Comparing Istation and easyCBM as Progress-Monitoring Tools in a Rural School in New Mexico

Sunaina Shenoy, Theresah Boateng, and Allison Nannemann
University of New Mexico

Abstract

This study compared students' performance on Istation, a computer-based progress-monitoring tool, and easyCBM, a paper-pencil progress-monitoring tool. Participants included 106 students in Grades K-6 from a rural school in New Mexico. Teachers in these grade levels administered both assessments three times during the school year. Results suggested that for measuring progress in reading, easyCBM was a preferred tool for students in kindergarten whereas Istation was a preferred tool for students in Grade 3. For students in Grade 5, easyCBM was preferred for measuring math and Istation was preferred for measuring reading. For the other grades, we observed no significant differences in students' scores on Istation and easyCBM for both reading and math measures. The measure used directly impacts intervention for students in reading and math at all levels, both in general and special education. Implications for research and practice within this context are discussed.

Keywords: Curriculum-based measures, Istation, easyCBM, rural, New Mexico

Curriculum-based measurement (CBM; Greenwood & Kim, 2012; Jin et al., 2015; Kendeou & Papadopoulous, 2012; Kim et al., 2016) is an alternative to standardized testing that has rich empirical support. It consists of brief measures of an academic construct such as reading, writing, or mathematics that can be administered repeatedly by the classroom teacher (Deno, 2003; Reschly et al., 2009; Tindal, 2013). Unlike other formative assessments, CBMs are backed by robust validity and reliability data and can be used to guide educational decisions by comparing student performance over time as well as to performance benchmarks (Miuara-Wayman et al., 2007).

CBMs serve several key purposes in a classroom. First, at the micro level, they help teachers with universal screening and early identification of students who might be at risk for learning disabilities (Fletcher & Vaughn, 2009; Fuchs & Fuchs, 2017; Sugai & Horner, 2009). Second, they help teachers plan interventions for students who are struggling to acquire grade-level standards in reading and math (Fletcher & Vaughn, 2009). Third, they provide teachers with an overview of what seems to be working in their classrooms as well as where they might need to focus their attention in whole- and small-group instructional planning (Fuchs & Fuchs, 2017). Fourth, they provide teachers with critical information on how to differentiate instruction (Gartland & Strosnider, 2020). Moreover, at the macro level, CBMs provide school districts with data on student progress in grade-level reading and math standards that can be compared to national norms, which, in turn, helps with targeted professional development, financial planning, and data-based decision-making.

In the following, we will look at two such tools, easyCBM (Anderson et al., 2014) and Istation (Istation, 2024: https://www.istation.com/).

easyCBM

easyCBM is a curriculum-based measure developed and revised by researchers at the University of Oregon to measure students' grade-level progress

Shenoy, S., Boateng, T., & Nannemann, A. (2024). Comparing Istation and easyCBM as Progress-Monitoring Tools in a Rural School in New Mexico. *International Journal for Research in Learning Disabilities, 7*(1), 14. https://doi.org/10.28987/ijrld.7.1.14

in reading and math (Anderson et al., 2014) with a particular focus on facilitating "data-driven instructional decision making through enhanced reporting options" (Anderson et al., 2014, p. 4) to promote progress monitoring and universal screening in schools (Deno, 2003; Keller-Margulis et al., 2008).

Several factors contributed to the development of easyCBM: (a) an adoption of the Common Core State Standards (CCSS; www.corestandards.org); and (b) a decision to embrace all learners, including typical students, students with disabilities, and English language learners (ELLs). Together, the response-to-intervention (RTI) model and the data-based features of easyCBM offer a tool for effective decision-making about students' reading and math progress overtime (Alonzo et al., 2012).

The reading assessment, consisting of test items and a link to content standards, was developed by a group of teachers with many years of experience working with students struggling with reading (Alonzo et al., 2012). Further, a research team of 3–5 members reviewed the items, using universal design for assessment as a guideline (Alonzo et al., 2012). A convenience sampling of students with and without disabilities and ELLs across different states participated in the item piloting phase. Data analyses were performed using item response theory, and 10 tests were created for formal progress monitoring and three for benchmark screening. The reading measures included Letter Names, Letter Sounds, Phoneme Segmenting, Word Reading Fluency, Passage Reading Fluency, Vocabulary, Basic Reading, Proficient Reading, Syllable Sounds, Syllable Segmenting, Word Reading Fluency, Sentence Reading Fluency and Vocabulary. The measures were developed in both English and Spanish (Alonzo et al., 2012).

The math assessment was developed to replicate the National Council of Teachers of Mathematics Focal Point Standards (Alonzo et al., 2012). Sixteen items were developed by mathematics teachers across grade levels K-8 to align with the CCSS, for the purpose of measuring students' progress every 3–4 weeks. The math measures included Numbers and Operations, Algebra, Geometry, and Measurement (Alonzo et al., 2012).

Measures of reliability of the easyCBM from an alternate form and test-retest reliability of student participants from Grades 1, 3, 5, and 8 were conducted in a school in the Pacific Northwestern part of the United States. The results of the alternate form showed the following outcome for Grade 1: Phoneme Segmenting (.86 to .91), Letter Names (.82 to .89), Word Reading Fluency (.95 to .96), and Passage Reading Fluency (.95 to .97). Similar results were reported for Grades 3 and 5. Grade 3 had co-efficient ranging from (.87 to .93) with a strong correlation in Passage Reading Fluency for Grade 3 (.94 to .95) and a strong correlation in Passage Reading Fluency for Grade 5 (.87 to .96). Grade 8 results also showed a strong correlation in Passage Reading Fluency (.87 to .95). The alternate form showed a weaker correlation with comprehension measure $R = .59$ for Grade 5 and $R = .35$ for Grade 8 (Alonzo & Tindal, 2009).

easyCBM was developed and integrated with a Universal Design for Assessment component and as a result reflects general applicability to a wide range of diverse learners. The assessment has been implemented in all U.S. states and across school districts primarily due to its alignment with the CCSS (Alonzo & Tindal, 2009).

Istation

Istation, a computer-based program, is becoming the most widely used progress-monitoring tool in the United States, currently adopted by approximately 24 states, serving 4 million students across (https://www.istation.com/). The Istation Early Reading (ISIP Early Reading) is designed for year-round continuous CBM progress monitoring of students' reading abilities in critical targeted domains (phonemic awareness, alphabetic knowledge and skills, fluency, vocabulary, and comprehension) using a computer adaptive testing mechanism.

A data-tracking system, Istation supports progress monitoring, analysis, and provision of immediate feedback in the form of web-based reports that inform about students' reading and math strengths and challenges accompanied by recommendations to increase student performance. It was developed to replicate the features of the CCSS and tailored to reflect the core principle of the *Elementary and Secondary Education Act* and *No Child Left Behind,* which is to focus on the progress of disadvantaged students (Mathes et al., 2016). Istation tests are available in English and Spanish.

The Istation Early Reading (ISIP Early Reading) is designed for pre-kindergarten through Grade 3 and includes phonemic awareness, alphabetic knowledge and skills, connected text fluency, vocabulary, and comprehension (Mathes et al., 2016). ISIP Advanced Reading is designed for students in Grades 4–8 and addresses word analysis, text fluency, vocabulary, and comprehension (Mathes, 2016). Finally, the math domain is comprised of ISIP Early Math (Pre-K to kindergarten) and ISIP Math for students in Grades 2–8. The math subtests include number sense, operations, algebra, geometry, measurement, and data analysis (Mathes et al., 2016).

Evidence for the validity of Istation was reported in 2008–2009 based on its application in a North Texas school district in Pre-K classes at five elementary schools. Participants spanned a diverse group of students, including "different ethnic groups (African American, Asian, Hispanic, Pacific Islander, White, Other), students qualifying for Free/Reduced Lunch, students receiving ESL Services, in bilingual classrooms, English Language Leaners (ELL), students with disabilities, and receiving Special Education Services" (Mathes, 2009, p. 5). The tests were administered monthly to students according to a set schedule, and the test items matched students' grade levels.

Based on data collated from kindergarten through Grade 3 during the school year, reliability evidence was tested with Cronbach's coefficient alpha and Pearson product-moment correlation coefficients. The results showed a 0.927 to 0.970 ($N = 416$) reading ability from the testing conducted from October to February. A test of validity was established from a content validity analysis. Two professional experts in early reading and a panel of specialists incorporated rigorous development of the content areas, a review process, and several revisions to establish accuracy across content and corresponding measurements of the test items (Mathes, 2009).

Several studies have reported on the ability of Istation scores to accurately predict reading and math outcomes (Campbell et al., 2018; Cook & Ross, 2021; Ford, 2014; Montelione, 2021; Wolf et al., 2022). Additionally, May et al. (2018) described the efficacy of the Istation reading inventory used in Idaho. Specifically, 32% of teachers reported that Istation did an excellent job, 51% reported that it did an adequate job, and 16% felt it performed poorly in matching their observations of students' overall performance in class; the results were similar for reading level, with 32% of teachers reporting that it was excellent, 45% reporting that it was adequate, and 20% reporting that it was poor.

Preferences for Paper-Pencil vs. Computer-Based Testing Modalities

Recently, there has been a move away from paper-pencil tests such as the easyCBM in favor of computer-based assessments such as the Istation, primarily based on the belief that technology provides teachers with a more convenient way to test all students at once without relying on individual administration time, and provides more valid and reliable results than paper-pencil tests (Wolf et al., 2022).

Wang et al. (2008) conducted a meta-analysis to study the differences between computer-based and paper-pencil tests on student reading and math achievement scores across Grades K-12. They found no significant differences based on the modality of testing for either reading or math outcomes across all grade levels. Similarly, Kingston (2008) studied the applicability of paper-based vs. computer-based measures across students in K-12 settings and found no grade- or subject-level differences; however, computer administration provided a slight advantage for students in English language arts and social studies whereas paper-pencil provided a small advantage for math. Chua (2012) found that computer-based tests were more reliable in terms of internal and external validity, significantly reduced testing time, and were preferred by both students and teachers over paper-pencil tests. Furthermore, the Istation manual (Mathes et al., 2016) reports correlations between Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2009), a paper-pencil test, and Istation scores, and overall moderate correlations ranging from 0.17–0.83 across subtests in Grades K-3, with smaller correlations reported in kindergarten and higher correlations in Grade 3.

Choi and Tinkler (2002) evaluated the two modalities in K-12 settings; several of their findings are as follows:

- Item difficulty estimates were greater for third graders than tenth graders; this effect was larger on reading tests than math tests; and items presented on the computer were more difficult than items presented in paper-pencil format.
- Across all grade levels, identical items administered using the two modalities were perceived by students as being more difficult on the computer version of the test than the paper-pencil version, and this effect was greater for reading than math tests.
- Taking tests online provided more novelty to younger than older students, and having to scroll through long reading comprehension passages had a negative effect on reading achievement scores, especially for younger students.

Gaps in Assessment Research

Despite the research mentioned above on the use of individual progress-monitoring assessments and comparisons of multiple assessment formats, there are some gaps in the literature worth noting. First, most of the existing studies were conducted in urban parts of the US, making generalization of findings across settings difficult, considering differences in context, student characteristics, and access to resources. Additionally, the body of research comparing the efficacy of computer-based and paper-pencil progress-monitoring tools is limited. Given that one format (computer-based) is replacing the other (Wolf et al., 2022) rather than leaving both options available, a more comprehensive body of research would be valuable.

Finally, both in research and in practice, assessments are chosen above the school level, diminishing the role of teachers' professional expertise and knowledge of students, especially in relation to students who come from culturally and linguistically diverse backgrounds and students who might have special education needs. Given that progress monitoring is designed to inform and evaluate classroom-level instruction, teachers would be expected to have valuable input on the assessments that are used.

In the current study, we used the principles of participatory action research (PAR) as a theoretical framework to address the aforementioned gaps in assessment research.

Current Study

Theoretical Framework

Emerging from the work of Freire (1974) and Habermas (1979), the basic premise of PAR is founded on the importance of creating knowledge that is a product of collective, self-reflective inquiry used to improve a situation (Street, 1995). According to Stringer (1996), PAR can be democratic, equitable, and life-enhancing. Further, the production of such knowledge is not only useful to groups of people but may also lead to empowerment through the reflective process of constructing this knowledge base (Reason, 1994). That is, participants are considered co-researchers in the process of knowledge construction, and they use this new knowledge to make improvements to their practice (Street, 1995).

The present study extended PAR by not only co-constructing knowledge with teachers and researchers but also building meaningful community-based research (CBR) partnerships between universities and schools as a model of social change (Barber, 1992; Brown, 2001; Kahne & Westheimer, 1996). According to Strand et al. (2003), community-based research traces back to Horton (1989), Hall (1992), and Lewin (1948), but a more recent understanding is that it is a collaboration between academic researchers and community members in which multiple methods of inquiry and knowledge acquisition are used for the purpose of achieving social justice. The role of the community member goes beyond identifying a research question, to include data collection and analysis, interpreting results, and implementing changes in practice (Strand et al., 2003). As such, it can become an important tool for learning, teaching, and empowerment and often takes several years to establish.

Context and Research Questions

The present study was an outcome of a university-school partnership based in New Mexico, a state located in the southwestern part of the United States. It is one of the poorest states in the country, with rural areas being especially poor. In the 2018–2019 school year, the participating public school district served 224 elementary school students (NCES, 2021). The district's overall average reading proficiency score was 25% compared to 29% statewide. Moreover, 30% of adults in the community have not received a high school diploma compared to 16% statewide; and 9% of adults are unemployed compared to 8% statewide (NCES, 2021). Finally, 75% of the population speaks a language other than English. The students in these schools come from various ethnic backgrounds, with 75% identifying as Hispanic, 22% identifying as Caucasian, and 3% identifying as Native American (NCES, 2021).

The principal at the school site invited the researchers to participate in a research study focused on implementing a tiered model of instruction for reading and math, with the goal of reducing the risk of misidentifying students with reading and math disabilities. This, in turn, led us to address two larger goals within our state: (a) complementing the multi-layered system of supports (MLSS) that was being piloted by the New Mexico Public Education Department (NM-PED) at the time of this study (NM-PED, 2021), and (b) providing students from low-income families, ELLs, and students with disabilities better educational outcomes, in response to the consolidated lawsuit *Yazzie/Martinez v. State of New Mexico* (New Mexico Center on Law and Poverty, n.d.).

The study was conducted during the 2021–2022 school year when students had just returned to in-person education after the COVID-19 lockdown. As a first step, we recommended bi-weekly professional learning communities (PLCs) with teachers in Grades K-6 that we hoped to serve through this project. PLCs are groups of teachers and school administrators who aim to improve their collective knowledge and practice to improve student outcomes (Hord, 1997).

Stoll et al. (2006) elaborate on this definition by detailing three essential aspects of PLCs. First, the learning undertaken by PLCs seeks to improve teaching practice, not perpetuate current practice. Second, the work of PLCs does not end with teacher learning; teachers are expected to utilize the new practices in their classrooms. Finally, the main purpose of engaging in PLCs is to promote greater student learning. Teachers largely agree that PLCs are an effective mechanism for improving educational practice given their student-centered focus on collaborative goal setting and problem solving (Farley-Ripple & Buttram, 2014; Hudson et al., 2013; Vescio et al., 2008).

As part of the bi-weekly professional learning communities in the current study, the teachers raised concerns about the use of Istation for lower-elementary grades, especially for students in kindergarten, as they found that it was not capturing students' "true" scores in reading and math for two reasons: (a) students were not familiar with the online format and were still learning basic computer functions like using a mouse, which decreased their reaction time in answering questions on timed tests; and (b) students had experienced a learning loss during the COVID-19 lockdown and were not meeting grade-level standards. The teachers believed the students felt intimidated by the computer-based assessments and would respond better to a one-on-one paper-pencil assessment.

In response to these concerns, we developed this study to compare student performance on the Istation, a computer-based test, with their performance on the easyCBM, a paper-pencil test. Our research questions sought to address the differences in student scores between the two measures:

1. Is there a correlation between the easyCBM and Istation reading and math tests across Grades K-6 in a rural school site in New Mexico?
2. Are there individual differences between students' reading and math scores on the Istation and the easyCBM across Grades K-6 in a rural school site in New Mexico?

Our choice of the above research questions was grounded in the literature review presented earlier and is a direct response to the teacher participants as part of the PAR model that we had set up at the school site to implement MLSS with fidelity. Given the research literature, we hypothesized that there would be a high degree of correlation between the easyCBM and the Istation, considering the high levels of reliability between paper-pencil and computer-based tests established in the literature. Moreover, we did not expect to find any differences between the two tests across grade levels.

Methods

Participants

Participants included (a) 7 grade-level teachers representing each grade in our sample; (b) the principal of the school; (c) 2 special educators/interventionists; and (d) 106 students from Grades K-6 (K: 17; Grade 1: 7; Grade 2: 16; Grade 3: 12; Grade 4: 16; Grade 5: 25; and Grade 6: 13). Of the 106 students, 63 (59.43%) were females, 6 (5.66%) were diagnosed with a disability and received special education services, and 9 (8.49%) were ELLs.

The school principal initiated the PAR partnership to support the selection and implementation of effective practices at the school. Through this partnership, the teachers were responsible for identifying specific needs, including the need to consider alternative assessment tools that would be more responsive to the unique population of their school. The teachers also provided information on the school community and student body that helped the university partners to identify alternative assessments.

School Setting

Our school site was in a district located in the rural part of northern New Mexico. During our first PLC meeting with teachers, they mentioned that many students in the school did not have access to laptops, computers, and internet in this remote part of the state and were starting the school year with a significant learning loss. The teachers had previously been trained in the Wilson Foundations reading curriculum but had not yet received all the grade-level materials. They were currently being trained in the LETRS reading curriculum (Moats & Sedita, 2004). For math, the teachers had access to Eureka Math (https://greatminds.org/math/eurekamath) and were trained in using Zearn (Knudsen et al., 2020; www.zearn.org).

Measures

Table 1 lists the description, administration time, and grade levels for the subtests that were administered in reading and math.

Data Collection

Teachers were instructed in how to administer the easyCBM subtests, and administered this test individually to all students in their classroom to measure grade-level skills in reading and math. easyCBM was administered three times during the 2021–2022

Table 1
Measures Administered Across Grade Levels

Name of Test	Name of Subtest	Description/Example	Administration Time	Grade Levels
easyCBM Reading	Letter Names	Identify lowercase and uppercase letters	1 minute	K-1
	Phoneme Segmenting	Breaking apart letter sounds (e.g., /c/a/t/)	1 minute	K-1
	Letter Sounds	Producing sounds of letters	1 minute	K-1
	Word Reading Fluency	Read decodable and sight-words (left to right)	1 minute	K-3
	Passage Reading Fluency	Read aloud short (250 words) narrative passage	1 minute	1-6
	Multiple Choice Reading Comprehension	Assesses comprehension of written texts	30 minutes	2-6
	Vocabulary	Assesses age-appropriate vocabulary proficiency	10-15 minutes	2-6
easyCBM Math	Numbers and Operations	Basic operations (addition, subtraction, multiplication, division)	8-15 minutes	K-6
	Geometry	Varies in complexity (example: identifying shapes to calculation of perimeter)	8-15 minutes	K, 1, 3
	Measurement	Assesses various aspects of measurement (calculation of area, telling time, etc.)	8-15 minutes	K, 2, 4
	Number and Operations, and Algebra	Assesses basic algebraic knowledge (addition, subtraction, multiplication, division)	8-15 minutes	1-4
	Geometry, Measurement, Algebra	Assesses geometry, measurement, and algebraic concepts	8-15 minutes	5, 7
	Algebra	Basic and complex algebraic equations and word problems	8-15 minutes	6
	Number and Operations, Ratios	Assesses numbers and operations, ratios	8-15 minutes	6
Istation Reading	Phonics	Assesses letter knowledge, alphabetic decoding, spelling	3-10 minutes	PreK-3
	Phonological Awareness	Spoken words as components of individual sounds	3-10 minutes	PreK-1
	Vocabulary	Assesses vocabulary	3-10 minutes	PreK- 6
	Comprehension	Assesses listening comprehension, reading comprehension	3-10 minutes	PreK- 6
	Fluency	Assesses text fluency, connected text fluency	3-10 minutes	1 (after evidence of alphabetic decoding) 2-6
	Word Analysis	Spelling and orthographic processing		4-6
	Connected Text Fluency	Assesses students on passages of equivalent difficulty to measure growth over time		4-6
Istation Math	Number Sense	Assesses number representation, number systems, and counting sequences	30 minutes	2-6
	Operations	Comprehension of mathematical operations, concepts, and relations	30 minutes	2-6
	Algebra	Comprehension of statements of relations, mathematical symbols, and rules	30 minutes	2-6
	Geometry	Assesses foundational skills such as describing shapes and angles, recognizing patterns and measurement	30 minutes	2-6
	Measurement	Assesses understanding of measurement concepts such as metric vs. customary measuring systems	30 minutes	2-6
	Data Analysis	Assesses the ability to form and evaluate numerical inferences to support accurate conclusions	30 minutes	2-6

academic year, in fall (August-September), winter (November-December), and spring (April-May). It measured reading and math subskills per grade level, including phonemic awareness, phonics, fluency, vocabulary, comprehension, operations, geometry, algebra, and measurement. It was administered 2–3 weeks after the district-mandated Istation measure to get a comparable score.

The school was already administering the Istation as a whole group per grade level three times during the school year, fall, winter, and spring, to measure beginning-of-the-year, middle-of-the-year, and end-of-the-year reading and math outcomes. They continued this for the 2021–2022 school year and gave us access to the data. The constructs measured on both easyCBM and Istation were comparable per grade level.

Data Analysis

Our data analysis included calculating composite percentiles for reading and math for the three administration time points. We used percentiles instead of raw scores because some subtests were administered during one time point but not the other, as prescribed by the easyCBM or Istation manuals, and percentiles were more valid for the purposes of making comparisons between scores. We then used these data to run correlations between the tests to address the first research question.

Finally, we used paired samples *t*-tests to measure individual differences between the two measures to address the second research question. This was an appropriate statistical measure for our purposes because in this type of test, each subject or entity is measured twice, resulting in pairs of observations, which was the case in our sample, where the same student was administered both easyCBM and Istation. Through this analysis, we were interested in measuring whether the mean difference between two sets of observations was zero. We ran the analyses using the SPSS software.

Results

Table 2 presents the correlation between easyCBM and Istation across grade levels and test administration periods. The average correlation between easyCBM and Istation across grade levels was 0.73 for reading and 0.45 for math, showing a strong correlation between the two tests on reading but a moderate correlation on math.

Tables 3 and 4 present the *t*-test comparisons of Istation and easyCBM composites for reading and math across grades K-6. For students in kindergarten, we found statistically significant differences between their reading scores on Istation and easyCBM during the fall and winter administration of the tests, but not the spring administration. Students performed better on the easyCBM during all three administration points. We did not find any significant differences between math scores on either test; however, they did score better on the easyCBM test than the Istation test.

For students in Grade 1, we did not find any statistically significant differences between reading and math scores; however, in general, they scored better on Istation for reading and better on easyCBM for math.

For students in Grade 2, we did not find any statistically significant differences between reading and math scores; however, in general, they scored better on the Istation test for both reading and math.

For students in Grade 3, we found a statistically significant difference between their reading scores on Istation and easyCBM during all three testing periods, with their performance on Istation surpassing their performance on easyCBM. We did not find any significant differences between their math scores on either test, though they did perform better on Istation when compared to easyCBM.

For students in Grade 4, we did not find any statistically significant differences between reading and math scores; however, in general, they scored better on the easyCBM test for reading and math.

For students in Grade 5, we found a statistically significant difference between their reading scores on Istation and easyCBM during fall and spring but not winter, with better scores recorded on Istation than easyCBM. We also found a statistically significant difference between their math scores on Istation and easyCBM during the winter and spring administrations of the tests and not the fall administration, with better performances recorded on easyCBM.

For students in Grade 6, we did not find any statistically significant differences between reading and math scores; however, in general, they scored better on easyCBM for reading and on Istation for math.

Discussion

This study was a product of teacher-guided research questions as part of the participatory action research framework that guided this work (Street, 1995). The goals of this project were centered around teacher voice and empowerment through answering the research questions, reminiscent of Reason's (1994) work, which showed how reflective practices led to teacher knowledge and empowerment. As mentioned, the questions arose at our bi-weekly pro-

Table 2
Correlation of easyCBM and Istation Across Administration Periods for Grades K-6

Grades	Administration Time Points	Correlation Coefficient Reading	Correlation Coefficient Math
Grade K	Fall	0.72	0.39
	Winter	0.80	0.33
	Spring	0.79	0.36
Grade 1	Fall	0.98	0.73
	Winter	0.86	0.69
	Spring	0.99	0.86
Grade 2	Fall	0.55	0.23
	Winter	0.63	0.11
	Spring	0.78	0.26
Grade 3	Fall	0.77	0.15
	Winter	0.67	0.14
	Spring	0.75	0.18
Grade 4	Fall	0.61	0.41
	Winter	0.49	0.58
	Spring	0.82	0.71
Grade 5	Fall	0.79	0.32
	Winter	0.62	0.40
	Spring	0.74	0.75
Grade 6	Fall	0.49	0.26
	Winter	0.75	0.76
	Spring	0.71	0.77

fessional learning communities, which focused on reading and math assessment and intervention.

Thus, our research questions were a product of teacher inquiry to gain more knowledge and better serve a community of learners. Teachers were concerned that computer-based tests were not capturing the true scores of students, especially in lower elementary grades, so we decided to test students in both computer-based and paper-pencil tests and compare the findings. Looking across grades, both progress-monitoring tools were strongly correlated for reading and moderately correlated for math. easyCBM was the preferred test to measure reading in kindergarten and math in Grade 5. Istation was the preferred test to measure reading in Grades 3 and 5. For the other grades, there were no statistically significant differences in scores between the two tests, and we did not find a trend that differentiated lower and upper-elementary grades. In contrast to our findings, Mathes et al. (2016) found a weaker correlation with a paper-pencil and a computer-based measure in kindergarten and higher correlations as students moved through the elementary grades.

Implications for Research and Practice

In terms of practice, we recommended that teachers in the school use the easyCBM test to measure reading in kindergarten, especially as these students learn how to navigate computer-based tests because, in general, they scored better on this test compared to Istation. We also recommended using easyCBM to measure math in Grade 5. However, for Grades 1–4, we would recommend that the teachers continue to use Istation, which is state-mandated. This directly impacts the reading and math interventions for students both in general education and special education classrooms. Moreover, it helps teachers make data-based decisions about how to group students based on their performance on reading and math assessments.

In terms of research, first, we were not able to find any comparative studies that included the specific pa-

Table 3
Comparison of Reading Composites on Istation and easyCBM for Grades K-6

Grade Levels	Data Collection Time Periods	Mean (*SD*) Istation	Mean (*SD*) easyCBM	*t*-statistic	*p*-value	Cohen's *d*
Kindergarten (*N* = 16)	Fall	42.25 (21.21)	52.25 (19.07)	2.65	0.018**	0.66
	Winter	47.33 (23.79)	58.87 (23.42)	3.01	0.009**	0.77
	Spring	52.13 (26.01)	58.88 (21.92)	1.69	0.112	0.42
Grade 1 (*N* = 7)	Fall	31.71 (37.71)	26.00 (33.18)	1.83	0.116	0.69
	Winter	36.71 (31.36)	39.43 (30.81)	0.44	0.676	0.17
	Spring	37.57 (32.17)	34.57 (26.78)	1.24	0.263	0.46
Grade 2 (*N* = 15)	Fall	39.93 (37.56)	36.13 (31.47)	0.44	0.665	0.11
	Winter	43.33 (32.42)	49.40 (26.73)	0.90	0.382	0.23
	Spring	43.67 (32.14)	46.33 (30.09)	0.44	0.669	0.11
Grade 3 (*N* = 10)	Fall	58.20 (22.94)	42.60 (20.16)	3.36	0.008**	1.06
	Winter	79.55 (18.93)	56.27 (24.32)	4.25	0.002**	1.28
	Spring	84.33 (19.9)	48.50 (23.42)	7.98	<.001***	2.30
Grade 4 (*N* = 13)	Fall	38.46 (32.24)	41.92 (19.32)	0.490	0.633	0.14
	Winter	46.82 (30.93)	52.55 (24.55)	0.67	0.519	0.20
	Spring	40.19 (30.04)	40.13 (26.00)	0.12	0.989	0.00
Grade 5 (*N* = 20)	Fall	36.85 (22.95)	25.65 16.57)	3.56	0.002**	0.80
	Winter	41.15 (20.56)	39.85 (18.28)	0.341	0.737	0.76
	Spring	30.39 (21.16)	37.35 (21.77)	2.15	0.43*	0.45
Grade 6 (*N* = 13)	Fall	43.54 (26.12)	48.15 (23.64)	0.66	0.524	0.18
	Winter	44.83 (23.16	49.92 (28.18)	0.95	0.363	0.27
	Spring	50.23 (21.12)	58.23 (18.44)	1.88	0.084	0.52

$^{*}p<0.05$. $^{**}p<0.01$. $^{***}p<0.001$.

Table 4
Comparison of Math Composites on Istation and easyCBM for Grades K-6

Grade Levels	Data Collection Time Periods	Mean (*SD*) Istation	Mean *(SD)* easyCBM	*t*-statistic	*p*-value	Cohen's *d*
Kindergarten (*N* = 16)	Fall	49.87 (20.43)	58.67 (3.95)	1.69	0.112	0.44
	Winter	49.27 (18.86)	55.20 (14.92)	1.16	0.266	0.29
	Spring	63.38 (17.26)	56.56 (10.86)	1.63	0.124	0.41
Grade 1 (*N* = 7)	Fall	28.57 (19.72)	41.86 (24.12)	2.11	0.080	0.79
	Winter	62.14 (18.62)	63.00 (22.54)	0.14	0.896	0.05
	Spring	61.14 (35.16)	55.86 (19.21)	0.67	0.529	0.25
Grade 2 (*N* = 15)	Fall	39.93 (37.56)	32.27 (18.54)	0.79	0.445	0.20
	Winter	43.33 (32.42)	52.80 (17.97)	0.95	0.360	0.24
	Spring	48.67 (32.14)	40.87 (25.08)	0.86	0.404	0.22
Grade 3 (*N* = 10)	Fall	38.50 (16.59)	31.20 (14.47)	1.14	0.286	0.36
	Winter	51.73 (25.66)	43.09 (14.62)	1.04	0.324	0.31
	Spring	53.33 (22.66)	54.58 (20.84)	0.16	0.879	0.45
Grade 4 (*N* = 12)	Fall	35.75 (22.66)	27.17 (19.46)	1.29	0.223	0.37
	Winter	21.17 (14.71)	25.33 (14.41)	1.09	0.301	0.31
	Spring	17.75 (17.11)	23.94 (18.30)	1.83	0.087	0.46
Grade 5 (*N* = 19)	Fall	19.32 (11.93)	18.32 (15.00)	0.27	0.787	0.06
	Winter	9.24 (6.13)	19.00 (10.71)	4.49	<.001***	0.98
	Spring	9.67 (5.75)	19.54 (14.45)	4.46	<.001***	0.91
Grade 6 (*N* = 12)	Fall	31.25 (24.01)	18.42 (11.15)	1.88	0.087	0.54
	Winter	18.92 (16.04)	15.67 (6.57)	0.95	0.362	0.28
	Spring	26.62 (14.91	23.15 (13.66)	1.27	0.228	0.35

***$p<0.001$.

per-pencil and computer-based measures used in this study, easyCBM and Istation. This constitutes a research gap that future studies need to address. Second, the research questions for the study were derived from working with the teachers at the participating school and not solely the research team, making them context-specific and directed at addressing a specific problem in the school. We stress this because we found it to be one of the benefits of community-based research projects; it not only empowered the teachers to share their experiences with assessment, it also gave our research team an opportunity to address their needs through a collaborative partnership. That is, we used the study as a platform for addressing the teachers' involvement in research projects and assisting them in making informed decisions about their practice. As such, the study represents a small step in forging the research-to-practice component that is often missing in large-scale projects. Third, the study is based on data from a rural school, which is an understudied population, and we hope to extend our knowledge base in this area.

Lastly, the study addresses the validity of test scores that are generated by paper-pencil vs. computerized modalities as an important factor to consider for interpreting results and for ensuring more targeted interventions for students. We use the definition of validity put forth by the Standards for Educational and Psychological Testing by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME) as the "degree to which evidence and theory support the interpretations of test scores for the proposed uses of tests" (AERA, APA, & NCME, 2014, p. 11).

Limitations and Future Directions

First, the study is a product of PAR in a community setting, which limits its generalizability. We aimed to answer a research question that was of interest to this particular school and this particular group of teachers and students. While we were successful in doing that, the sample size is small, so our findings cannot be generalized to other rural schools or school districts. A direction for future research would be to conduct several such studies, especially in rural schools, to learn more about the applicability of these progress-monitoring tools in these contexts.

Second, we conducted the study right after the pandemic, which might explain some of the differences observed across testing modalities. A future direction would be to extend this work to other rural schools as well as to test the same students again at a later stage to see if there would be any changes in our findings.

Finally, we detected an overall lack of research comparing these progress-monitoring tools. Even though we found them to measure similar constructs as depicted by our correlation matrix and research in the field, our search revealed few research studies in the extant literature that compared the two modalities in any school setting. Several states have moved from paper-pencil to computer-based testing to save time and for the convenience it affords, but it is imperative to study the effects of this change on teachers and students. Many teachers we have worked with in this and other schools have expressed that the move to computer-based testing has limited the immediate feedback they used to get on student performance afforded by paper-pencil tests. This has serious implications for instructional planning, identifying students at risk for reading and math disabilities, and providing early intervention to address the needs of students. As a future direction, therefore, we recommend that more studies examine this issue in different parts of the country, so we can get a better understanding of the universality of these assessment trends.

References

Alonzo, J., Park, B. J., & Tindal, G. (2012). *The development of the easyCBM CCSS reading assessments: Grade 3* (No. 1221) [Technical report]. University of Oregon.

Alonzo, J., & Tindal, G. (2009). *Alternate form and test-retest reliability of easyCBM reading measures.* University of Oregon, Behavioral Research and Teaching.

American Educational Research Association, American Psychological Association, & National Council for Measurement in Education. (2014). *Standards for educational and psychological testing.* American Educational Research Association.

Anderson, D., Alonzo, J., Tindal, G., Farley, D., Irvin, P. S., Lai, C. F., … & Wray, K. A. (2014). *Technical manual: easyCBM* [Technical report #1408]. University of Oregon, Behavioral Research and Teaching.

Barber, B. (1992). *An aristocracy for everyone: The politics of education and the future of America.* Oxford University Press.

Brown, D. (2001). *Putting it together: A method for developing service-learning and community partnerships based in critical pedagogy.* Corporation for National Service.

Campbell, L. O., Howard, C., Lambie, G. W., & Gao, X. (2022). The efficacy of a computer-adaptive reading program on grade 5 students' reading achievement scores. *Education and Information Technologies*, *27*(6), 8147–8163. https://doi.org/10.1007/s10639-022-10953-5

Choi, S. W., & Tinkler, T. (2002). Evaluating comparability of paper-and-pencil and computer-based assessment in a K-12 setting. In *Proceedings of annual meeting of the National Council on Measurement in Education,* New Orleans, LA.

Chua, Y. P. (2012). Effects of computer-based testing on test performance and testing motivation. *Computers in Human Behavior*, *28*(5), 1580–1586. https://doi.org/10.1016/j.chb.2012.03.020

Cook, M., & Ross, S. M. (2021). *The efficacy of Istation's early reading assessment and curriculum on MAP reading growth in a South Carolina school district*. Retrieved from www.istation.com/.

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, *37*(3), 184–192. https://doiorg.libproxy.unm.edu/10.1177/00224669030370030801

Farley-Ripple, E. N., & Buttram, J. L. (2014). Developing collaborative data use through professional learning communities: Early lessons from Delaware. *Studies in Educational Evaluation*, *42*, 41–53. https://doi.org/10.1016/j.stueduc.2013.09.006

Fletcher, J. M., & Vaughn, S. (2009). Response to intervention: Preventing and remediating academic difficulties. *Child Development Perspectives*, *3*(1), 30–37. https://doi.org/10.1111/j.1750-8606.2008.00072.x

Ford, M. L. (2014). *Active learning strategies and first grade reading achievement using the TAKE 10! program and Istation assessment tool: A correlation study* [Doctoral dissertation]. Jones International University.

Freire, P. (1974). *Pedagogy of the oppressed.* Seabury.

Fuchs, D., & Fuchs, L. S. (2017). Critique of the national evaluation of response to intervention: A case for simpler frameworks. *Exceptional Children, 83*(3), 255–268. doi:10.1177/0014402917693589

Gartland, D., & Strosnider, R. (2020). The use of response to intervention to inform special education eligibility decisions for students with specific learning disabilities. *Learning Disability Quarterly, 43*(4), 195–200. doi:10.1177/0731948720949964

Good, R. H., & Kaminski, R. A. (2009). *Dynamic indicators of basic early literacy skills*. University of Oregon.

Greenwood, C. R., & Kim, J. M. (2012). Response to intervention (RTI) services: An eco-behavioral perspective. *Journal of Educational and Psychological Consultation*, *22*(1–2), 79–105. https://doi-org.libproxy.unm.edu/10.1080/10474412.2011.649648

Habermas J. (1979). *Communication and the evolution of society.* Beacon.

Hall, B. L. (1992). From margins to center? The development and purpose of participatory research. *American Sociologist, 23*, 15–28. https://doi.org/10.1007/BF02691928

Hord, S. M. (1997). *Professional learning communities: Communities of continuous inquiry and improvement.* Southwest Educational Development Laboratory, Office of Educational Research and Improvement, U.S. Department of Education.

Horton, A. I. (1989). *The Highlander Folk School: A history of its major programs, 1932–1961.* Carlson Publishing.

Hudson, P., Hudson, S., Gray, B., & Bloxham, R. (2013). Learning about being effective mentors: Professional learning communities and mentoring. *Procedia-Social and Behavioral Sciences*, *93*, 1291–1300. https://doi.org/10.1016/j.sbspro.2013.10.031

Jin, H., Shin, H., Johnson, M. E., Kim, J., & Anderson, C. W. (2015). Developing learning progression-based teacher knowledge measures. *Journal of Research in Science Teaching*, *52*(9), 1269–1295. https://doi-org.libproxy.unm.edu/10.1002/tea.21243

Kahne, J., & Westheimer, J. (1996). In the service of what? *Phi Delta Kappan, 77*, 592–600.

Keller-Margulis, M. A., Shapiro, E. S., & Hintze, J. M. (2008). Long-term diagnostic accuracy of curriculum-based measures in reading and mathematics. *School Psychology Review*, *37*(3), 374–390. https://doi-org.libproxy.unm.edu/10.1080/02796015.2008.12087884

Kendeou, P., Papadopoulos, T. C., & Spanoudis, G. (2012). Processing demands of reading comprehension tests in young readers. *Learning and Instruction*, *22*(5), 354–367. https://doi.org/10.1016/j.learninstruc.2012.02.001

Kim, J. S., Vanderwood, M. L., & Lee, C. Y. (2016). Predictive validity of curriculum-based measures for English learners at varying English proficiency levels. *Educational Assessment*, *21*(1), 1–18. https://doi-org.libproxy.unm.edu/10.1080/10627197.2015.1127750

Kingston, N. M. (2008). Comparability of computer-and paper-administered multiple-choice tests for K- 12 populations: A synthesis. *Applied Measurement in Education*, *22*(1), 22–37. https://doi-org.libproxy.unm.edu/10.1080/08957340802558326

Knudsen, J., Lundh, T. P., Education, S.R.I., Hsiao, M., & Saucedo, D. (2020). *Zearn math curriculum study professional development final report*. SRI Education.

Lewin, K. (1948). *Resolving social conflicts*. Harper & Brothers.

Mathes, P. (2009). *Istation's indicators of progress: Early reading reliability and validity evidence.* Southern Methodist University.

Mathes, P. (2016). *Istation's indicators of progress (ISIP) advanced reading technical manual. Version 4.* Southern Methodist University.

Mathes, P., Torgesen, J., & Herron, J. (2014). *Istation's indicators of progress (ISIP) early reading technical manual.* Southern Methodist University.

May, M., Sargent-Hu, S., Hill, G., & Crossgrove Fry, V. (2018). *Istation program implementation evaluation.* Southern Methodist University.

Miura Wayman, M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education*, *41*(2), 85–120. https://doi-org.libproxy.unm.edu/10.1177/00224669070410020401

Moats, L., & Sedita, J. (2004). *LETRS: A road to reading comprehension*. Sopris West.

Montelione, T.A.A. (2021). *Influence of Istation formative assessment on reading student achievement* [Doctoral dissertation]. Walden University.

National Center for Education Statistics. (2021). *Fast facts*. U.S. Department of Education, Institute of Education Sciences. Retrieved from http://nces.ed.gov/

New Mexico Center on Law and Poverty. (n.d.). *Yazzie/Martinez v. state of New Mexico decision.* https://nmpovertylaw.org/wp-content/uploads/2018/09/Graphic-Yazzie-Martinez-Decision.pdf

New Mexico Public Education Department (NM-PED). (2021). *New Mexico's multi-layered system of supports*. https://webnew.ped.state.nm.us/bureaus/multi-layered-system-of-supports-mlss/

Reason, P. (1994). Three approaches to participative inquiry. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 324–339). Sage.

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, *47*(6), 427–469. https://doi.org/10.1016/j.jsp.2009.07.001

Stoll, L., Bolam, R., McMahon, A., Wallace, M., & Thomas, S. (2006). Professional learning communities: A review of the literature. *Journal of Educational Change*, *7*(4), 221–258. https://doi.org/10.1007/s10833-006-0001-8

Strand, K., Marullo, S., Cutforth, N., Stoecker, R., & Donohue, P. (2003). Principles of best practice for community-based research. *Michigan Journal of Community Service Learning*, *9*(3), 5–15. http://hdl.handle.net/2027/spo.3239521.0009.301

Street, A. (Ed.). (1995). Establishing a participatory action research group. In A. Street (Ed.), *Nursing replay: Researching nursing culture together* (pp. 59–78). Churchill Livingstone.

Stringer, E. (1996). *Action research: A handbook for practitioners.* Sage.

Sugai, G., & Horner, R. H. (2009). Responsiveness-to-intervention and school-wide positive behavior supports: Integration of multi-tiered system approaches. *Exceptionality*, *17*(4), 223–237. https://doi-org.libproxy.unm.edu/10.1080/09362830903235375

Tindal, G. (2013). *Curriculum-based measurement: A brief history of nearly everything from the 1970s to the present.* International Scholarly Research Notices.

Vescio, V., & Adams, A. (2015). *Learning in a professional learning community: The challenge evolves.* In D. Scott & E. Hargreaves (Eds.), *The SAGE handbook of learning* (pp. 274–284). Sage Publications.

Wang, S., Jiao, H., Young, M. J., Brooks, T., & Olson, J. (2008). Comparability of computer-based and paper-and-pencil testing in K-12 reading assessments: A meta-analysis of testing mode effects. *Educational and Psychological Measurement*, *68*(1), 5–24. https://doi-org.libproxy.unm.edu/10.1177/00131644073055

Wolf, R., Locke, V., & Lewis, S. (2022). *Predicting TELPAS overall category rankings using ISIPTM reading*. Southern Methodist University.