Hammill Institute on Disabilities

The Journal of Special Education
2022, Vol. 56(1) 14–24
© Hammill Institute on Disabilities 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/00224669211014168
journalofspecialeducation.sagepub.com
SAGE

Comparing Technology-Based Reading Intervention Programs in Rural Settings

Brit'ny Stein, PhD[1], Benjamin G. Solomon, PhD[2], Chase Kitterman, MS[3], Debbie Enos, MLIS[1], Elizabeth Banks, MS[3], and Sierra Villanueva, MS[3]

Abstract

An ever-growing call for the use of evidence-based practice has come up against the logistical hurdles of a lack of resources and expertise, particularly in rural schools that work with historically underserved students. Although integrated learning systems (ILSs)—stable and likely requiring fewer resources than personnel—do not offer a complete solution to this problem, they may serve as a useful resource, particularly for milder literacy deficits. And yet, there is a surprising lack of empirical research on their effectiveness, particularly for contemporary programs. This study examines the effectiveness and efficiency of two popular ILSs, *Lexia* and *iStation*, both of which use a blended model of computer and traditionally delivered instruction, and compares them against business-as-usual (BAU) conditions across a variety of outcomes. Results suggest both programs resulted in meaningful growth across an academic year of implementation, although generally no more so than that observed in the BAU condition. However, *Lexia* yielded the highest level of instructional efficiency. That is, despite comparable growth across conditions, *Lexia* required less staff time to implement per student participant.

Keywords

integrated learning systems, computer-adaptive intervention, rural education, reading intervention

Teaching children to read with fluency and comprehension is a fundamental goal of the elementary school curriculum (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). For children exhibiting slow growth in their early reading development, evidence-based intervention has been found effective in resolving instructional deficits (Coyne et al., 2019; Vellutino et al., 2006). Unfortunately, there has been a national struggle to implement best practice, which is likely caused by a lack of resources and appropriate faculty/staff training (Benner et al., 2011; Piasta et al., 2020). Integrated learning systems (ILSs), as a primary or supplementary support, may compensate for these vulnerabilities by allowing for a higher teacher to student ratio for intervention and substantially lower training needs. However, the evidence base for ILSs is concerningly thin, despite their near unrestrained use within schools. To address this gap, this study investigated the effectiveness of two ILSs to resolve reading deficits among students at-risk.

Integrated Learning Systems

An ILS is defined presently as instruction delivered to students via a personal computer or tablet, with a 1:1 computer to student pairing, which is *adaptive* in the leveling and/or repetition of lessons and activities based on ongoing student performance (Lee & Park, 2007; Putman, 2017). The promise of ILSs has been hailed for decades. For example, Torgesen and Barker (1995) noted that a computer's unique ability to rapidly administer consistent learning trials fit well within the developing understanding of the instructional needs of students with learning disabilities, and they identified emerging programming to fit this need. An earlier study by Torgesen and colleagues reported positive effects for varied approaches to computerized instruction for teaching sight words (Torgesen et al., 1988). Later developers touted individualized, intensive instruction that is not constrained to one-to-one or small-group administration (Macaruso & Rodman, 2011; Putman, 2017). School-based ILSs often align content with the foundations of reading found through research to be critical to overall proficiency (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010; National Reading Panel, 2000).

[1]Osage County Interlocal Cooperative, Hominy, USA
[2]University at Albany, State University of New York, USA
[3]Oklahoma State University, Stillwater, USA

Corresponding Author:
Benjamin G. Solomon, University at Albany, State University of New York, Albany, NY 12222, USA.
E-mail: bgsolomon@albany.edu

Traditional reading interventions typically require implementation in a small group or in a one-on-one faculty to student setting, limiting the number of students who can receive intensive instruction and the overall efficiency and fidelity of administration (Scanlon et al., 2016). Implementation of such interventions must be feasible for trained interventionists to conduct and accessible to the majority of the students in need for there to be a meaningful distributed effect. Due to the persistently high level of reading deficits observed across students educated within the national public school system (National Center for Educational Statistics, 2019), it is economically essential that such interventions not be solely conducted by highly trained specialists, particularly in underresourced areas, where human capital may be scarce and there may be a high turnover of staff. Current literature suggests there is a paucity of research examining the utilization of available technology as a resource to deliver intervention so as to compensate for these logistical challenges.

ILSs Nested Within a Tiered Framework. ILSs have been conceptualized as a primary or supplemental means of delivering instruction across levels of student need (Cheung & Slavin, 2013; McDermott & Gormley, 2015). Commonly, within the nomenclature of response to intervention (RTI), the general educational environment—what all students receive—is referred to as Tier 1 instruction, whereas intensified instruction that a minority of students at-risk receive in addition to Tier 1 is considered Tier 2 (Fuchs et al., 2003). Presently, we describe ILSs as a mechanism for the delivery of Tier 2 supports (Torgesen & Barker, 1995). Importantly, one of the fundamental assumptions of RTI is that the intervention being applied has a preponderance of evidence attesting to its effectiveness (Fuchs et al., 2003; Individuals with Disabilities Education Improvement Act, 2004). The rapid adoption of ILSs within an RTI framework creates a demand for such empirical research, and we distinguish this use of ILSs from prior research, which positioned it as a supplement to teacher-driven in-person instruction (e.g., Torgesen et al., 2010).

Instruction in Rural Settings

A particular appeal of ILSs is that they may offer cost-efficient, standardized, and standards-aligned instruction in geographical areas where such support is otherwise particulary difficult to provide. Rural educational systems often must work with tight budgetary restrictions that limit hiring. They must be responsive to the needs of students whose families often come from poverty, such as the need for high-quality evidence-based reading instruction and intervention. A given rural system often must service a geographically wide community, making it difficult for specialists to access students. ILSs offer a partial solution to these challenges. Such programs, although expensive when coupled with the necessary hardware (e.g., computer; hotspot), are cost-efficient per student relative to the resources needed to purchase traditional intervention curricula/materials and train and support dedicated intervention staff, who also may turnover at a high rate. Such programs also may be accessed from the student's home, provided basic hardware and connectivity needs are met.

It also is often the case that students who identify as American Indian—a historically underserved demographic that constituted a large portion of our present sample—live within rural communities. Prior research has identified American Indian students as underperforming, on average, in basic academic domains relative to majority/privileged groups in the United States, as well as having fewer technological resources at home to access information (Freeman & Fox, 2005). Additional research, although sparse, suggests this underperformance can be explained primarily by higher mean levels of poverty within this group, and in rural communities generally (Hibel et al., 2008).

Lexia and iStation. Two examples of contemporary ILSs are *Lexia* (Lexia Learning, 2019) and *iStation* (iStation, 2019). Both programs are cloud-based, individualize literacy instruction based on initial and ongoing student performance, and align activities around the five pillars of reading and the Common Core state standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010; National Reading Panel, 2000). Furthermore, both programs weave in supplemental scripted teacher-guided lessons for students non-responsive to specific modules, referred to as a blended model of delivery (Schechter et al., 2015). Recommended time on intervention ranges from 20 to 90 min per week, differing by vendor and the student's initial and ongoing performance (iStation, 2019; Lexia Learning, 2019).

There is promising, but limited, evidence for the effectiveness of either program. In a set of initial studies, Macaruso and Rodman (2011) found that use of *Lexia* approximately 15 min a day, 3 times a week from beginning to end of the school year, resulted in significant average improvement on standardized reading assessments compared with a no-treatment control group for a small heterogenous group of preschool and kindergarten students. The control and experimental groups included students in Tier 1 class-wide instruction, with students in the experimental group receiving *Lexia* as a supplement. The effect was pronounced for initially low-performing students. A similar study found comparable results for students in first and second grades (Schechter et al., 2015). Like Macaruso and Rodman (2011), *Lexia* was offered as a supplement, with the control group receiving business-as-usual (BAU) Tier 1 instruction. A third study, O'Callaghan et al. (2016), examined the effects of *Lexia* on a larger group of Irish students identified as at-risk of reading, with intervention applied daily over the course of 2 months. Again, the control group received only Tier 1 instruction. Again, results favored the use of *Lexia*, with effects in the small range. Of these three studies, only the second mentioned the use of a blended model.

Less research appears to exist regarding *iStation*. One quasi-experimental study, Putman (2017), compared general education kindergarten students receiving the program in addition to Tier 1 instruction against a matched contrast who received only Tier 1 instruction. Students participating in *iStation* outperformed contrast students on a standardized measure of reading. Blended learning as a response to individual performance on *iStation* was not mentioned. Taken together, this body of research suggests further research is warranted. Limited research exists on the effectiveness of ILSs for students at-risk and students in rural areas, compared with traditional in-person Tier 2 intervention, and with the programs' blended model implemented.

Instructional Time

Skinner et al. (1996), Skinner (2008), and Poncy et al. (2015) discussed how studies of comparative intervention effectiveness can be misleading unless they account for implementation time (IT) expended in intervention. This is a pertinent consideration in light of the authors' observation that the primary instructional objective of intervention is to accelerate learning rates beyond that which occurs as a result of core instruction. Simply, two interventions may be of comparable effectiveness, but if one takes longer to implement, then parsimony would dictate the simpler intervention of shorter duration is superior, which allows for more time in core instruction. This lens on intervention superiority—examining intervention effects in light of the required time needed to implement them—is applicable to the comparative study of ILSs, as they require varying amounts of time to implement with fidelity and may use a blended delivery model. This may result in meaningful differences in the minimum IT required to implement the interventions with fidelity, even if outcome effects are similar, and this should be taken into account so as to maximize the validity of experimental results.

Purpose of the Current Study

Current research suggests that commercialized ILSs are underresearched. This is problematic given their widespread use. Furthermore, little research to our knowledge has directly compared ILSs within an experimental framework, and none within the lens of IT, that is, learning rates. In addition, the current study also took place in a highly rural location. Therefore, the purpose of the current study was to assess the effectiveness and efficiency of two ILSs, *iStation* and *Lexia*, when compared against traditional BAU Tier 2 supports. Our research questions were:

- **Research Question 1 (RQ1):** Does the implementation of *Lexia* and/or *iStation*—including the blended model supplements—for students identified as at-risk for reading deficits result in gains in literacy beyond that observed in a BAU condition?
- **Research Question 2 (RQ2):** Which of the two programs result in the greatest gains in literacy when IT is considered (i.e., efficiency)?

Given the lack of research in this area, we made no directional hypotheses.

Method

Design

This study employed two concurrent 2 × 2 experimental factorial designs (Shadish et al., 2001) where the first independent variable was condition assignment and the second was time of data collection, that being either the pretest or posttest. One of the ILSs was randomly assigned to each school, with one school hosting *iStation* (Study 1) and the other *Lexia* (Study 2). Within those schools, eligible students were individually randomly assigned to either a BAU condition or the assigned intervention condition for that school. Pretest assessments took place in September of 2018, with both intervention conditions commencing immediately thereafter and continuing until posttest, which was in April 2019. An a priori power analysis was conducted to guide sample size based on an anticipated effect of .25, power of .80, and an anticipated .50 correlation between pre- and post-test data. The resulting anticipated sample size was 34 students within each school.

Participants and Setting

The two participating sites were both mid-western rural public elementary schools in neighboring districts, the first serving Grades 1 through 5 and the second serving Grades 1 through 4. The participating schools had populations between 200 and 300 students and were indexed as low performing by the state. The demographics of students in the first school, where the experimental group received *iStation*, was 54% American Indian, 2% Hispanic, 40% Caucasian, and 4% other/multiracial. Fifty-two percent of students were male, and 69% of students received free or reduced-price lunch (FRL). Chi-square tests yielded no significant difference in FRL rate, $\chi^2 = 1.78$, $p = .18$, or gender, $\chi^2 = .083$, $p = .77$, between the experimental and control groups.

Participants in the second school, where students in the experimental condition received *Lexia*, had a similar composition. Forty-eight percent were male, 25% were American Indian, 2% African American, 69% Caucasian, and 4% other/multiethnic. Sixty-one percent received FRL. No significance difference was observed between groups on gender, $\chi^2 = 0.00$, $p = 1.00$, but a significant difference was observed for FRL, $\chi^2 = 4.76$, $p = .03$. Specifically, there were 11 students who received FRL in the BAU and 17 in the *Lexia* condition. Last, student characteristics were similar across the overall samples comprising the two independent studies: FRL rate, $\chi^2 = .55$, $p = .46$, gender, $\chi^2 = .04$, $p = .83$.

In each school, 24 students were randomly assigned to the school's respective experimental condition (i.e., *Lexia* or *iStation*) and 24 to the BAU condition. Students eligible for study inclusion were initially identified as at-risk via the school's respective fall literacy screening. One school used DIBELS (University of Oregon, 2018) and the other STAR (Renaissance Learning, 2015) to conduct such screenings. In the first school, 17, 13, 12, and 5 students were found eligible from first, second, third, and fourth grade, respectively, and in the second, 20, 11, 10, and 7 students were identified from these associated grades. Researchers independently verified that each student was at risk when collecting pretest data by confirming their performance was below the 25th percentile across pretest measures.

Dependent Variables

Woodcock–Johnson IV. Included students were administered select subtests of the Woodcock–Johnson IV Tests of Achievement (WJ IV; Schrank, McGrew, Mather, & Woodcock, 2014). The following clusters were used as outcomes: *Broad Reading, Basic Reading*, and *Reading Fluency*. The *Broad Reading* cluster includes the *Letter-Word Identification, Passage Comprehension*, and *Sentence Reading Fluency* subtests. The *Basic Reading* cluster includes the *Letter-Word Identification* and *Word Attack* subtests. Finally, the *Reading Fluency* cluster includes the *Oral Reading* and Sentence *Reading Fluency* subtests. The WJ is a long-standing and popular broadband measure of academic achievement for students across the age span. The WJ has strong reliability, with reported internal consistency and test–retest coefficients above .80. The WJ also has demonstrated validity for the purposes of indexing student proficiency across curriculums. Norm-referenced scores are based on a large sample of students stratified across numerous demographic variables (Schrank et al., 2014). Three parallel versions of the WJ (A, B, and C) are available. In the current study, all versions were used, randomly selected for a given administration.

Fastbridge curriculum-based measurement. The Fastbridge (FastBridge Learning, LLC, 2015) family of curriculum-based measurement (CBMs) includes an array of traditional CBMs and computer-adaptive tests. These tests are reported to be reliable and valid for the purposes of screening and progress-monitoring (National Center for Intensive Intervention, 2019). The *readingCBM* test, a traditional paper-based oral reading fluency measure, and *COMPefficiency*, a computer-adaptive measure of reading comprehension, were administered. ReadingCBM has been found valid when used with a large minority demographic in our sample, American Indian youth (Pearce & Gayle, 2009). *readingCBM* is scored as orally read words identified correctly per minute, which is a well-researched index of student reading ability. *COMPeffeciancy* is graded as a percentage accuracy score. For the current study, *readingCBM* scores were converted to *z*-scores within grade and then pooled, as students in different grades likely began intervention at different performance levels and grew at different rates. Difference scores from pretest to posttest were examined.

On-task behavior. To assess whether observed differences among conditions could be attributed to the level of attention to instruction, class-wide direct observation of engagement was conducted using an adapted version of the *Behavioral Observation of Students in Schools* (Shapiro, 2011). This instrument is used to engage in a simultaneous recording of on-task and off-task behavior, where off-task behavior is further delineated into verbal, motor, and passive elements (see Shapiro, 2011, for operational definitions). Students within a given class were individually observed in a randomly sequenced rotating fashion using a 15-s momentary time-sampling schedule over the course of a 30-min observation. As outlined by Shapiro (2011), on-task behavior was recorded using momentary time-sampling, and the off-task behaviors were coded using partial interval recording. One observation was completed each month for each condition by an independent observer.

Independent Variable

iStation. iStation (2019) is a suite of interactive assessment and intervention modules that individual students engage using either a tablet or computer, and which is designed for pre-kindergarten through eighth grade. Each student worked one-on-one with a tablet and used headphones so as not to distract peers. Each student had an individual login and password to access their personalized *iStation* program. The teacher provided occasional assistance when students engaged in the primary program, and she redirected students if they appeared off-task.

iStation initially begins with a broadband preassessment of reader skills. An adaptive course of cycled instruction is then administered, moving across foundations and interactive activities both within and across lessons. This content is explicitly designed around the five pillars of reading: phonemic awareness, alphabetic knowledge, fluency, vocabulary, and reading comprehension, also in consideration of the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Formative assessment is embedded in day-to-day instruction. If, after repeated trials, a student does not reach threshold on a given skill, *iStation* recommends and provides print-out lessons for teachers to run individually with students. All recommendations for these supplementary lessons were followed in the current study. *IStation*'s assessment component, the *Indicators of Progress*, identifies baseline starting points for instruction, tracks student data, and provides overviews of student growth.

There was one primary interventionist who oversaw the grade-level groups of students who received *iStation*. This interventionist, who served as support staff at the school, as well as members of the research team, went through professional development conducted by an *iStation* representative on-site. This included sharing and review of an implementation checklist to guide the use of the program. Students engaged in *iStation* up to 90 min a week, based on recommendations from the vendor, distributed across the school week (see Note 1). Students did not make up missed sessions due to student absence, scheduled holidays, or school breaks. The implementing teacher then delivered supplemental lessons to nonresponsive students, as identified by the program, in the afternoon, which took approximately eight additional hours per week to implement, or 25–30 min per small group. Approximately 10–12 students from the intervention group required supplemental lessons each day.

Lexia. *Lexia* shares several structural components with *iStation*. This includes a preassessment that places students at an appropriate level, formative adaptive testing, and extension exercises for nonresponsive students. *Lexia* organizes instruction around six thematic strands: phonological awareness, phonics, structural analysis, fluency, vocabulary, and comprehension, which increase in difficulty across 18 leveled sets of activities (Macaruso et al., 2019; Schechter et al., 2015). *Lexia* offers recommendations for extended time on the program, follow-up activities, and generates class and student progress reports.

Two groups of students (first/second and third/fourth) engaged in *Lexia* in their computer lab for a total of 60 to 80 min of school time across the 5 days of the school's week. As with *iStation*, students did not make up missed sessions. Like *iStation*, students engaged *Lexia* one student to one tablet, wearing headphones, and using unique login and passwords to sign into the program. Supplemental recommend lessons were completed within the designated time reserved for intervention, which was a 45-min instructional block. A reading specialist at the school was trained by a *Lexia* representative and ran all groups, with support from the researchers, who also participated in the training. Like with *iStation*, an implementation plan was shared by the vendor, which anchored the training. The participating teacher reported that approximately half of the group required supplemental lessons on average each day.

Business-as-usual. The BAU condition consisted of the schools' typical pull-out small group remedial services for students flagged as at-risk in the fall. This included a mix of sight word instruction, explicit phonics, and reading fluency training prepared in consultation with researchers and which were not drawn from any particular curriculum. These activities were structurally similar across the two participating schools. Chosen activities were based on skill deficits as identified by the school's CBM assessment. The interventionists had one or multiple assistants, including support from *Americorps* volunteers. The primary interventionists for Study 1 were the classroom teachers. The primary interventionist for Study 2 led both *Lexia* and the BAU. Time spent in BAU ranged from 20 to 40 min per day and occurred every day of the school week.

Fidelity of administration. Both *iStation* and *Lexia* passively calculate time spent on intervention per week. For both *iStation* and *Lexia*, 100% of participating students met the criteria for fidelity of usage each week. Fidelity of supplemental activities was reviewed via analysis of permanent products. This analysis suggested 100% of supplemental activities across *iStation* and *Lexia* were completed as intended. As reported above, the interventionists and members of the research team who provided on-site support also reviewed implementation checklists for both programs. These were used to guide daily implementation, and all participating staff/researchers went through half to full day in-person trainings with the vendors.

Analysis. We first compared how each experimental group compared with their matched in-school BAU group across time and condition. Therefore, three sets of effects were reviewed: the main effects for time and condition and the interaction. Because each outcome variable was restricted to certain grade levels, and therefore could not be measured for every student involved, analysis of variance (ANOVA) was chosen over multivariate analysis of variance (MANOVA) due to the issues with listwise deletion this would introduce.

Second, we examined time spent in intervention so as to better understand the relative efficiency of each of the intervention conditions (Poncy et al., 2015; Skinner et al., 1996). In doing so, we considered (a) the overall school time allocated per intervention session, (b) this allocated time

Table 1. Descriptives Across Outcome Metrics.

	Pre-test			Post-test			
Outcome	*n*	*M* (*SD*)	*Skew*	*n*	*M* (*SD*)	*Skew*	Average difference
Control (School 1)							
WJ Broad Reading	23	77.88 (13.51)	−1.10	23	82.23 (13.50)	−1.82	4.35
WJ Basic Reading	24	83.58 (14.70)	−1.23	24	87.65 (14.38)	−1.68	4.07
WJ Reading Fluency	19	78.95 (11.23)	−0.52	22	83.14 (14.03)	−1.81	4.19
R-CBM	24	35.92 (34.76)	0.87	24	55.91 (40.03)	0.53	19.99
COMPeff	16	0.60 (0.17)	0.76	15	0.66 (0.18)	−0.22	0.06
Control (School 2)							
WJ Broad Reading	24	76.96 (11.60)	−0.06	23	79.91 (12.76)	0.10	2.95
WJ Basic Reading	24	85.92 (13.54)	−0.99	23	87.87 (15.36)	−0.68	1.95
WJ Reading Fluency	20	74.10 (12.17)	0.30	19	80.79 (13.78)	0.12	6.69
R-CBM	24	35.13 (35.98)	0.96	23	73.48 (39.10)	0.29	38.35
COMPeff	14	0.63 (0.15)	0.28	14	0.64 (0.18)	0.68	0.01
iStation (School 1)							
WJ Broad Reading	24	75.17 (8.80)	0.29	22	81.00 (9.76)	0.07	4.83
WJ Basic Reading	24	83.63 (7.45)	0.11	22	86.36 (10.07)	0.16	2.73
WJ Reading Fluency	22	72.86 (8.69)	0.82	22	81.32 (11.00)	−0.46	8.46
R-CBM	24	28.58 (26.32)	1.05	22	49.95 (26.80)	0.24	21.37
COMPeff	15	0.52 (0.13)	1.17	14	0.59 (0.19)	0.65	0.07
Lexia (School 2)							
WJ Broad Reading	24	78.83 (12.09)	−0.79	23	83.43 (10.36)	−0.53	4.60
WJ Basic Reading	24	86.67 (12.30)	−1.38	23	90.26 (9.43)	−0.10	3.59
WJ Reading Fluency	22	79.23 (13.19)	−0.81	23	82.35 (11.68)	−0.60	3.12
R-CBM	24	44.08 (40.24)	0.46	23	74.61 (37.76)	0.27	30.53
COMPeff	14	0.71 (0.18)	−0.73	12	0.73 (0.15)	−0.57	0.02

Note. COMPeff = Fastbridge comprehension efficiency; R-CBM = reading curriculum-based measurement; WJ = Woodcock–Johnson.

divided by the number of participating students, and (c) required overspill time needed to implement the programs with fidelity, which included the supplemental lessons (in the case of the ILSs) and which was summed across days of intervention.

Results

Descriptives for all outcomes and groups across pretest and posttest are shown in Table 1. Figure 1 reports differences in two major outcomes, *ReadingCBM* and the largest WJ cluster, *WJ Broad Reading*. All outcome data fulfilled traditional statistical assumptions required for ANOVA.

On-Task Behavior

On-task behavior was comparable across conditions. Averaged across the six observations, median *iStation* on-task behavior was 91%, minimum = 72%, maximum = 96%, and for *Lexia* the median was 94%, minimum = 83%, maximum = 97%. The median on-task behavior of the *iStation*'s contrasting BAU was 71%, minimum = 61%, maximum = 77%, and for *Lexia*'s BAU contrast, the median on-task behavior was 89%, minimum = 82%, maximum = 95%. All other recorded behaviors had median frequency rates below 10%, with the exception of verbal off-task behavior for the *iStation* BAU contrast, median = 16.50%, minimum 6%, maximum = 20%.

Examination of Outcomes

Study 1. Results are reported for each of the within-school studies (see Table 2). For the first study, *iStation* versus BAU, students grew at a moderate and significant level on the WJ *Broad Reading,* $t = 22.42$, $p < .01$, $d = .33$, and *Fluency,* $t = 28.66$, $p < .01$, $d = .57$, clusters. A small main effect for group was also observed, $t = 8.82$, $p = .01$, $d = .13$, for the WJ *Reading Fluency* cluster, favoring *iStation*, and no interactions were observed.

Study 2. Results of the *Lexia* study were similar. Students grew at a moderate and significant level in both conditions on all WJ clusters: *Broad,* $t = 12.88$, $p < .01$, $d = .31$; *Basic,* $t = 11.94$, $p < .01$, $d = .22$; and *Fluency,* $t = 17.29$, $p < .01$, $d = .38$. There were neither main effects for group nor interactions.

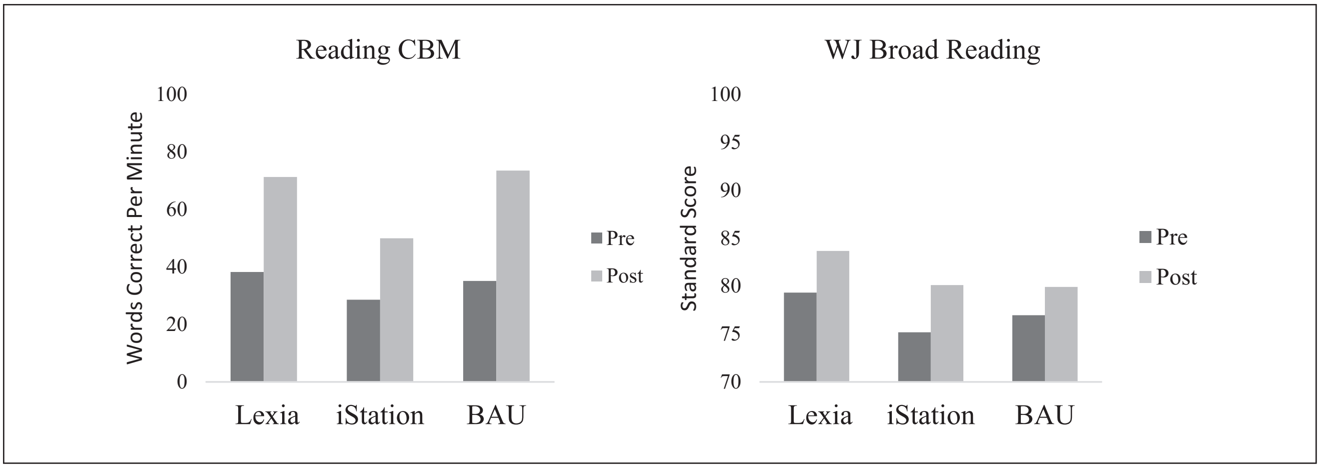

Figure 1. Descriptive differences on two study outcomes, readingCBM and WJ broad reading.
Note. BAU = business-as-usual; CBM = curriculum-based measurement; WJ = Woodcock–Johnson.

Table 2. Results of the Factorial ANOVAs Within School Conditions.

	School 1 (iStation)			School 2 (Lexia)		
Outcome	*t*	*df*	*p*	*t*	*df*	*p*
WJ Broad						
Time	22.42	1	<.01	12.88	1	<.01
Group	0.67	1	.42	0.64	1	.43
Time × Group	0.21	1	.65	0.61	1	.44
WJ Basic						
Time	0.55	1	.46	11.94	1	<.01
Group	0.01	1	.94	0.18	1	.68
Time × Group	0.55	1	.46	0.80	1	.38
WJ Fluency						
Time	28.66	1	<.01	17.29	1	<.01
Group	8.82	1	.01	1.16	1	.29
Time × Group	0.01	1	.93	0.15	1	.70
R-CBM						
Group	0.22	43	.83	0.76	44	.45
Compeff						
Time	4.37	1	.05	0.32	1	.58
Group	2.13	1	.16	1.79	1	.19
Time × Group	0.92	1	.92	0.17	1	.69

Note. ANOVA= analysis of variance; R-CBM = reading curriculum-based measurement; WJ = Woodcock–Johnson.
[a]Significant after family-wise error correction across outcome families (Benjamini–Hochberg procedure; Benjamini & Hochberg, 1995).

Analysis of Instructional Efficiency

We investigated time spent in intervention descriptively across conditions as a means to make further inferences regarding student performance in various conditions. This analysis is summarized in Table 3. It can be seen that the BAU groups, overall, required the most amount of time to implement, with an average range of 333.86 to 469.33 min of IT expended per student. This was expected because instruction was done in small groups of four to seven students, thus requiring significant allocated time across the school day for all students across all grades.

Regarding the two computer-based intervention, *Lexia* required less than half the amount of time to implement with fidelity relative to *iStation*, 155.01 average minutes of IT versus 414.30 average minutes of IT, and also required less time to implement than the BAU, when considering allocated time divided by the number of participating students. This was primarily due to the reported time required to implement *iStation*'s supplemental lessons with fidelity

Table 3. Analysis of Instructional Efficiency.

Group	Cumulative minutes implemented[a]	Average cumulative minutes per student[b]	Number of days implemented
iStation	9,939	414.13	88
iStation Matched Control	28,160	469.33	88
Lexia	4,410	155.01	99
Lexia Matched Control	19,090	333.86	99

[a]This represents the recommended allocated time for intervention (*Lexia* = 45 min, *iStation* = 40 min) and additional time, documented by the teacher each day, required to implement supplemental intervention when recommended by the program ($\bar{X}_{istation}$ = 88.48 min; *Lexia* required no additional time). [b]This represents cumulative minutes implemented, corrected for the number of participating students.

beyond the designated allocated time for the computer-based intervention. Lexia's supplemental lessons, in contrast, were recommended by the program with less frequency, as reported by the implementor, and required less time to organize and implement, as documented by study records and teacher report. That is, differences in IT occurred primarily as a result of implementation of the blended model.

Discussion

The purpose of the current study was to examine the effectiveness of two ILSs relative to traditional remedial instruction for students identified as at-risk in literacy. Given the widespread use of ILSs, the paucity of research on contemporary programs is concerning. Existing research generally focuses on the use of ILSs as a Tier 1 supplement, compared with Tier 1 instruction. In contrast, the current study situated ILSs as a primary intervention for students identified as at-risk, in comparison with students receiving BAU Tier 2 intervention. This study also occurred in a rural setting with students of a high poverty rate and of a unique demographic make-up relative to prior studies. Overall, results tell an interesting story that supports the use of ILSs in schools and highlights the importance of considering IT.

Summary of Outcomes

The two ILSs performed similarly to traditional pull-out intervention across both studies. One main effect was found favoring *iStation*, and no interactions were observed in either case. In a sense, this is reassuring. Technology is expensive (i.e., purchasing software and tablets), but likely less so than hiring, training, and maintaining additional faculty and staff. Underserved schools also have an elevated risk of teacher turnover (Redding & Henry, 2018). For schools facing financial constraints, rural schools particularly so, ILSs hold promise for serving as a supplement or replacement to traditional Tier 2 intervention. Although unexplored in the current study, ILSs can also be accessed from home, and therefore might be particularly suitable for students with higher absence rates, who otherwise have difficulty getting to school, or when transitioning from in-person to remote instruction. Effect sizes across conditions were moderate, suggesting practically meaningful levels of student growth occurred across conditions. Importantly, WJ cluster scores are age-adjusted, so the average participating student increased their proficiency relative to the appropriate norming group for the battery.

However, the two ILSs did present themselves with unique challenges. The programs required a stable and high-speed internet connection. Due to instability of the rural networks, there were times intervention could not be implemented, whereas this did not affect the BAU groups. Granted, across the entire school year, these instances were relatively few. Students also had to be accustomed to working on a tablet with an external keyboard. There were several students who, as anecdotally reported, struggled initially simply because they were not familiar with the use of a QWERTY keyboard. For these students, this resulted in additional IT being expended through teaching the log-in process alone and may have reduced the validity of the pre-assessments. This highlights the importance of the supervising teacher remaining vigilant during the administration of ILSs, which may not be appropriate for all students at-risk. It also is important to reiterate that this study examined the effects of blended models, unlike prior studies. Results should not be generalized to the circumstance of exclusively using *Lexia* or *iStation*'s computer-based interface, as the effect of one or the other component could not be isolated.

Instructional Efficiency

A unique contribution of this study was the analysis of IT (Skinner et al., 1996). One ideal quality of the two ILSs was that these data were passively generated. At any time, the experimenters could remotely call up the amount of time an individual, or group of students, spent on intervention, which we supplemented with estimates of on-task behavior. This was ideal for rural consultation, allowing for the remote monitoring of intervention performance and fidelity by the researchers. In practice, this would make analysis of implementation fidelity when determining RTI easier, given the difficulty of accessing these rural sites.

The analysis of IT suggested that *iStation* and the BAU, surprisingly, took comparable amounts of time to implement. Although Lexia and *iStation* offer slightly different recommendations for minimum IT, the vast majority of this difference was accounted for by the use of supplemental activities. The implementing *iStation* teacher expressed concern over being able to implement the fully blended model with fidelity and reported that, in addition to the time needed to run the in-person lessons, preparation for the lessons took a significant amount of time, which was not considered in our estimates. The *Lexia* interventionist expressed no such concerns and stated that when supplemental lessons were produced, they were able to be prepared and completed within the designated time for intervention. Importantly, this is the verbal report of only one interventionist from each condition. Furthermore, the *Lexia* teacher had additional instruction/intervention training and credentials relative to the *iStation* teacher, although researchers did support preparation/implementation in both conditions.

This analysis does not consider the ratio of students to teachers. Although *iStation* and the BAU took similar amounts of time to implement, the ratio of teachers to students still favored *iStation*. That is, while the BAU was constrained to the small-group scenario, the practical cap for both ILSs is higher. We refrained from factoring this into our analysis so that our estimates remained unambiguous and based on easily understood data.

Limitations and Future Directions

The results of this study should be considered in light of several limitations. First, even though power analysis supported the overall size of our sample, our sample within grades was smaller. Thus, we could not consider differentiated effects across grades. This was an unfortunate necessity. Available schools were small and widely spread apart, as is the nature of rural education. Relatedly, staff resources in these schools were limited, and it is for this reason one intervention condition was facilitated by a reading specialist and the other by a school support staff member. This differential in professional training may have affected results, although research staff were trained by both vendors, monitored activities in both conditions, and assisted when needed. Another limitation was that we calculated IT for the BAU based on the designated time, not actual time, spent in intervention. In contrast, our IT data for the two experimental conditions were based on precise report generated by the programs themselves, in addition to the supplemental time reported to us by the interventionist in delivering the blended activities. Next, this study reported on short-term gains. However, it is possible long-term results are different. Finally, it is critical to stress that the ILSs performed reasonably well against BAU conditions that were not precisely scripted or that relied on a guiding curriculum. Scripted in-person programs are likely to have a stronger effect and may outperform the examined ILSs. This important question requires additional research.

Implications

This study has direct implications for schools considering the nature of their Tier 2 support services. Our findings converge with prior research (e.g., Cheung & Slavin, 2013; Torgesen et al., 2010) to suggest that well-designed, comprehensive, and theoretically grounded ILSs may perform similarly to a common approach to Tier 2 remedial literacy intervention implemented individually or in small groups. However, this recommendation assumes (a) a blended model of delivery, (b) a 1:1 student: tablet ratio, and (c) appropriate vendor-supplied training. Further research is needed, including a parceling of results for individual grades, for students who enter intervention with varying levels of initial performance and who experience virtual or blended delivery, in varying environments (e.g., in school vs. partially or fully remote) and contrasted against more intensive in-person intervention approaches. We therefore refrain from recommending one or the other ILS, but believe there is evidence for the effectiveness of these programs when used at Tier 2, provided a specific set of implementation conditions.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Research reported in this publication was funded by the Education and Innovation Research Program, U.S. Department of Education, grant number U411C170219. The opinions expressed are those of the authors and do not represent views of the U.S. Department of Education.

ORCID iD

Benjamin G. Solomon https://orcid.org/0000-0002-8457-1112

Note

1. The district had recently adopted a 4-day, Monday through Thursday, school week.

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B: Methodological*, *57*, 289–300.

Benner, G. J., Nelson, J. R., Stage, S. A., & Ralson, N. C. (2011). The influence of fidelity of implementation on the reading outcomes of middle school students experiencing reading difficulties. *Remedial and Special Education*, *32*, 79–88.

Cheung, A. C. K., & Slavin, R. E. (2013). Effects of educational technology applications on reading outcomes for struggling readers: A best-evidence synthesis. *Reading Research Quarterly*, *48*, 277–299.

Coyne, M. D., McCoach, D. B., Ware, S., Austin, C. R., Loftus-Rattan, S. M., & Baker, D. L. (2019). Racing against the vocabulary gap: Matthew effects in early vocabulary instruction and intervention. *Exceptional Children*, *85*, 163–179.

FastBridge Learning, LLC. (2015). *Formative Assessment System for Teachers (FAST): Benchmarks and norms version 2.3, 2015-2016*. https://www.fastbridge.org/wp-content/uploads/2015/12/2015-16FastBridgeNormsandBenchmarksAllMeasuresFINAL.pdf

Freeman, C., & Fox, M. A. (2005). *Status and trends in the education of American Indians and Alaska natives* (Report NCES 2005-108). U.S. Department of Education, National Center for Education Statistics. https://nces.ed.gov/pubs2005/2005108.pdf

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice*, *18*, 157–171.

Hibel, J., Faircloth, S. C., & Farkas, G. (2008). Unpacking the placement of American Indian and Alaska Native students in special education programs and services in the early grades: School readiness as a predictive model. *Harvard Educational Review*, *78*, 498–528.

Individuals with Disabilities Education Improvement Act, 20 U.S.C. § 1400 (2004).

iStation. (2019). *iStation reading*. https://www.istation.com/Reading

Lee, J., & Park, O. (2007). Adaptive instructional systems. In J. M. Spector, M. D. Merrill, J. V. Merrienboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed., pp. 469–484). Lawrence Erlbaum.

Lexia Learning. (2019). *Lexia Core5 reading*. http://lexialearning.com/product/core5

Macaruso, P., & Rodman, A. (2011). Efficacy of computer-assisted instruction for the development of early literacy skills in young children. *Reading Psychology*, *32*, 172–196.

Macaruso, P., Wilkes, S., Franzen, S., & Schechter, R. (2019). Three-year longitudinal study: Impact of a blended learning program—Lexia®Core5® reading—On reading gains in low-SES kindergartners. *Computers in the School*, *36*, 2–18.

McDermott, P., & Gormley, K. A. (2015). Teachers' use of technology in elementary reading lessons. *Reading Psychology*, *37*, 121–146.

National Center for Educational Statistics. (2019). *2019 mathematics assessment report card: Summary data tables with additional detail for average scores and achievement levels for states and jurisdictions*. Institute of Education Sciences, U.S. Department of Education.

National Center for Intensive Intervention. (2019). *Academic screening tools chart*. https://charts.intensiveintervention.org/chart/academic-screening

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*.

National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.

O'Callaghan, P., McIvor, A., McVeigh, C., & Rushe, T. (2016). A randomized controlled trial of an early intervention, computer-based literacy program to boost phonological skills in 4- to 6-year-old children. *British Journal of Educational Psychology*, *86*, 546–558.

Pearce, L. R., & Gayle, R. (2009). Oral reading fluency as a predictor of reading comprehension with American Indian and White elementary students. *School Psychology Review*, *38*, 419–427.

Piasta, S. B., Farley, K. S., Mauck, S. A., Ramirez, P. S., Schachter, R. E., O'Connell, A. A., . . . Weber-Mayrer, M. (2020). At-scale, state-sponsored language and literacy professional development: Impacts on early childhood classroom practices and children's outcomes. *Journal of Educational Psychology*, *112*, 329–343.

Poncy, B. C., Solomon, B., Duhon, G., Skinner, C., Moore, K., & Simons, S. (2015). An analysis of learning rate and curricular scope: Caution when choosing academic interventions based on aggregated outcomes. *School Psychology Review*, *44*, 289–305.

Putman, R. S. (2017). Technology versus teachers in the early literacy classroom: An investigation of the effectiveness of the Istation integrated learning system. *Educational Technology Research and Development*, *65*, 1153–1174.

Redding, C., & Henry, G. T. (2018). Leaving school early: An examination of novice teachers' within- and end-of-year turnover. *American Educational Research Journal*, *56*, 204–236.

Renaissance Learning. (2015). *STAR reading: Technical manual*.

Scanlon, D. M., Anderson, K. L., & Sweeney, J. M. (2016). *Early intervention for reading difficulties: The interactive strategies approach* (2nd ed.). The Guilford Press.

Schechter, R., Macaruso, P., Kazakoff, E. R., & Brooke, E. (2015). Exploration of a blended learning approach to reading instruction for low SES students in early elementary grades. *Computers in the Schools*, *32*, 183–200.

Schrank, F. A., McGrew, K. S., Mather, N., & Woodcock, R. W. (2014). *Woodcock-Johnson IV tests of achievement*. Riverside.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Shapiro, E. S. (2011). *Academic skills problems: Direct assessment to intervention* (4th ed.). The Guilford Press.

Skinner, C. H. (2008). Theoretical and applied implications of precisely measuring learning rates. *School Psychology Review*, *37*, 309–314.

Skinner, C. H., Fletcher, P. A., & Henington, C. (1996). Increasing learning rates by increasing student response rates: A summary of research. *School Psychology Quarterly*, *11*, 313–325.

Torgesen, J. K., & Barker, T. A. (1995). Computers as aids in the prevention and remediation of reading disabilities. *Learning Disabilities Quarterly*, *18*, 76–87.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Herron, J., & Lindamood, P. (2010). Computer-assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches. *Annals of Dyslexia*, *60*, 40–56.

Torgesen, J. K., Waters, M. D., Cohen, A. L., & Torgesen, J. L. (1988). Improving sight-word recognition skills in LD children: An evaluation of three computer program variations. *Learning Disability Quarterly*, *11*, 125–132.

University of Oregon. (2018). *8th edition of dynamic indicators of basic early literacy skills (DIBELS®)*. https://amplify-com-mktg.imgix.net/app/uploads/2018/10/27133348/D8-Administration-Scoring-Guide-Jul-2019-1.pdf

Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grader interventions. *Journal of Learning Disabilities*, *39*, 157–169.