Can educational technology effectively differentiate instruction for reader profiles?

Lauren S. Baron[1,2,3] · Tiffany P. Hogan[1] · Rachel L. Schechter[2,4] ·
Pamela E. Hook[2] · Elizabeth C. Brooke[2]

Abstract

Teachers are responsible for identifying and instructing an increasingly diverse population of student readers. Advances in educational technology may facilitate differentiated instruction. Using data from a large, population-based sample of third-grade students, we investigated *what works for whom* in technology-based literacy instruction. We classified 594 students into four reader subgroups or *profiles* based on two scores from a readily-available, commonly-used progress monitoring tool (aimsweb©). Using this simple and accessible method, we identified profiles of readers that resemble the poor decoder, poor comprehender, mixed deficit, and typical reader subgroups found in past studies that used extensive test batteries or advanced statistics. We then employed nonparametric analyses to examine both proximal and distal outcomes across one academic year. First, we compared the relative progress of reader profiles on a technology-based reading program (Lexia® Core5® Reading). Second, we determined whether each reader profile made gains on aimsweb at the end of the year. We found that Core5 effectively differentiated online instruction and contributed to improved aimsweb performance for most reader profiles. The findings from this study help inform educational best practices for efficient identification of and effective intervention for all students.

Keywords Simple view of reading · Reading profiles · Educational technology · Poor comprehenders · Poor decoders

✉ Tiffany P. Hogan
   thogan@mghihp.edu

Extended author information available on the last page of the article

Introduction

Today's educators are faced with an increasingly complex student population. Inclusive classrooms are comprised of students with various learning differences and disabilities (Subban, 2006; Tomlinson, 2000). When it comes to reading, students may struggle in a variety of ways and for a number of reasons. As such, teachers need to accurately and efficiently identify students with reading difficulties and then differentiate instruction to meet the individual needs of each student. Many teachers are turning to technology for instructional support, but the effects of technology-based interventions are not well understood. In this study, we sought to understand what works for whom in technology-based literacy instruction. To accomplish this goal, we followed a three-step line of inquiry. First, we classified students into four reader profiles based on scores from a readily-available, commonly-used progress monitoring tool (aimsweb©). Second, we examined the relative progress of reader profiles on a technology-based reading program (Lexia® Core5® Reading). Third, we evaluated whether individual reader profiles made gains on the progress monitoring tool at the end of the school year.

Identifying reader profiles

Multiple cognitive and linguistic abilities are required for successful reading. These abilities generally contribute to one of two primary skills: word reading or linguistic comprehension (Catts, Hogan, & Fey, 2003; Oakhill, Cain, & Bryant, 2003). Accurate word reading relies on the knowledge of letter-sound relationships and integration of information about phonology, orthography, and morphology. Linguistic comprehension is the product of foundational skills such as vocabulary and syntax as well as discourse level processing such as inferencing, comprehension monitoring, and story grammar knowledge (Language and Reading Research Consortium, (LARRC) & Chiu, 2018). In summary, skilled readers must be able to decode words as well as derive meaning from words and connected text.

Poor readers might struggle because of difficulties with word reading, comprehension, or both. The Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990) is an empirically-supported framework commonly used to guide assessment and subgrouping of readers based on their individual strengths and weaknesses (Aaron, Joshi, & Williams, 1999; Catts, Hogan, & Fey, 2003). Using this theory, students may present with concordant or discordant abilities. It is most common to find concordant abilities: students with good word reading and good comprehension abilities are considered *typical readers*, whereas students who struggle with both word reading and comprehension are labeled as having a *mixed deficit*. Students with the mixed deficit profile often have more severe reading difficulties than their peers with a single deficit (Catts, Hogan, & Fey, 2003). Studies have also identified students with deficits in a single component of reading or discordant reader profiles. Some students struggle with word reading but have a relative strength in comprehension. These students are often labeled *poor decoders* and may even fit the criteria for dyslexia (Catts, Hogan, & Fey, 2003; Petscher, Justice, &

Hogan, 2018; Elwér, Keenan, Olson, Byrne, & Samuelsson, 2013; Hagtvet, 2003). On the other hand, students who have relatively good word reading, but poor comprehension are labeled *poor comprehenders.* These students can read accurately and fluently, but they have difficulty extracting meaning from text.

Many studies necessarily focus on a single reader profile (e.g., Nation, Clarke, Marshall, & Durand, 2004) or recruit a specific number of students for each reader profile (e.g., Cutting, Materek, Cole, Levine, & Mahone, 2009), while few studies identify and examine all reader profiles within a single sample. Shankweiler and colleagues (1999) were among the first to use the Simple View of Reading to classify their full sample of students into each reader profile. Their clinical sample included 361 students in second-grade or above, aged 7.5–9.5 years (Shankweiler and colleagues 1999). Students were classified according to performance on eight different assessments of word reading and reading comprehension using composite scores for each construct of interest and then utilizing cut-scores and a "buffer-zone" to determine group assignment. Students whose scores fell in the buffer-zone were not assigned to a reader profile. This method lessened the influence of individual variation and measurement error for students who performed at or near the boundary between subgroups, thus alleviating some of the disadvantages associated with using arbitrary cut-scores (Dwyer, 1996; Shankweiler et al., 1999). This method resulted in the following reader profile proportions: 44% had mixed deficits, 11% were poor decoders, 6% were poor comprehenders, and 39% were typical readers (Shankweiler et al., 1999). Leach et al. (2003) used a similar methodology with a population-based sample of 289 students in fourth- through fifth-grade. They found the following reader profile proportions: 16% had mixed deficits, 17% were poor decoders, 7% were poor comprehenders, and 59% were typical readers (Leach, Scarborough, & Rescorla, 2003).

The relative proportion of these profiles is fairly consistent across studies, even when more advanced statistical analyses are employed and different languages are examined (Lerkkanen, Rasku-Puttonen, Aunola, & Nurmi, 2004; Torppa et al., 2007). For example, Torppa et al. (2007) used mixture modelling to categorize 1750 Finnish-speaking students into reader profiles according to word reading and reading comprehension abilities. In their sample, 15% had mixed deficits, 28% were poor decoders (i.e., they were slow readers in Finnish), 11% were poor comprehenders, and 46% were typical readers (Torppa et al., 2007). In general, typical readers comprise the largest or most-common profile, while either poor decoders or students with mixed deficits are the next largest group. Poor decoders and poor comprehenders are typically the smallest or least-common reader profiles. Minor variations in this pattern may occur due to differences in study sample, age or grade-range, language, and finally yet importantly, profile selection criteria (Aaron et al., 1999).

Unlike the studies described here, schools have a limited capacity to identify reader profiles by administering a large assessment battery or running advanced statistical tests due to constraints on staff availability or expertise, instructional time, and/or budget. In-depth diagnostic assessments are typically reserved for students who demonstrate significant difficulty with reading that cannot be remediated within a classroom setting (see Gersten et al. 2008 for an overview of multi-tier interventions). There is a need to determine whether more practical methods can be used

to screen large populations and identify reader profiles. The aimsweb assessment system is a standardized measure of reading ability that is commonly-used by teachers for progress monitoring at the beginning, middle, and end of the school year. It contains two norm-referenced subtests of word reading and reading comprehension abilities but does not contain any subtests that measure oral language (e.g., listening comprehension, vocabulary, or syntax). Therefore, aimsweb can be a useful screening tool, but it is not comprehensive enough to replace diagnostic assessment of reading and language abilities.

Notably, past studies of poor reader subgroups—many of which subscribe to the Simple View of Reading—differ in whether they measure listening comprehension (e.g., Catts, Hogan, & Fey, 2003; Elwér et al., 2013) or reading comprehension (e.g., Catts, Adlof, & Weismer, 2006; Leach et al., 2003; Nation et al., 2004; Shankweiler et al., 1999; Torppa et al., 2007), depending on the study aims, research design, and participant sample. Listening comprehension is not synonymous with reading comprehension; however, by third grade, the majority of reading comprehension abilities are explained by individual differences in listening comprehension (Catts, Hogan, & Adlof, 2005; LARRC, 2015; Cain, 2015; Catts et al., 2006; Hagtvet, 2003; Justice, Mashburn, & Petscher, 2013; Nation, et al., 2004). Despite not providing very specific diagnostic information, reading comprehension measures are highly relevant to educational performance. They also fit within the constraints of school settings because they are often available for group administration. In comparison, few listening comprehension measures are available for group administration, and unfortunately, the available measures are often more difficult and time consuming to score than reading comprehension measures. As such, listening comprehension measures are not routinely administered in the school setting (Hendricks,Adlof,Alonzo,Fox, & Hogan, 2019).

In the present study, we examined an existing dataset comprising a large, population-based sample. We used scores from the aimsweb word reading and reading comprehension subtests and applied cut-points used in previous studies (Adlof et al., 2006; Catts et al., 2006) to categorize students into one of the four reader profiles: mixed deficits, poor decoders, poor comprehenders, and typical readers. We hypothesized that reader profiles captured through aimsweb would be similar in proportion to those found in past studies.

Instructing reader profiles

Decades of research and nationally commissioned reports have agreed upon the importance of explicit and systematic reading instruction that emphasizes phonemic awareness, phonics, fluency, and comprehension (e.g., National Reading Panel, 2000). Unfortunately, there is still debate about *how* to provide these comprehensive and effective reading programs (e.g., Foorman, Breier, & Fletcher, 2003). Differentiated instruction is the practice of modifying the focus or format of instruction to meet the needs of individual students; it is typically implemented in small groups and employs frequent progress monitoring (Foorman et al., 2003; Subban, 2006; Tomlinson et al., 2003). Many teachers and schools are turning to a blended learning

model to differentiate instruction and meet the needs of all students (Staker & Horn, 2012). Blended learning incorporates both educational technologies and traditional, offline teaching methods. The technology-based portion of blended learning programs can often support the most time-consuming aspects of differentiated instruction, such as data collection and ongoing progress monitoring. Educational technologies can also provide a systematic scope and sequence for reading skill acquisition as well as targets for additional offline instruction.

One technology-based intervention that appears to be both comprehensive and effective is the Lexia Core5 Reading (Core5) program (Schechter, Macaruso, Kazakoff, & Brooke, 2015; Prescott, Bundschuh, Kazakoff, & Macaruso, 2017). Core5 was designed to accelerate mastery of reading skills for students of all abilities in preschool through fifth-grade. The online component of Core5 contains systematic, personalized, and engaging instructional activities. The offline component of Core5 utilizes student performance data to create detailed progress reports linked to targeted resources for teacher-led lessons or independent paper-and-pencil activities. This study will examine the online portion of Core5, which contains 18 Levels corresponding to specific grade-level material: preschool (Level 1), kindergarten (Levels 2–5), first grade (Levels 6–9), second grade (Levels 10–12), third grade (Levels 13–14), fourth grade (Levels 15–16), and fifth grade (Levels 17–18). The content in each level is aligned with the Common Core State Standards. The program activities are organized into six strands of reading skills: Phonological Awareness, Phonics, Structural Analysis, Automaticity/Fluency, Vocabulary, and Comprehension.

There are two key program features that personalize the learning experience for each student: auto placement and instructional branching. All students begin Core5 by taking an *auto placement* test. This brief screening tool determines the appropriate Start Level for each student based on their reading ability. Thus, students may begin Core5 by working on material that is below, in, or above their current grade level. Students then work on units within activities at the designated Core5 level until they achieve 90–100% accuracy and advance to the next level. When students have difficulty in a given unit of Core5, an *instructional branching* feature automatically differentiates task presentation. Figure 1 provides a screenshot for each instructional mode for an example unit. All students begin the program in the *standard* instructional mode. If students make 1–2 errors in a unit, they are automatically moved to the *guided practice* mode with fewer stimuli and more structure. If

Fig. 1 Screen shots of a multiple meaning words sorting task from the vocabulary strand of Core5 in each of the instructional modes (from left to right: standard, guided practice, direct instruction)

students are successful, they return to the standard mode. If students continue to make errors, they are moved to the *direct instruction* mode that explicitly teaches the skill in the activity (e.g., multiple meanings) and the type of error the student made (e.g., bat).

No one has explicitly studied the effects of Core5 for poor reader profiles. In this study, we compared the performance of reader profiles on Core5 for one school year. Performance was gauged by three key measures: accuracy in the standard mode, time spent in the guided practice mode, and time spent in the direct instruction mode. The performance of typical readers served as a comparison group for the impaired reader profiles. We hypothesized that typical readers would achieve greater accuracy in the standard mode of Core5 while spending less time in the guided practice or direct instruction modes. In contrast, students with mixed deficits might achieve lower accuracy in the standard mode compared to the other profiles. We predicted that students with discrepant reading skill profiles would achieve lower accuracy in the standard mode compared to their typical peers, but only in their area of weakness. For example, poor decoders might achieve lower accuracy than typical readers on Phonics activities. Another possibility however, is that these students could achieve similar accuracy as the typical readers if time in the guided practice or direct instruction modes facilitates performance in the standard mode. Using the same example, poor decoders might demonstrate equal accuracy during Phonics activities if they spend significantly more time in the guided practice mode of Phonics instruction.

Proximal and distal measures of reading performance

Reading performance can be assessed using proximal or distal measures. The terms proximal and distal typically refer to the relative distance between two objects. In assessment, distance is comparable to how closely a measure is associated with the mechanism of change; proximal is close or highly-related and distal is more separated. If the mechanism of change is a particular program of literacy instruction, the proximal measures are the embedded program metrics, and the distal measures are outside assessments. In other words, the proximal measures tell us about performance in the reading program, but we need distal measures to determine whether any changes were made in a more ecologically valid context. Regarding Core5 metrics, we anticipated that significant between-group differences in Fall performance variables would disappear in the Spring, which would indicate improvement. If such a pattern of change occurred proximally, we could potentially see gains on the more distal reading assessment, aimsweb.

In summary, the goal of this paper was to determine *what works for whom* in technology-based literacy instruction. The investigation was organized around three key questions. First, can scores from a readily-available, commonly-used progress monitoring tool (aimsweb) be used to classify third-grade students into four reader profiles commensurate with past research? Second, on average, do students in the four reader profiles perform differently from one another on a technology-based reading program (Lexia Core5 Reading) in word reading and comprehension activities, when controlling

for program start level? Third, how does each reader profile perform on aimsweb in the spring compared to the fall? Results of this study will contribute to practical and efficient assessment and intervention for students with reading difficulties.

Method

Participants

Participants were selected from a large sample that experienced a statewide initiative to improve reading outcomes for elementary students. The participants in this study were third-grade students who used Lexia Core5 Reading with fidelity for one school year. We selected third-grade students for this study because it is a critical period in school when students are shifting from "learning to read" to "reading to learn," and many struggling readers are identified (Chall, 1996). We also wanted to isolate a single developmental period for language and literacy learning given the differential impact of component reading skills over time (Catts et al., 2006). Next, we only included students who had complete assessment data from the Fall testing period. Finally, we required students to use the intervention program with fidelity; that is, use Core5 for a minimum of 20 weeks and meet the recommended usage (20–80 min) for at least 50% of those weeks. The amount of recommended weekly minutes varied based on student risk level, as determined in Core5. Past studies of Core5 suggest that using the program with this fidelity criterion is required for efficacy (Prescott et al., 2017; Schechter et al., 2015). After applying these inclusion criteria, 1119 students were considered for reader profile classification, which is detailed at the end of the Methods section.

Procedures

The aimsweb was administered to students in the Fall and Spring of the 2014–2015 school year. In this midwestern state, the school year begins in August and ends in May; thus, Fall scores were obtained between mid-August and mid-October, while Spring scores were obtained between beginning-April and mid-May. During the school year, two primary reading curricula were used in a blended learning approach: *Treasures* McGraw-Hill© 2011; now sold as *Wonders*) and *Lexia Core5 Reading* (Core5). Treasures is a more traditional, teacher-guided educational program for students in kindergarten through sixth-grade containing leveled texts and instructional materials. Core5 was implemented as a supplemental, technology-based instructional program. For additional information about the implementation of Lexia Core5 Reading in this sample, see Kazakoff, Macaruso, and Hook (2018). On average, students spent 20–80 min weekly in the online component of Core5, and approximately 60–120 min daily in offline instruction (i.e., Treasures and/or Core5). The Fall instructional period lasted from September through December, while the Spring instructional period lasted from January through April.

Measures

Both subtests of the aimsweb are comprised of Standard Reading Assessment Passages that aimsweb developers created for each grade level and modified for use with either the R-CBM or Maze procedures. There are 33 passages for third-grade students that are designed to have consistent readability based on the Lexile-graded standards. As such, all passages can be administered at any time during the school year (Howe & Shinn, 2002). Of note, the aimsweb passages resulted in significantly less measurement error when used to measure word reading as compared to the DIBELS passages—another progress monitoring tool that is commonly used in schools (Ardoin & Christ, 2009).

Word reading

The *Reading Curriculum-Based Measure* (R-CBM) from aimsweb is an individually administered test of oral word reading fluency. Notably, fluency or efficiency—as opposed to accuracy—is a better indicator of word reading ability for students in or above third-grade (LARRC, 2015). In the R-CBM, students read a 300-word passage aloud for 1 min. Words were scored as correct if they were pronounced accurately and in the correct order. Minor speech distortions were not counted against the score, and self-corrections were allowed if they occurred within 3 s of the error. The final score was the total number of words read correctly. For typical progress monitoring, three passages are administered, and the median score is reported.

According to the aimsweb technical manual (NCS Pearson, 2012), R-CBM scores correlated with North Carolina End of Grade Test at approximately $r = .70$ for third-grade students. The manual also reported correlations between R-CBM and other reading assessments (including the Measures of Academic Progress) from published studies that ranged from $r = .65–.81$, demonstrating criterion validity. According to the technical manual, alternate-form, test–retest, and split-half reliability was approximately .94 for third-grade students.

Reading comprehension

The *Reading Maze Test* (Maze) from aimsweb is an individual or group administered test of reading comprehension. Students had 3 min to read a 300-word passage silently and complete a multiple-choice cloze task. The first sentence in the passage was unaltered; thereafter, every seventh word was replaced with three options (i.e., one exact match and two distractor words) inside parentheses. Students were instructed to circle the word that belongs in the sentence. The final score was the total number of words circled correctly.

According to the aimsweb technical manual (NCS Pearson, 2012), Maze scores correlated with end of year state reading assessments at approximately $r = .59$ for third-grade students, demonstrating criterion validity. Test–retest reliability was approximately .70 for third-grade students. It should be noted that there are multiple ways to assess reading comprehension with variable influence from word reading ability (see: Keenan, Betjemann, & Olson, 2008). The Maze subtest was our only

option for measuring comprehension within aimsweb; however, it has been shown that this type of assessment may be influenced by word reading more than other reading comprehension assessments using multiple choice or open response questions (Keenan et al., 2008; Muijselaar, Kendeou, de Jong, & van den Broek, 2017).

Core5 performance

For this study, we combined relevant Core5 strands (i.e., Phonological Awareness, Phonics, Structural Analysis, Automaticity/Fluency, Vocabulary, and Comprehension) to create performance domains that aligned with the Simple View of Reading. We combined Sight Words (part of the Automaticity/Fluency strand) and Phonics to create a *Word Reading Domain*, and we combined Vocabulary and Comprehension (including both listening and reading comprehension activities, depending on the level) to create a *Comprehension Domain*. Many third-grade students will be too advanced for the Phonological Awareness activities and not yet at the level for Structural Analysis activities, which is why we did not include these strands in the domains.

Within the Word Reading and Comprehension Domains, we examined the two key features of Core5 that personalize the learning experience. First, we used Start Level, as determined by the auto placement tool, as a control variable in all Core5 analyses. This is because each Core5 level has a unique set of activities that vary in complexity, which prohibits direct comparisons from one level to another. For example, the relative difficulty of 15 min on a Level 13 (beginning of Grade 3) Passage Comprehension activity with informational text is not directly comparable to 15 min on a Level 6 (beginning of Grade 1) Picture-Phrase Match activity. Therefore, controlling for Start Level allowed us to evaluate profile performance on Core5 as though all students were working within the same level of difficulty.

Second, we examined performance in each of the instructional branching modes. Performance variables include: accuracy in the standard mode, and time (minutes) spent in the guided practice and direct instruction modes for the Word Reading and Comprehension Domains during the Fall and Spring. We analyzed a total of 12 dependent variables (3 instructional modes × 2 domains × 2 time periods). We interpreted various patterns of instructional branching to represent profile strengths and weaknesses. For example, high accuracy in the standard mode indicated strength in a given domain, whereas large amounts of time spent in practice or instruction modes indicated weakness.

Criteria for reader profile classification

We created reader profiles based on aimsweb subtest performance for the 1119 students who met the inclusionary criteria described above. See Fig. 2 for a visual summary of the profile classification method. First, we excluded 140 students who achieved a score at or above the 85th percentile on either the R-CBM or Maze. By removing exceptionally strong readers, the typical readers were comparable to
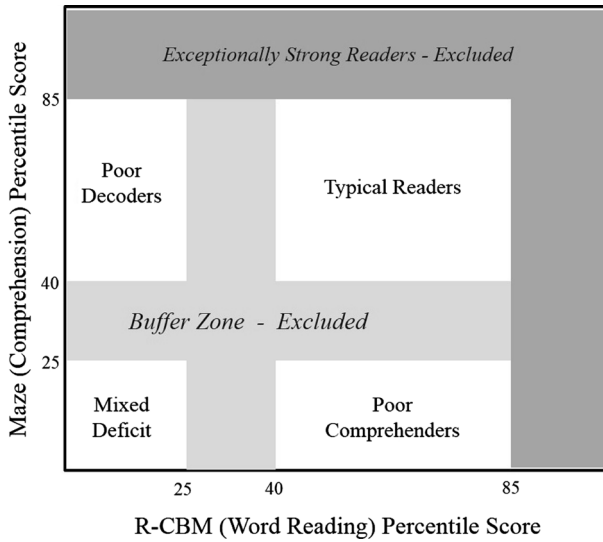
Fig. 2 Reader profile classification method

the poor decoders and poor comprehenders in their areas of relative strength (Catts et al., 2006). We then classified the remaining 979 students into four profiles by imposing cut-points at the 25th and 40th percentiles of the aimsweb subtests. We chose these cut-points and applied profile labels to maintain consistency with prior studies of poor reader subgroups (Adlof et al., 2006; Catts et al., 2006). Students who performed below the 25th percentile on the R-CBM and above the 40th percentile on the Maze were classified as *poor decoders*. Students who performed above the 40th percentile on the R-CBM and below the 25th percentile on the Maze were classified as *poor comprehenders.* Students who performed below the 25th percentile on both subtests were classified as having a *mixed deficit.* Finally, students who performed above the 40th percentile and below the 85th percentile on both subtests were classified as *typical readers.*

This sub-grouping method utilized a *buffer zone* in which students who scored between the 25th and 40th percentile on either subtest did not get classified into a profile. Other studies of reader profiles—studies of poor comprehenders in particular—have validated this method (e.g. Adlof & Catts, 2015; Elwér et al., 2013, Shankweiler et al., 1999). While the buffer zone results in some data loss, the benefit of having clearly delineated subgroups outweighs the cost.

Analytic plan

Data analysis proceeded in three steps—one for each research question. First, we classified readers into profiles using the criteria described in the Method. Second, we compared average between-group profile performance on Core5 (i.e., accuracy or time in each instructional mode and domain) in the Fall and Spring. Third, we compared within-group profile performance on aimsweb scores from the Fall to the

Spring. For the second and third steps, we planned to conduct multivariate analyses of covariance; however, the resulting reader profiles were unequal in size and did not meet the assumptions for parametric analyses. Therefore, we used nonparametric methods.

Core5 performance was analyzed using Quade's procedure for rank analysis of covariance (Quade, 1967). Each of the Core5 performance variables was ranked and then regressed on Fall Level. Using the unstandardized residuals for each Core5 performance variable as dependent variables, we conducted two MANOVAs—one for Fall and one for Spring. We analyzed Fall and Spring effects separately to control for possible differences between time periods and to increase power (see van de Sande, Segers, & Verhoeven, 2016). The independent variable was the four profile groups. When the MANOVA was significant, as indicated by Wilk's lambda, we investigated univariate tests for each of the variables. Effect sizes were interpreted using Richardson's (2011) guidelines for partial eta-squared (η_p^2), where a value of less than .01 was considered small, .06 was considered medium, and .14 was considered large. When univariate analyses were significant, post hoc tests were used to examine pairwise profile comparisons. Bonferroni corrections were applied to all univariate and post hoc analyses.

The data from aimsweb were not appropriate for Quade's procedure because the distribution of the covariate (Fall aimsweb score) was not equal across groups, which violated the nonparametric assumptions (Quade, 1967). As such, we employed a nonparametric repeated measures analysis, or Friedman's Test, for each profile. The percentile scores for each aimsweb subtest (R-CBM and Maze) in the Fall and Spring were the variables of interest.

Results

Reader profiles

The reader profile proportions for our sample are displayed in Table 1. There were 271 students with mixed deficits, 44 students with poor word recognition but relatively good reading comprehension ability (i.e., poor decoders), 45 students with poor reading comprehension but relatively strong word reading skills (i.e., poor

Table 1 Reader profiles based on Fall aimsweb performance

Profile	N	% of total	% of profiles
Mixed deficit	271	28	46
Poor decoder	44	4	7
Poor comprehender	45	5	8
Typical reader	234	24	39
Sub-total	594	61	100
Unspecified	385	39	
Total	979	100	

comprehenders), and 234 students with strong word recognition and reading comprehension (i.e., typical readers).

Over 60% of the third-grade students fit the criteria for reader profiles in this study. About one-third of the students (39%) scored within the buffer zone on one or more aimsweb subtests, and therefore, were not included in further analyses. Of the 594 students who were categorized within a profile, the largest proportions had concordant skill levels on both subtests: 46% had mixed deficits and 39% were considered typical readers. A relatively small proportion of students demonstrated skill discrepancies: 7% of students presented with a poor decoder profile and 8% of students presented with a poor comprehender profile. In general, these profile proportions are in line with past studies of population-based samples (Leach et al., 2003; Shankweiler et al., 1999; Torppa et al., 2007).

Availability of demographic data for students was limited, but when compared with the 2015 Census, data were representative of the midwestern state from which they were collected. Eighty-nine percent of the sample reported gender, but only 24% reported primary language and only 8% reported race. From the reported data, 50% of students were female, 82% identified as white, and 94% reported English as their first language. Per the U.S. Census Bureau, in 2015, 50% of residents in this state were female, 84% were white, and 89% spoke only English.

Core5 performance (between-group analyses)

As mentioned in the Analytic Plan section, we used the unstandardized residuals of the Core5 performance variables in all Core5 performance analyses. Results using these values are reported in the text; complete descriptive statistics for these values are reported in "Appendix". The contextual meaning of the residual values is not easily interpreted, nor are they necessarily meaningful when considering Core5 performance. Therefore, to increase interpretability, figures contain estimated marginal means in their actual unit of measure (e.g., time in minutes) by profile, controlling for Fall Level. These means were obtained by running a one-way analysis of covariance for each dependent variable with profile as the independent variable and Fall Level as the covariate. Markers of significance are based on the nonparametric results.

Fall results

The MANOVA for Core5 performance in the Fall yielded a significant effect of profile, $F(18, 1341) = 6.310$, $p < .001$, partial eta2 = .074. Univariate follow-up tests revealed significant differences, $F(3, 479)$, all $p's < .001$, for all six dependent variables: in the Comprehension Domain: (1) proportion correct in the standard instructional mode, (2) time spent in guided practice, (3) time spent in direct instruction, and in the Word Reading Domain: (4) proportion correct in the standard instructional mode, (5) time spent in guided practice, (6) time spent in direct instruction. Results of Bonferroni corrected post hoc profile comparisons are presented in Fig. 3.
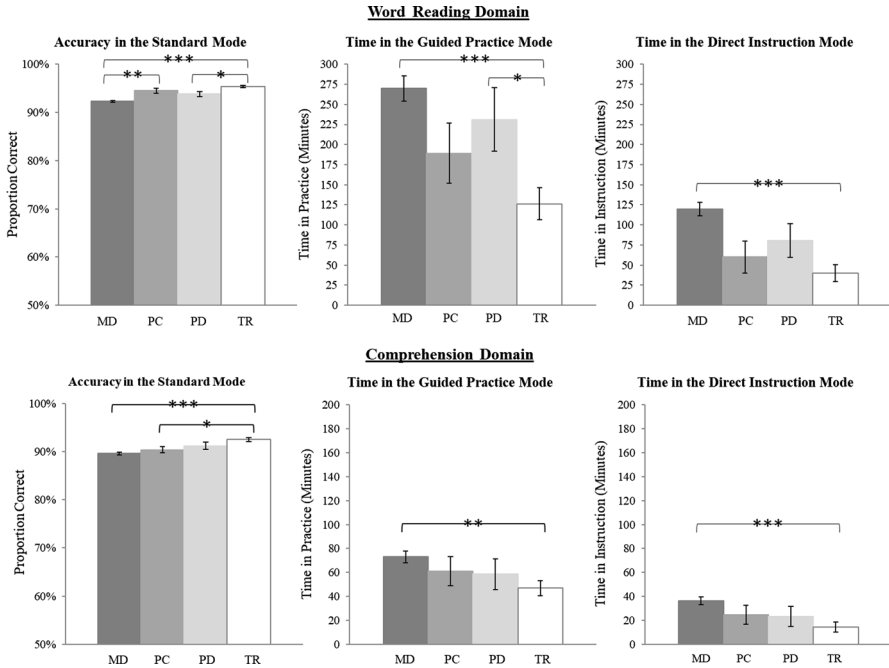
Fig. 3 Fall Core5 Performance by Profile on the Word Reading and Comprehension Domains. Note. MD = Mixed Deficits, PC = Poor Comprehenders, PD = Poor Decoders, TR = Typical Readers. Covariates appearing in the model are evaluated at the following values: Fall Level = 9.82. Error bars represent standard error of the mean. *$p<.05$, **$p<.01$, and ***$p<.001$ for between-group differences on post hoc tests with Bonferroni correction

In the Word Reading Domain, students who were typical readers achieved a significantly higher proportion correct in the standard mode of instruction compared to students with mixed deficits ($p<.001$) and poor decoders ($p=.039$). Poor comprehenders also achieved a significantly higher proportion correct in the standard mode compared to students with mixed deficits ($p=.002$). Typical readers spent significantly less time in the guided practice mode ($p<.001$) and the direct instruction mode ($p<.001$) than their peers with mixed deficits. Typical readers also spent significantly less time in the guided practice mode than poor decoders ($p=.048$).

In the Comprehension Domain, students who were typical readers achieved a significantly higher proportion correct in the standard mode of instruction when compared to students with mixed deficits ($p<.001$) and poor comprehenders ($p=.023$). Typical readers also spent significantly less time in the guided practice mode ($p=.004$) and the direct instruction mode ($p<.001$) than their peers with mixed deficits.

Spring results

The MANOVA for Core5 performance in the Spring yielded a significant effect of profile, $F(18, 1095)=3.822$, $p<.001$, partial eta2 = .056. Univariate follow-up tests

revealed significant differences for 4 of the 6 dependent variables. The effect of profile was significant, $F(3, 392)$, all p's $< .001$, for the following four dependent variables: in the Comprehension Domain: (1) proportion correct in the standard mode of instruction; in the Word Reading Domain: (2) proportion correct in the standard mode of instruction, (3) time spent in the guided practice mode, and (4) time spent in the direct instruction mode. The effect of profile was not significant for remaining two variables in the Comprehension Domain: (1) time spent in the guided practice mode ($p = .782$), and (2) time spent in the direct instruction mode ($p = .109$).

Results of Bonferroni corrected post hoc profile comparisons are presented in Fig. 4. In the Word Reading Domain, students who were typical readers achieved a significantly higher proportion correct in the standard mode of instruction compared to students with mixed deficits ($p < .001$). Similar to the Fall, typical readers spent significantly less time in the guided practice mode ($p < .001$) and the direct instruction mode ($p < .001$) than their peers with mixed deficits.

In the Comprehension Domain, students who were typical readers achieved a significantly higher proportion correct in the standard mode of instruction when compared to students with mixed deficits ($p < .001$). There were no significant differences bbetween profiles for time spent in the guided practice or direct instruction modes during the Spring.
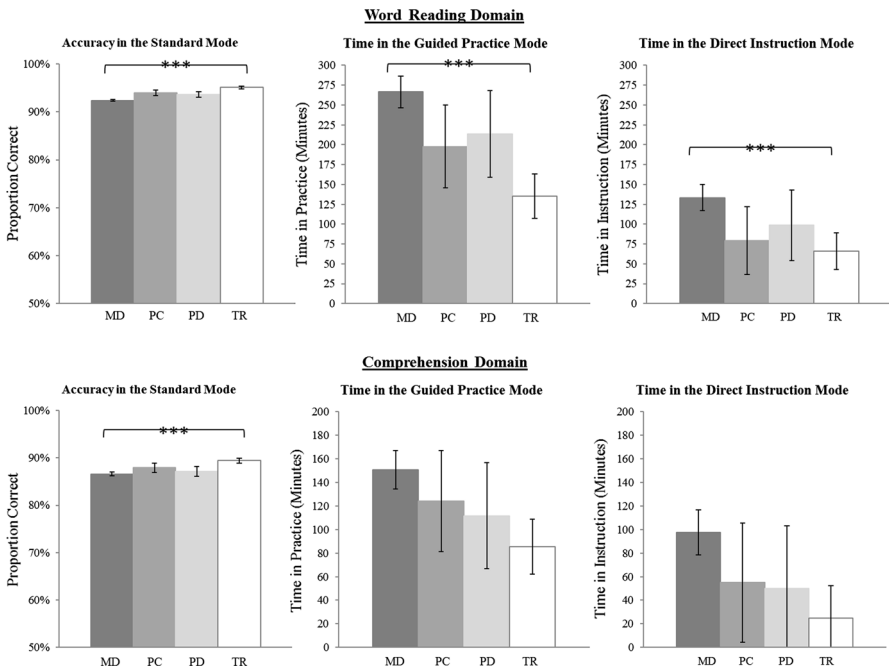
Fig. 4 Spring Core5 Performance by Profile on the Word Reading and Comprehension Domains. Note. MD = Mixed Deficits, PC = Poor Comprehenders, PD = Poor Decoders, TR = Typical Readers. Covariates appearing in the model are evaluated at the following values: Fall Level = 9.82. Error bars represent standard error of the mean. *$p < .05$, **$p < .01$, and ***$p < .001$ for between-group differences on post hoc tests with Bonferroni correction

## aimsweb performance (within-group analyses)

R-CBM results

On the R-CBM subtest of the aimsweb, which measured word reading, all profiles made significant gains from Fall to Spring except the poor comprehenders (See Fig. 5). The poor comprehenders performed at the 51st percentile in the Fall and the 52nd percentile in the Spring ($\chi^2(1)=0.36$, $p=0.546$). On average, the students with mixed deficits increased from the 12th percentile in the Fall to the 18th percentile in the Spring ($\chi^2(1)=57.07$, $p<.001$). The poor decoders increased from the 18th percentile to the 31st percentile ($\chi^2(1)=19.56$, $p<.001$), and the typical readers increased from the 60th percentile to the 63rd percentile ($\chi^2(1)=10.11$, $p=.001$).

Maze results

On the Maze subtest of the aimsweb, which measured reading comprehension, all profiles performed significantly different in the Spring compared to the Fall except the typical readers (See Fig. 6). Typical readers were at the 63rd percentile in the Fall and the 60th percentile in the Spring ($\chi^2(1)=1.60$, $p=.205$). On average, the students with mixed deficits increased from the 10th percentile in the Fall to the 21st percentile in the Spring ($\chi^2(1)=49.71$, $p<.001$). The poor comprehenders increased from the 13th percentile to the 36th percentile ($\chi^2(1)=28.90$, $p<.001$). The poor decoders performed at the 52nd percentile in the Fall and declined to the 40th percentile in the Spring ($\chi^2(1)=6.42$, $p=.011$).
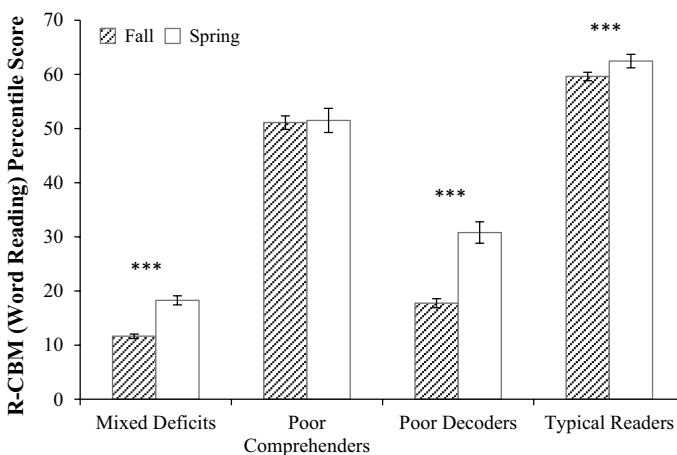
Fig. 5 Average profile performance on the R-CBM (Word Reading) subtest of the aimsweb in the Fall (shaded) and Spring (empty)
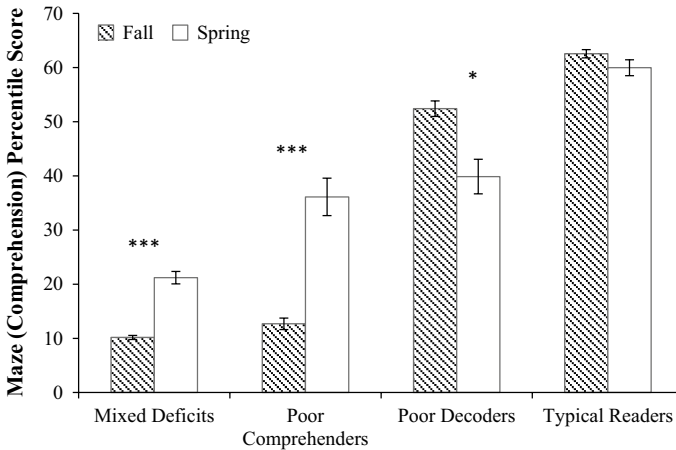
Fig. 6 Average profile performance on the Maze (Comprehension) subtest of the aimsweb in the Fall (shaded) and Spring (empty)

Discussion

We investigated a statewide sample of third-grade students to identify reader profiles and examine their relative performance on a technology-based reading program. Using a single progress monitoring tool, we were able to identify reader profiles that roughly aligned with previous studies using similar sample characteristics and subgrouping methodology, but substantially larger assessment batteries or complex statistical methods. Average profile performance on a technology-based program reflected profile weaknesses in the Fall; however, differences in accuracy compared to typical readers were resolved in the Spring for most students. Readers with mixed deficits appeared to be the least responsive to the technology-based instruction. Furthermore, profile performance on more distal measures generally indicated maintenance or growth of component reading skills from Fall to Spring, with the exception of poor decoders on the reading comprehension measure. This study provides preliminary evidence that a technology-based reading program can effectively personalize instruction for reader profiles and may contribute to improved reading skills. Importantly, there are some caveats to interpreting our results and applying them to classroom practices that we will provide throughout the discussion.

Identifying reader profiles using a single progress monitoring tool

Prior studies have documented the presence of reader profiles in large samples (e.g., Catts, Hogan, & Fey, 2003; Shankweiler et al., 1999; Torppa et al. 2007); however, few have done so using a single, readily-available progress monitoring tool to

classify students (cf. Foorman, Petscher, Stanley, & Truckenmiller, 2018). In the present study, we applied previously established cut-points for reader profile classification to an extant dataset that measured word reading and reading comprehension with subtests from the aimsweb progress monitoring tool. A strength of the present study was the population sample, which was large and representative of students in the midwestern state from where they were recruited. We applied strict inclusion criteria to this initial sample to create well-defined reader profiles. The reader profiles we found were mostly comparable to studies with similar classification criteria and sample characteristics. The largest proportions of students were classified as having concordant reading abilities, fitting a typical reader or mixed deficits profile. Small but relatively substantial proportions of students were classified as having discordant reading abilities, fitting the poor decoder or poor comprehender profile. However, reading is a complex skill that is dependent on multiple subskills, and thus, all profile classification methods are subject to error.

First, we will describe two key points of comparison between our study and existing studies of reader subgroups, and then we will discuss some limitations of our study findings. The first key point is that the proportion of students in the buffer zone (39% of the total sample) was higher than the 11–20% of students in the buffer zone of prior studies (Shankweiler et al., 1999; Catts et al., 2006). Some studies reported using a profile classification method with a buffer zone but do not provide the number of excluded students (e.g., Leach et al., 2003). Other studies use a buffer zone but only recruit and enroll students who can be classified into one of the four reader profiles (e.g., Cutting et al., 2009), so further comparison is limited. The second key point is that our study found a larger proportion of readers with mixed deficits (46%) than some of the past studies: 16% (Leach et al., 2003), 15% (Torppa et al., 2007). However, Shankweiler et al. (1999) also found a similarly large proportion of readers with mixed deficits (44%) using a composite score for multiple word reading and reading comprehension assessments. One possible explanation for these cross-study differences could be the variable influence of word reading abilities on different reading comprehension tasks (Keenan et al., 2008; Muijselaar et al., 2017). Given that Maze-type tasks are known to depend more on word reading skills compared to other reading comprehension tasks, it is possible that we identified an inflated number of students with mixed deficits and a deflated number of poor comprehenders in our sample due to the measure we used.

We believe that our overall method is a feasible approach for screening, such as in a school-wide response to intervention framework. However, we acknowledge that aimsweb does not measure oral language or listening comprehension, and therefore, it cannot provide a truly differential diagnosis for the cause of reading difficulty according to the Simple View of Reading. This is a particularly important caveat for applying this method of poor reader classification to classroom practice. If instructional targets are decided based on aimsweb performance alone (which we know is

highly dependent on word reading ability, regardless of the subtest), students' listening comprehension skills are not well defined. Thus, students classified as poor decoders are at a higher risk for not receiving the appropriate instruction if lessons only target their area of weakness as prescribed by aimsweb. Schools have an obligation to educate all students, including those who are not easily classified into a reader profile or those who do not respond to initial interventions. For these students, and poor decoders in particular, further assessment of listening comprehension and the oral language components that underlie it (e.g., vocabulary, syntax, and inferencing) is imperative to determine a more precise cause of reading difficulties (see Gersten et al., 2008; National Center on Intensive Intervention, 2013). In conclusion, we maintain there is practical value in using a single progress monitoring tool to classify students into reader profiles under the assumption that further diagnostic assessment is utilized to fully characterize students' strengths and weaknesses and inform individualized instruction.

We identified a couple possible directions for future research on reader profiles that we could not address within the design or scope of this paper. First, studies could explore the use of other screening measures or progress monitoring tools to identify reader profiles from large, population-based samples (e.g., Hendricks et al., 2019). It is apparent that the use of different measures precludes direct comparison across studies; however, the reality of school budgets and curricula is that different measures of reading comprehension are in use across the country. Second, studies could examine if reader profiles remain stable within and across grades. For example, Catts, Hogan, & Fey (2003) found that reader profiles remained relatively stable from second- to fourth-grade, with almost two-thirds of the sample maintaining their original profile classification. It follows then, that intervention history is an important element to consider when examining reader profiles over time.

Educational technology can effectively differentiate instruction for reader profiles

Differentiated reading instruction is an elusive goal for many teachers, and it is apparent that measures such as aimsweb provide useful but limited information for guiding instruction. Alternatively, the Lexia Core5 Reading (Core5) technology-based program appears to meet the needs of individual students in all reader profiles by adapting the level of difficulty, complexity of content, and amount of direct teaching across a comprehensive program of reading components. In this study, we compared the performance of poor reader profiles on Core5 to that of typical readers for two time periods. The performance measures included: accuracy in the standard mode of instruction, time spent in the guided practice mode, and time spent in the direct instruction mode, controlling for the level at which students started the

program. Overall, there was increased variability in these measures during the Spring compared to the Fall. It is possible that students who were classified in a poor reader profile in the Fall may have transitioned to a typical reader profile in the Spring. This study did not examine profile stability, but this is a worthy topic for future research, as described in the previous section.

Importantly, Core5 adapted to students' individual differences in performance rather than a categorical label. In the Word Reading Domain, we saw that poor decoders were less accurate in the standard mode of instruction than typical readers during the Fall usage period. This difference was not significant in the Spring, which means the poor decoders' accuracy in Core5 improved during the year. Notably, poor decoders also spent more time in the guided practice mode than the typical readers. It is likely that this increased support in the Fall allowed poor decoders to close the accuracy gap with typical readers in the Spring. A similar pattern of results was found for poor comprehenders in the Comprehension Domain. We saw that poor comprehenders were less accurate in the standard mode of instruction than typical readers in the Fall, but this difference was no longer significant in the Spring. Students with mixed deficits were less accurate in the standard mode of instruction for the Word Reading Domain when compared to both poor comprehenders and typical readers. Poor comprehenders have a relative strength in word reading, so it makes sense that they achieved higher accuracy than their peers with mixed deficits. Students with mixed deficits were also less accurate in the standard mode of instruction for the Comprehension Domain when compared to typical readers. Across both domains, the differences in accuracy for students with mixed deficits compared to typical readers remained significant in the Spring, despite spending more time in the scaffolded instructional modes for the duration of the school year.

In summary, it appears that increased time spent in the scaffolded instructional modes for the domain of difficulty during the Fall allowed poor comprehenders and poor decoders to close the accuracy gap with typical readers in those domains in the Spring. This pattern was more apparent for word reading than comprehension. Students with mixed deficits did not benefit as much from the technology-based instruction and would likely benefit from additional teacher-led instruction (cf. Connor et al., 2018). Programs like Core5 provide valuable information to teachers about the reading and language abilities of students by collecting and reporting student performance data. The added benefit is that the students do not miss any instructional time to participate in formal assessment.

With regards to intervention for reader profiles, we identified a few directions for future work. First, the classification of students into reader profiles, such as those

presented in this study, could be built into technology-based literacy programs to further inform offline instruction and generalization of skills. The wealth of student performance data, inclusive of word reading, listening comprehension, reading comprehension abilities and more, could yield far more accurate reader profiles than traditional progress monitoring tools. In addition, future studies could investigate behavioral predictors of which students, across all reader profiles, will respond best to technology-based interventions and which will require a different approach. Another interesting direction is to examine the mechanism of change within the intervention program. Given the retrospective design of this study, we were unable to determine which features of the scaffolded instructional modes were the most beneficial. For example, was it the reduction of task complexity (e.g., the decreased number of response choices in the guided practice mode) or simply the increased time spent with a given concept that lead to improved accuracy? Thus, future work could examine the mechanism of change by controlling various task features.

Core5 related to improved aimsweb performance

There is prior evidence that Core5 performance is highly correlated with (Mitchell, Baron, & Macaruso, 2018) and predictive of (Prescott et al., 2017) performance on distal reading assessments. We anticipated that Spring scores on our distal measure, aimsweb, would be the same or better than Fall scores given the positive proximal effects of Core5 and the expected academic gains for a population-based sample across one school year. In other words, grade-based percentile scores should be unchanged from Fall to Spring if a selected sample makes progress at the same rate as the normed sample. Ideally though, we expected to see growth in the percentile scores that were below average in the Fall, such as the word reading scores of poor decoders. In this study, we compared the Fall and Spring scores for each profile on the two aimsweb subtests.

For the oral word reading fluency measure (R-CBM), all profiles either maintained their scores or made gains from Fall to Spring. The students with identified deficits in word reading—the poor decoders and readers with mixed deficits—made significant gains on this measure. Typical readers also made small, but significant improvements in word reading. On average, poor comprehenders remained the same—solidly within the average range—in the Fall and Spring. We did not expect poor comprehenders to significantly improve in word reading because it is not an area of weakness. For the comprehension measure (Maze), the students with identified deficits in comprehension—poor comprehenders and readers with mixed deficits—made significant gains from Fall to Spring. Notably, poor comprehenders progressed from below average to well-within the average range. On average, typical readers remained the same in the Fall and Spring. Finally, poor decoders showed a significant decline in comprehension from Fall to Spring, but maintained a typical-borderline ability level. This decline could be expected due to ongoing difficulty with word reading that impedes successful comprehension of text as it increases in

difficulty. However, aimsweb passages are not necessarily designed to be more difficult later in the school year (Howe & Shinn, 2002). If they were, we would expect the increased complexity of text to result in a processing bottleneck for poor decoders whereby their inability to decode accurately and fluently prohibits accurate comprehension. In fact, the poor decoders in our study improved their scores on the word reading fluency measure, but not beyond the low average range. It is likely that these students have weaknesses in oral language skills that were not captured by either aimsweb subtest, which provides further support for why diagnostic assessment is a crucial follow-up to progress monitoring assessments.

Overall, it appears that Core5 usage contributed to increased word reading and comprehension performance on a distal measure at the end of the school year for the majority of students. For students with the most severe difficulties or mixed deficits, the technology-based instruction could not close the gap with typical peers (i.e., increase accuracy in the Standard mode of instruction). Of note, readers with mixed deficits did make significant gains on aimsweb from Fall to Spring. While we don't have a comparison condition, we know that all students in our sample received the same core curriculum and achieved different outcomes. Future work should examine the proximal and distal effects of technology-based interventions for reader profiles under controlled, experimental design conditions.

Conclusion

We conclude that a single progress monitoring tool, such as aimsweb, can be used to identify reader profiles mostly in line with past studies. The primary limitation of aimsweb is that it does not directly measure language ability, which would more accurately capture reader profiles and better guide instruction. Lexia Core5 Reading was able to effectively differentiate instruction for the majority of reader profiles; however, some students, especially those with mixed deficits, may require additional diagnostic assessment and teacher-led instruction to remediate reading difficulties. Finally, reader profiles generally maintained or increased average percentile scores on aimsweb subtests from Fall to Spring. In combination, these findings provide evidence that efficient identification and personalized intervention are possible for all students with reading difficulties.

Appendix

See Table 2.

Table 2 Descriptive statistics for the ustandardized residuals of each Core5 performance variable by profile, ranked and then regressed on Fall Level

Fall variables	Profile	N	Mean	SD	Median	Q1	Q3	IQR	Min	Max
Word reading activities	Accuracy in the standard mode of instruction									
	Mixed deficit	248	−54.41	120.95	−64.98	−149.35	35.01	184.36	−335.45	237.84
	Poor comprehenders	39	20.96	110.73	23.32	−39.49	112.67	152.16	−246.33	204.37
	Poor decoders	35	−2.07	129.93	−15.63	−100.86	118.71	219.57	−224.63	259.07
	Typical readers	161	60.21	125.23	92.37	26.98	150.32	123.34	−336.68	274.28
	Time in the guided practice mode									
	Mixed deficit	248	26.84	110.68	19.67	−59.32	120.52	179.84	−255.15	292.72
	Poor comprehenders	39	3.46	108.77	−1.33	−87.04	67.89	154.93	−155.33	255.7
	Poor decoders	35	21.95	108.68	14.76	−37.11	96.71	133.82	−217.68	257.7
	Typical readers	161	−28.83	84.19	−42.98	−92.24	26.7	118.94	−233.08	261.72
	Time in the direct instruction mode									
	Mixed deficit	248	29.28	111.01	35.33	−44.69	111.2	155.88	−328.85	278.85
	Poor comprehenders	39	−7.71	94.22	−37.58	−70.1	61.19	131.29	−199.74	188.85
	Poor decoders	35	3.74	116.93	−2.72	−70.02	94.58	164.59	−239.01	220.89
	Typical readers	161	−38.87	83.89	−59.19	−82.19	7.42	89.61	−260.85	261.48
Comprehension activities	Accuracy in the standard mode of instruction									
	Mixed deficit	248	−30.49	133.33	−31.06	−120.24	63	183.24	−382.29	290.28
	Poor comprehenders	39	−11.01	126.29	−11.41	−107.29	86.44	193.73	−227.86	208.59
	Poor decoders	35	6.9	103.49	15.3	−80.83	78.93	159.76	−197.72	169.3
	Typical readers	161	56.11	128.36	48.57	−47.43	162.57	210	−316.91	367.13
	Time in the guided practice mode									
	Mixed deficit	248	19.43	146.89	31.43	−100.91	143.15	244.06	−281.56	323.21
	Poor comprehenders	39	12.36	149.93	18.88	−112.46	155.16	267.62	−249.21	243.99
	Poor decoders	35	12.82	145.36	13.44	−87.96	124.62	212.58	−306.99	280.53
	Typical readers	161	−33.21	157.13	−34.01	−167.85	98.15	266	−311.85	277.7
	Time in the direct instruction mode									
	Mixed deficit	248	25.13	150.89	42.18	−100	156.11	256.11	−265.25	294.38
	Poor comprehenders	39	12.39	155.66	4.72	−95.06	155.23	250.29	−259.65	263.9
	Poor decoders	35	−8.4	149.98	12.9	−128.94	88.31	217.25	−259.65	290.06
	Typical readers	161	−47.14	152.39	−68.02	−176.1	73.9	250	−288.28	245.98

Table 2 (continued)

Spring variables	Profile	N	Mean	SD	Median	Q1	Q3	IQR	Min	Max
Word reading activities	Accuracy in the standard mode of instruction									
	Mixed deficit	223	−59.21	116.14	−71.24	−142.75	19.48	162.23	−322.89	245.48
	Poor comprehenders	30	−11.61	135.06	−0.25	−89.49	56.42	145.91	−265.46	233.01
	Poor decoders	27	−12.8	118.94	19.01	−123.56	59.51	183.07	−217.99	199.01
	Typical readers	116	29.8	127.5	42.98	−52.43	126.04	178.46	−335.51	259.44
Time in the guided practice mode	Mixed deficit	223	42.44	95.88	57.14	−20.42	112.35	132.77	−264.86	252.25
	Poor comprehenders	30	13.12	128.31	12.59	−54.01	102.75	156.76	−244.66	263.25
	Poor decoders	27	13.63	96.53	5.25	−30.15	87.79	117.94	−199.86	195.14
	Typical readers	116	−16.83	107.38	−9.36	−87.24	57.66	144.9	−268.66	326.94
Time in the direct instruction mode	Mixed deficit	223	22.23	92.38	34.36	−44.49	84.39	128.88	−220.25	213.12
	Poor comprehenders	30	−18.84	112.65	−30.75	−87.98	61.89	149.87	−245.35	202.12
	Poor decoders	27	−15.94	91.51	−19.64	−83.86	41.39	125.24	−199.35	181.75
	Typical readers	116	−37.44	89.02	−42.61	−95.86	20.89	116.75	−233.35	312.61
Comprehension activities	Accuracy in the standard mode of instruction									
	Mixed deficit	223	−31.06	166.06	−35.36	−175.25	119.19	294.43	−377.84	338.75
	Poor comprehenders	30	−2.09	155.58	−7.52	−129.21	137.4	266.61	−239.03	263.75
	Poor decoders	27	−27.55	143.8	−22.99	−125.62	89.49	215.11	−313.99	230.97
	Typical readers	116	52.91	139.93	60.15	−59.77	176.36	236.13	−266.46	291.59
Time in the guided practice mode	Mixed deficit	223	12.86	172.07	10.66	−142.65	181.03	323.68	−308.43	302.07
	Poor comprehenders	30	5.85	168.87	−6.15	−134.17	139.89	274.07	−245.34	295.07
	Poor decoders	27	30.46	160.59	66.57	−69.58	150.03	219.61	−285.34	259.35
	Typical readers	116	−1.34	139.34	−4.19	−105.02	109.19	214.21	−321.96	276.66
Time in the direct instruction mode	Mixed deficit	223	15.87	171.97	15.34	−123.61	183.39	307	−288.52	286.99
	Poor comprehenders	30	−2.77	170.55	21.95	−157.61	129.08	286.69	−256.26	281.99
	Poor decoders	27	30.95	151.82	62.78	−68.61	148.94	217.55	−258.11	260
	Typical readers	116	−26.13	136.07	−29.31	−140.97	77.2	218.17	−284.48	268.99

References

Aaron, P. G., Joshi, M., & Williams, K. A. (1999). Not all reading disabilities are alike. *Journal of Learning Disabilities, 32*(2), 120–137. https://doi.org/10.1177/002221949903200203.

Adlof, S. M., & Catts, H. W. (2015). Morphosyntax in poor comprehenders. *Reading and Writing, 28,* 1051–1070. https://doi.org/10.1007/s11145-015-9562-3.

Adlof, S. M., Catts, H. W., & Little, T. D. (2006). Should the simple view of reading include a fluency component? *Reading and Writing, 19,* 933–958. https://doi.org/10.1007/s11145-006-9024-z.

Ardoin, S. P., & Christ, T. J. (2009). Curriculum-based measurement of oral reading: Standard errors associated with progress monitoring outcomes from DIBELS, AIMSweb, and an experimental passage set. *School Psychology Review, 38*(2), 266–283.

Catts, H. W., Hogan, T., & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities*, *36*(2), 151–164. https://doi.org/10.1177/002221940303600208

Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In H. W. Catts & A. G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 25–40). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Cain, K. (2015). Literacy development: The interdependent roles of oral language and reading comprehension. In R. H. Bahr & E. R. Silliman (Eds.), *Routledge handbook of communication disorders* (pp. 204–214). London: Routledge.

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49*(2), 278–293. https://doi.org/10.1044/1092-4388(2006/023).

Chall, J. S. (1996). *Stages of reading development*. Fort Worth, TX: Harcourt Brace College Publishers.

Connor, C. M., Phillips, B. M., Kim, Y. S. G., Lonigan, C. J., Kaschak, M. P., Crowe, E., … & Al Otaiba, S. (2018). Examining the efficacy of targeted component interventions on language and literacy for third and fourth graders who are at risk of comprehension difficulties. *Scientific Studies of Reading*, *22*(6), 462–484.

Cutting, L. E., Materek, A., Cole, C. A., Levine, T. M., & Mahone, E. M. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Annals of dyslexia, 59*(1), 34–54. https://doi.org/10.1007/s11881-009-0022-0.

Dwyer, C. A. (1996). Cut scores and testing: Statistics, judgment, truth, and error. *Psychological Assessment, 8*(4), 360.

Elwér, Å., Keenan, J. M., Olson, R. K., Byrne, B., & Samuelsson, S. (2013). Longitudinal stability and predictors of poor oral comprehenders and poor decoders. *Journal of Experimental Child Psychology, 115*(3), 497–516. https://doi.org/10.1016/j.jecp.2012.12.001.

Foorman, B. R., Breier, J. I., & Fletcher, J. M. (2003). Interventions aimed at improving reading success: An evidence-based approach. *Developmental Neuropsychology, 24*(2–3), 613–639. https://doi.org/10.1207/S15326942DN242&3_06.

Foorman, B. R., Petscher, Y., Stanley, C., & Truckenmiller, A. (2018). Latent profiles of reading and language and their association with standardized reading outcomes in kindergarten through tenth grade. *Journal of Research on Educational Effectiveness, 10*(3), 619–645. https://doi.org/10.1080/19345747.2016.1237597.

Gersten, R., Compton, D., Connor, C.M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2008). *Assisting students struggling with reading: Response to Intervention and multi-tier intervention for reading in the primary grades. A practice guide.* (NCEE 2009-4045). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6–10.

Hagtvet, B. E. (2003). Listening comprehension and reading comprehension in poor decoders: Evidence for the importance of syntactic and semantic skills as well as phonological skills. *Reading and Writing, 16,* 505–539. https://doi.org/10.1023/A:1025521722900.

Hendricks, A. E,Adlof, S. M.,Alonzo, C. N.,Fox, A. B., & Hogan, T. P. (2019). Identifying children at risk for developmental language disorder using a brief, whole-classroom screen. *Journal of Speech, Language, and Hearing Research* 62(4), 896–908. https://doi.org/10.1044/2018_JSLHR-L-18-0093

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

Howe, K. B., & Shinn, M. M. (2002). *Standard reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features*. Eden Prairie, MN: Edformation.

Justice, L., Mashburn, A., & Petscher, Y. (2013). Very early language skills of fifth-grade poor comprehenders. *Journal of Research in Reading, 36*(2), 172–185. https://doi.org/10.1111/j.1467-9817.2011.01498.x.

Kazakoff, E. R., Macaruso, P., & Hook, P. (2018). Efficacy of a blended learning approach to elementary school reading instruction for students who are English learners. *Educational Technology Research and Development, 66*(2), 429–449. https://doi.org/10.1007/s11423-017-9565-7

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12*(3), 281–300. https://doi.org/10.1080/10888430802132279.

LARRC (2015). Learning to read: Should we keep things simple? *Reading Research Quarterly*, *50*(2), 151–169. http://doi.org/10.1002/rrq.99

Language and Reading Research Consortium (LARRC), & Chiu, Y. D. (2018). The simple view of reading across development: Prediction of grade 3 reading comprehension from prekindergarten skills. *Remedial and Special Education, 39*(5), 289–303.

Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology, 95*(2), 211–224. https://doi.org/10.1037/0022-0663.95.2.211.

Lerkkanen, M.-K., Rasku-Puttonen, H., Aunola, K., & Nurmi, J.-E. (2004). Reading performance and its developmental trajectories during the first and the second grade. *Learning and Instruction, 14*(2), 111–130. https://doi.org/10.1016/j.learninstruc.2004.01.006.

Mitchell, A., Baron, L. & Macaruso, P. (2018). Assessment without testing: Using performance measures embedded in a technology-based instructional program as indicators of reading ability. *Journal of Educational Multimedia and Hypermedia, 27*(2), 179–192.

Muijselaar, M. M. L., Kendeou, P., de Jong, P. F., & van den Broek, P. W. (2017). What does the CBM-Maze test measure? *Scientific Studies of Reading*, *0*(0), 1–13. http://doi.org/10.1080/10888438.2016.1263994

Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language, and Hearing Research, 47,* 199–211. https://doi.org/10.1044/1092-4388(2004/017).

National Center on Intensive Intervention. (2013). *Data-based individualization: A framework for intensive intervention*. Washington, DC: Office of Special Education Programs, U.S. Department of Education.

National Reading Panel (US), National Institute of Child Health, & Human Development (US). (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 Institutes of Health: National Institute of Child Health and Human Development.

NCS Pearson. (2012). aimsweb: Technical Manual. Bloomington, MN: NCS Pearson. Retrieved from http://www.aimsweb.com/wp-content/uploads/aimsweb-Technical-Manual.pdf.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and cognitive processes, 18*(4), 443–468.

Prescott, J. E., Bundschuh, K., Kazakoff, E. R., & Macaruso, P. (2017). Elementary school–wide implementation of a blended learning program for reading intervention. *The Journal of Educational Research*. https://doi.org/10.1080/00220671.2017.1302914.

Petscher, Y., Justice, L. M., & Hogan, T. P. (2018). Modeling the early language trajectory of language development when the measures change and its relation to poor reading comprehension. *Child development*, *89*(6), 2136–2156. https://doi.org/10.1111/cdev.12880

Quade, D. (1967). Rank analysis of covariance. *Journal of the American Statistical Association, 62*(320), 1187–1200.

Richardson, J. T. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review, 6*(2), 135–147. https://doi.org/10.1016/j.edurev.2010.12.001.

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*(3–4), 183–200. https://doi.org/10.1080/07380569.2015.1100652

Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., Brady, S., et al. (1999). Comprehension and decoding: Patterns of association in children with reading difficulties. *Scientific Studies of Reading, 3*(1), 69–94. https://doi.org/10.1207/s1532799xssr0301.

Staker, H., & Horn, M. B. (2012). *Classifying K-12 blended learning*. Mountain View, CA: Innosight Institute. https://doi.org/10.1007/s10639-007-9037-5.

Subban, P. (2006). Differentiated instruction: A research basis. *International Education Journal, 7*(7), 935–947.

Tomlinson, C. A. (2000). *Differentiation of instruction in the elementary grades*. Champaign, IL: Clearinghouse on Elementary and Early Childhood Education. https://eric.ed.gov/?id=ED443572.

Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., et al. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of the literature. *Journal for the Education of the Gifted, 27*(2–3), 119–145. https://doi.org/10.1177/016235320302700203.

Torppa, M., Tolvanen, A., Poikkeus, A. M., Eklund, K., Lerkkanen, M. K., Leskinen, E., et al. (2007). Reading development subtypes and their early characteristics. *Annals of Dyslexia, 57*(1), 3–32. https://doi.org/10.1007/s11881-007-0003-0.

van de Sande, E., Segers, E., & Verhoeven, L. (2016). Supporting executive functions during children's preliteracy learning with the computer. *Journal of Computer Assisted Learning, 32*(5), 468–480.

Affiliations

Lauren S. Baron[1,2,3] ⬤ · **Tiffany P. Hogan**[1] ⬤ · **Rachel L. Schechter**[2,4] · **Pamela E. Hook**[2] · **Elizabeth C. Brooke**[2]

[1] Department of Communication Sciences and Disorders, MGH Institute of Health Professions, Charlestown Navy Yard, 36 1st Ave, Boston, MA 02129, USA

[2] Lexia® Learning, A Rosetta Stone® Company, Concord, USA

[3] Present Address: Department of Communication Sciences and Disorders, University of South Carolina, Columbia, USA

[4] Present Address: Houghton Mifflin Harcourt, Boston, USA