading measure was utilized, whereas a direct measure of decoding skill was not. Therefore, specific deficits in word decoding may not have been detected, even though students who were retained performed similarly to matched-promoted students on the word reading measure. Possible unidentified deficits in word decoding may have contributed to students' difficulty in completing the comprehension task and may explain a need for teachers to have provided more code-focused instruction. In addition, the sample size was small because only a small percentage of the original sample was retained. However, this means that statistical power was an issue, and there may have been group differences that we were unable to detect. Likewise, assignment of treatment condition was at the school level. Although schools were matched, differences in how schools interpret student data and make retention decisions above and beyond the district guidelines for retention may differ from school to school and may have impacted our results. When we tried to address Aim 1 using HLM with a Bernoulli distribution, where retained = 1 and promoted = 0, neither the two- nor three-level unconditional model would converge. Thus, unfortunately, our sample size and small ICCs precluded HLM analyses for Aim 1. For the same reasons, analyses were conducted at the student level for Aim 3. We did address Aim 2 using two-level models. Preliminary analyses suggested that three-level models yielded similar results. However, given the small sample size and concerns about power, we decided to use the more parsimonious two-level models. Even with these limitations, we felt that the opportunity to examine multiple factors that might contribute to decisions of retention within the context of a randomized control trial was important enough to warrant this study.

Implications

The findings from this study indicate that differentiating literacy instruction in first-grade classrooms to meet the needs of individual students whose reading and vocabulary skills differ may result in fewer students retained at the end of first grade. Many studies have highlighted the negative effects of grade retention on student outcomes, both academic and non-academic, as well as its substantial cost (Reynolds, Temple, Robertson, & Mann, 2002), but very few have addressed the types of instruction students who were retained were receiving. Examining the achievement of students who were retained and received the ISI intervention extends this research and demonstrates the potential value of continuing to explore the association between classroom instruction and grade retention. With the adoption of Response to Intervention (Gersten et al., 2008), teachers and administrators are increasingly called on to tailor instruction based on children's reading and language skills. One additional benefit may be that fewer students will be retained. Moreover, students' self-regulation and passage comprehension skills in the fall predicted which students would be retained at the end of first grade. Neither construct is regularly part of universal screening or progress monitoring. Thus, including such skills in assessment and progress monitoring of students at risk for reading difficulties and retention might help identify such students for early intervention services. Future research on the topic of retention might consider monitoring the types of instruction students are receiving in the classroom before and after retention occurs. For example, TCM-MF

instruction appears to predict students' reading comprehension growth (Connor et al., 2009), and children in treatment classrooms were more likely to receive this type of instruction. Furthermore, it may be informative to follow students through elementary school and beyond to determine whether the instructional experiences a student encounters before and after retention occurs has made a difference in future academic achievement. Ideally, we would be able to predict which instructional practices result in the fewest number of students being retained and then provide the instruction. By reducing rates of retention, which is a costly and largely ineffective intervention, we may be able to improve the ongoing reading skills of students who are at serious risk for reading difficulties. This is important because students who are proficient readers are more likely to be successful in school, less likely to be referred for special education, and generally have stronger outcomes than are students with weak reading skills.

References

Abbott, M., Willis, H., Greenwood, C. R., Kamps, D., Heitzma-Powell, L., & Selig, J. (2010). The combined effects of grade retention and targeted small-group intervention of students' literacy outcomes. Reading and Writing Quarterly, 26(1), 2–25.

Beebe-Frankenberger, M., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Sorting second-grade students: Differentiating those retained from those promoted. Journal of Educational Psychology, 96(2), 204–215.

Bos, C., Mather, N., Narr, R. F., & Babur, N. (1999). Interactive, collaborative professional development in early literacy instruction: Supporting the balancing act. Learning Disabilities Research and Practice, 14(4), 227–238.

Connor, C. M. (2011). Child by Instruction interactions: Language and literacy connections. In S. B. Neuman & D. K. Dickinson (Eds.), Handbook on early literacy (3rd ed., pp. 256–275). New York: Guilford.

Connor, C. M., Morrison, F. J., Fishman, B., Giuliani, S., Luck, M., Underwood, P.,et al. (2011). Testing the impact of child characrteristics X instruction interactions on third graders' reading comprehension by differentiating literacy instruction. Reading Research Quarterly, 46(3), 189–221.

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

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

Connor, C. M., Morrison, F. J., & Katch, L. E. (2004). Beyond the reading wars: Exploring the effect of child-instruction interactions on growth in early reading. Scientific Studies of Reading, 8(4), 305–336.

Connor, C. M., Morrison, F. J., & Petrella, J. N. (2004). Effective reading comprehension instruction: Examining Child by Instruction interactions. Journal of Educational Psychology, 96(4), 682–698.

Connor, C. M., Morrison, F. J., Schatschneider, C., Toste, J., Lundblom, E. G., Crowe, E., et al. (2011). Effective classroom instruction: Implications of child characteristic by instruction interactions on first graders' word reading achievement. Journal for Research on Educational Effectiveness,4(3), 173–207.

Connor, C. M., Morrison, F. J., & Underwood, P. (2007). A second chance in second grade: The independent and cumulative impact of first and second grade reading instruction and students' letter-word reading skill growth. Scientific Studies of Reading, 11(3), 199–233.

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

Connor, C. M., Ponitz, C. C., Phillips, B. M., Travis, Q. M., Glasney, S., & Morrison F. J. (2010). First graders' literacy and self-regulation gains: The effect of individualized student instruction. Journal of School Psychology, 48, 433–455.

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

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, et al. (2008). Assisting students struggling with reading: Response to Intervention and multi-tier intervention for reading in the primary grades. A practice guide. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Gleason, K. A., Kwok, O., & Hughes, J. N. (2007). The short-term effect of grade retention on peer relations and academic performance of at-risk first graders. The Elementary School Journal, 107(4), 327–340.

Gresham, F. M., & Elliot, S. N. (1990). Social Skills Rating System. Bloomington, MN: Pearson Assessments.

Griffith, C. A., Lloyd, J. W., Lane, K. L., & Tankersley, M. (2010). Grade retention of students during grades K-8 predicts reading achievement and progress during secondary schooling. Reading and Writing Quarterly, 26(1), 51–66.

Holmes, C. T., & Matthews, K. M. (1984). The effects of nonpromotion on elementary and junior high school pupils: A meta-analysis. Review of Educational Research, 54(2), 225–236.

Hong, G., & Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. Educational Evaluation and Policy Analysis, 27(3), 205–224.

Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. School Psychology Review, 30(3), 420 437.

Jimerson, S., Carlson, E., Rotert, M., Egeland, B., & Sroufe, L. A. (1997). A prospective, longitudinal study of the correlates and consequences of early grade retention. Journal of School Psychology, 35(1), 3–25.

Jimerson, S.R., & Ferguson, P. (2007). A longitudinal study of grade retention: Academic and behavioral outcomes of retained students through adolescence. School Psychology Quarterly, 22(3), 314–339.

Jimerson, S. R., Pletcher, S. M. W., Graydon, K., Schnurr, B. L., Nickerson, A. B., & Karweit, N. (1991). Repeating a grade: Time to grow or denial of opportunity? Baltimore: Center for Research on Effective Schooling for Disadvantaged Students.

Karweit, N. (1999). Grade retention: Prevalence, timing, and effects. Baltimore: Johns Hopkins University, Center for Research on Students Placed at Risk.

Kundert, D. K. (2006). Beyond grade retention and social promotion: Promoting the social and academic competence of students. Psychology in the Schools, 43(1), 85–97.

Mantzicopoulos, P., & Morrison, D. (1992). Kindergarten retention: Academic and behavioral outcomes through the end of second grade. American Educational Research Journal, 29(1), 182–198.

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

McCoy, A., & Reynolds, A. (1999). Grade retention and school performance: An extended investigation. Journal of School Psychology, 37(3), 273–298.

National Institute of Child Health and Human Development. (2000). National Reading Panel report: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

National Reading Panel. (2000). Teaching children to read: An evidence-based assessment of the scientific research literature and its implications for reading instruction (NIH Publication No. 00-4769). Washington, DC: U.S. Department of Health and Human Development.

Noldus Information Technology. (2010). The Observer Video-Pro XT 10.0. Retrieved December, 2010, from http://www.noldus.com/human-behavior-research/products/the-observer-xt

O'Connor, R. E., Bacian, K., Beebe-Fcankenberger, M., & Linklater, D. L. (2010). Responsiveness of students with language difficulties to early intervention in reading. The Journal of Special Education, 43(4), 220–235.

Peterson, S. E., DeGracie, J. S., & Ayabe, C. R. (1987). A longitudinal study of the effects of retention/promotion on academic achievement. American Educational Research Journal, 24(1), 107–118.

Peterson, S. E., & Hughes, J. N. (2011). The differences between retained and promoted children in educational services received. Psychology in the Schools, 48(2), 156–165.

Picklo, D. M., & Christenson, S. L. (2005). Alternatives to retention and social promotion. Remedial and Special Education, 26(5), 258–268.

Ponitz, C. C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. Early Childhood Research Quarterly, 23, 141–158.

Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral regulation and its relation to kindergarten outcomes. Developmental Psychology, 45, 605–619.

Reutzel, D. R., Smith, J. A., & Fawson, P. C. (2005). An evaluation of two approaches for teaching reading comprehension strategies in the primary years using science texts. Early Childhood Research Quarterly, 20, 276–305.

Reynolds, A. J., Magnuson, K. A., & Ou, S.R. (2010). Preschool-to-third grade programs and practices: A review of research. Children and Youth Services Review, 32, 1121–1131.

Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2002). Age 21 cost-benefit analysis of the Title I Chicago child-parent centers. Educational Evaluation and Policy Analysis, 24(4), 267–303.

Roderick, M., & Nagaoka, J. (2005). Retention under Chicago's high-stakes testing program: Helpful, harmful, or harmless? Educational. Evaluation and Policy Analysis, 27(4), 309–340.

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

Schnurr, B. L., Kundert, D. K., & Nickerson, A. B. (2009a). Grade retention: Current decision making practices and involvement of school psychologists. Psychology in the Schools, 46, 410–419.

Silberglitt, B., Appleton, J., Burns, M., & Jimerson, S. (2006). Examining the effects of grade retention on student reading performance: A longitudinal study. Journal of School Psychology, 44, 255–270.

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

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

Stone, S., & Engel, M. (2007). Same old, same old? Students' experiences of grade retention under Chicago's Ending Social Promotion Policy. American Journal of Education, 113, 605–634.

Taylor, B. M., Pearson, D. P., Clark, K., & Walpole, S. (2000). Effective schools and accomplished teachers: lessons about primary-grade reading instruction in low-income schools. Elementary School Journal, 101(2), 121–165.

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

Torgesen, J. K. (2002). The prevention of reading difficulties. Journal of School Psychology, 40(1), 7–26.

Westbury, M. (1994). The effect of elementary grade retention on subsequent school achievement and ability. Canadian Journal of Education, 19(3), 241–250.

Wharton-McDonald, R., Pressley, M., & Hampston, J. M. (1998). Literacy instruction in nine first-grade classrooms: Teacher characteristics and student achievement. Elementary School Journal, 99(2), 101–128.

Wu, W., West, S. G., & Hughes, J. N. (2008). Effect of retention in first grade children's achievement trajectories over 4 years: A piecewise growth analysis using propensity score matching. Journal of Educational Psychology, 100(4), 727–740.

Wu, W., West, S. G., & Hughes, J. N. (2010). Effect of grade retention in first grade on psychosocial outcomes. Journal of Educational Psychology, 102(1), 135–152.