



THE EDTECH COLLECTIVE
Instructure Partner Ecosystem

BRAINSCIENCE WORKS

Spring 2026 Implementation Study

Prepared by Instructure:
Meetal Shah, Ph.D., Lead Researcher

August 2026



EXECUTIVE SUMMARY

BrainScience Works contracted with Instructure to examine the effect of its early-literacy program on transitional-kindergarten students' letter and literacy outcomes through a small, randomized pilot. Literacy achievement is defined broadly to include standardized early-literacy skills as well as program-aligned letter recognition. This study was designed as a feasibility pilot to estimate effects and inform a future adequately powered evaluation.

Study Sample and Methodology

The *BrainScience Works* program engages students in systematic, digital literacy instruction that incorporates multisensory activities such as building letters with clay, writing in journals, and playing interactive games. Educators support student learning by monitoring usage through the platform, documenting student work, completing training modules, and participating in coaching sessions.

This study draws on data from the spring 2026 semester in one transitional-kindergarten classroom in a California public school district and includes 18 students. English-speaking students were stratified by baseline achievement and randomly assigned to the user group, which received the *BrainScience Works* program (users, $n = 7$), or a comparison group, which received the district's standard transitional-kindergarten program (non-users, $n = 7$). Four Spanish-speaking students received a separate, video-based version of the program and are reported descriptively in the appendix.

Researchers used the STAR Early Literacy assessment (Renaissance) unified scaled score as the primary standardized outcome, an on-platform *BrainScience Works* letter recognition check as a secondary outcome, and STAR skill-area scores as exploratory outcomes. A baseline equivalence test confirmed the stratified randomization produced comparable groups, with no statistically significant difference between the user and non-user groups on the STAR pretest ($p = .89$). For each outcome, an analysis of covariance compared the users and non-users after adjusting for students' pretest scores.

Conclusions

Because the randomized comparison includes only 14 students, the study was not powered to detect statistical significance, and results should be read as preliminary and directional. Across outcomes, the pattern consistently favored the users. After adjusting for baseline, the users scored about 26 STAR scaled-score points higher than the non-users, a difference driven largely by the non-users declining over the study window while the users held steady, and the users scored higher on every exploratory literacy domain, with several differences approaching significance (p -values between .07 and .12). The *BrainScience Works* letter check showed no group difference, consistent with a ceiling effect on that measure. These encouraging, directional findings support proceeding to a larger, adequately powered study; they do not by themselves establish an ESSA evidence tier.

TABLE OF CONTENTS

INTRODUCTION	4
IMPLEMENTATION	7
OUTCOME FINDINGS	8
CONCLUSIONS AND RECOMMENDATIONS	9
REFERENCES	10
APPENDIX A. BRAINSCIENCE WORKS LOGIC MODEL	11
APPENDIX B. EXPLORATORY OUTCOMES	12
APPENDIX C. DESCRIPTIVE FINDINGS FOR THE SPANISH-LANGUAGE GROUP	13



INTRODUCTION

Effective interventions for dyslexic students should not only provide phonics instruction but should also target visual-spatial skills. However, the primary focus of many interventions for dyslexic students is on phonemic awareness. To address this gap in education for dyslexic individuals, *BrainScience Works* provides a structured program that systematically builds literacy skills by implementing visual-spatial instruction to support reading fluency and comprehension, through explicit instruction and tailored assessments to ensure mastery and success in reading and writing.

As part of their ongoing efforts to demonstrate the efficacy of the *BrainScience Works* program, the team contracted with Instructure to conduct a small, randomized pilot examining the program's effect on transitional-kindergarten students' early-literacy outcomes. Building on a prior correlational study of program usage and achievement, this pilot introduced a randomized control condition to estimate the program's effect and to inform the design of a future, adequately powered evaluation. The study was framed as a feasibility pilot rather than a confirmatory trial, given the small number of students available in a single classroom.

The following research questions guided this study:

Student Outcomes

1. After adjusting for students' baseline achievement, how did the *BrainScience Works* program affect students' standardized early-literacy achievement, as measured by the STAR Early Literacy scaled score, relative to non-users?
2. After adjusting for baseline, how did the program affect students' program-aligned letter recognition, as measured by the *BrainScience Works* letter check?

Exploratory Questions

3. How was the program related to specific early-literacy skill domains — alphabetic sequence, alphabetic knowledge, letter matching, letter recognition, and sound recognition?
4. Among the Spanish-speaking students who received the Spanish-language video adaptation, what descriptive patterns emerged in their early-literacy outcomes?

This report details the study design and methods, implementation, findings, and conclusions.



STUDY DESIGN AND METHODS

This section of the report briefly describes the study participants, measures, and analysis methods.

Study Design

This study used a small randomized controlled design, with students randomly assigned to receive the *BrainScience Works* program or to a control condition. Because only a single classroom was available, the study was conducted as a feasibility pilot: it was intended to estimate the program's effect and inform the design of a larger evaluation, rather than to serve as a confirmatory or certification study.

Setting and Participants

This study draws on data from the spring 2026 semester in one transitional-kindergarten classroom in a California public school district and includes 18 students. English-speaking students were first stratified by baseline achievement — using a composite of STAR percentile rank and *BrainScience Works* letter mastery (high, above 90; medium, 80 to 89; low, below 80) — and then randomly assigned within each stratum to either use the *BrainScience Works* program (users, $n = 7$), or to a comparison group receiving the district's standard transitional-kindergarten program (non-users, $n = 7$). This stratified procedure was used to promote baseline equivalence across the two groups given the small sample. Four Spanish-speaking students received a Spanish-language adaptation of the program delivered via recorded video rather than the interactive platform; because this group had no comparison condition, its outcomes are reported descriptively in the appendix.

Measures

This study used standardized and program-aligned measures to assess students' early-literacy outcomes.

Standardized Assessment. The STAR Early Literacy assessment (Renaissance) provided the primary standardized outcome, the Unified scaled score, along with a percentile rank and a set of skill-area scores — alphabetic sequence, alphabetic knowledge, letter matching, letter recognition, and sound recognition — all drawn from the Renaissance research export.

BrainScience Works Letter Check. Administered on the *BrainScience Works* platform, this program-aligned measure assessed students' correct identification of lowercase and uppercase letters, each scored out of 26 possible and combined into a total out of 52.

Testing followed the same schedule for all students: the STAR pretest and platform pretest in early March, program instruction until late May, the STAR posttest in late April (before the program concluded), and the platform posttest in late May.

Data Analysis

Descriptive statistics summarized participant characteristics. A baseline equivalence test confirmed that the stratified randomization produced comparable groups, with no statistically significant difference between users and non-users on the STAR pretest ($p = .89$). For each outcome, an analysis of covariance compared the users and non-users after adjusting for students' corresponding pretest score, yielding a pretest-adjusted (marginal) mean for each group and an adjusted difference between them. Adjusting for the pretest is standard practice because baseline achievement is typically the strongest predictor of end-of-year performance.

Student demographics were not included as additional covariates: with only 14 students in the randomized comparison, and with home language and IEP status near-constant within the two English-speaking groups, adding them would exceed what the sample can support. Because of the small sample, all estimates should be interpreted as preliminary and directional rather than as tests of statistical significance.

IMPLEMENTATION

During the instructional window (March 12 to May 25, 2026), users received the *BrainScience Works* program in their transitional-kindergarten classroom, with lessons administered by the classroom teacher; program-aligned assessments were completed on the *BrainScience Works* platform. Non-users received the district's standard transitional-kindergarten program and did not complete *BrainScience Works* lessons, though they completed the same platform-based letter check at pretest and posttest to provide a comparable outcome measure.

The four Spanish-speaking students received a Spanish-language adaptation of the program delivered through recorded video rather than the interactive platform, supported by a bilingual instructor working with the students in a small group. Because the recorded videos required manual navigation and did not include the platform's built-in stopping points, these students needed more hands-on adult support than the standard platform experience provides.

Unlike the prior study, student-level implementation data for this pilot — such as modules completed and time spent in the program — was not provided to the evaluation team. As a result, this report does not include usage metrics or dosage analyses.

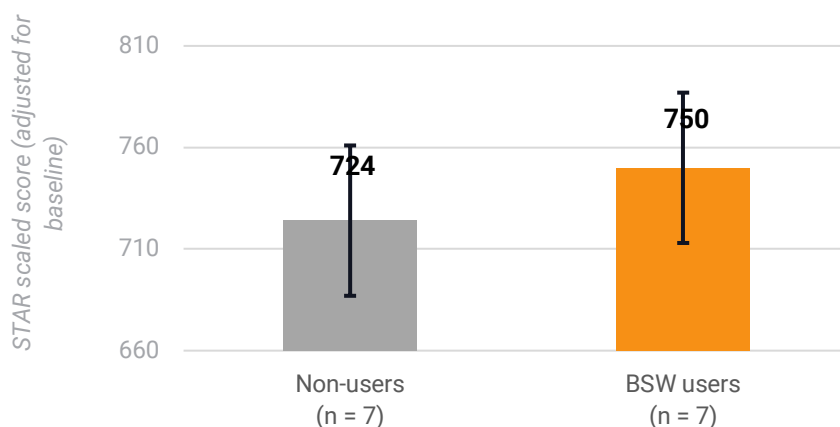
OUTCOME FINDINGS

To estimate the program's effect, researchers compared users and non-users on the primary standardized outcome, the STAR Early Literacy scaled score, and the secondary program-aligned outcome, the *BrainScience Works* letter check, in each case adjusting for the corresponding pretest. Results are reported as pretest-adjusted (marginal) means and the adjusted difference between groups. Because the randomized comparison includes only 14 students, these estimates are directional and were not expected to reach statistical significance.

Outcome	Non-users Adjusted Mean	BSW users Adjusted Mean	Difference	<i>p</i> - value
STAR scaled score (primary)	724	750	+26	.30
BSW letter check total, 0–52 (secondary)	47.9	48.3	+0.4	.86

Primary Outcome: STAR Early Literacy Scaled Score

After adjusting for baseline, the users' mean STAR scaled score was 750, compared with 724 for the non-users — an adjusted difference of 26 points favoring users ($p = .30$). This gap was driven largely by the non-users declining over the study window while the users held steady. The difference is not statistically significant at this sample size, but its direction and size are consistent with a meaningful program effect worth testing in a larger study.



Secondary Outcome: *BrainScience Works* Letter Check

After adjusting for baseline, the two groups scored almost identically on the letter check (users 48.3, non-users 47.9 out of 52; adjusted difference +0.4, $p = .86$). This is expected: both groups received letter instruction, the assessment runs on the platform, and scores clustered near the top of the scale, leaving little room to detect group differences. This ceiling effect limits the letter check's usefulness as an outcome and is one reason the standardized STAR measure is the more informative primary outcome.

CONCLUSIONS AND RECOMMENDATIONS

This feasibility pilot found consistent, positive, directional evidence favoring students who used the *BrainScience Works* program. After adjusting for baseline, users outscored non-users on the primary standardized outcome — the STAR Early Literacy scaled score — by 26 points and scored higher than non-users on every exploratory literacy domain. Because the randomized sample was small ($n = 14$), none of these differences reached statistical significance, and all should be interpreted as preliminary estimates rather than confirmed effects. The pattern is nonetheless encouraging, and the pilot demonstrated that a randomized design can be implemented in this setting — a solid foundation for a larger, adequately powered evaluation.

A note on evidence standards. Unlike the prior correlational study, this pilot does not by itself satisfy an ESSA evidence tier, because those tiers require at least one statistically significant, positive finding. The randomized design used here, however, aligns structurally with what ESSA Level II (Moderate Evidence) requires; carried out with a larger sample and adequate statistical power, a study of this design could support a Level II claim.

To strengthen future evaluations and better understand the program's impact, we offer the following recommendations:

- **Increase sample size and statistical power.** A single classroom of 14 randomized students cannot support reliable inference. A larger, multi-classroom sample would allow the directional effects seen here to be tested with adequate power.
- **Prioritize standardized, ceiling-resistant outcome measures.** The program-aligned letter check clustered near the top of its scale, leaving little room to detect group differences. Leaning on standardized measures such as STAR, which are less prone to ceiling effects, will yield more informative outcomes.
- **Investigate the decline among non-users.** Non-users declined on the STAR measures over the study window. A larger study could examine contextual factors - such as attendance, classroom disruption, and home language support - to understand whether this pattern is real and what drives it.

Continued evaluation with a larger sample and improved data completeness will be essential to confirm the promising, directional effects observed in this pilot.

REFERENCES

Renaissance Learning. (2024). *Star Assessments for Early Literacy technical manual*. Renaissance Learning, Inc.

What Works Clearinghouse. (2022). *What Works Clearinghouse procedures and standards handbook, version 5.0*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). This report is available on the What Works Clearinghouse website at [URL](#)

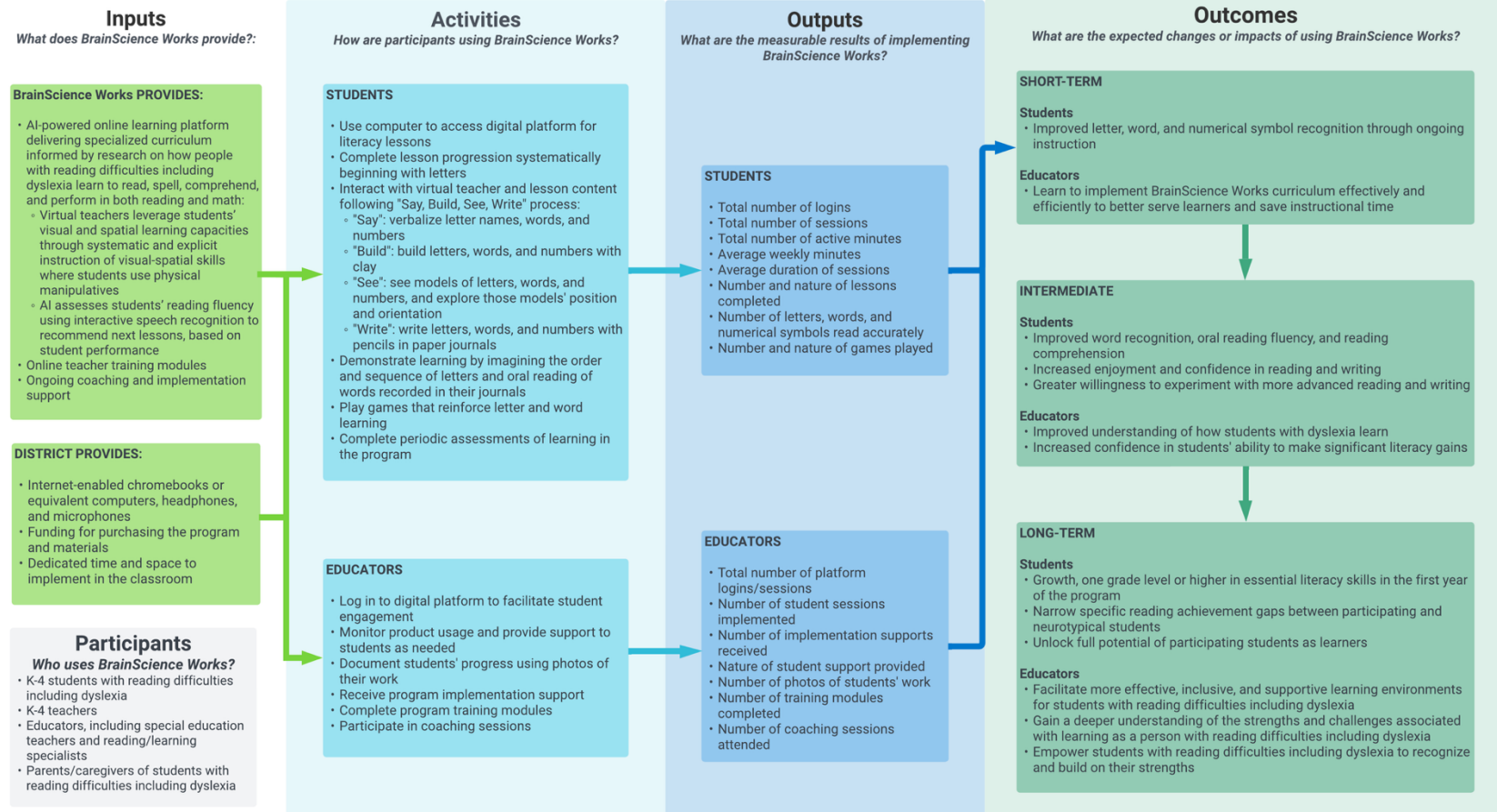


APPENDIX A. BRAINSCIENCE WORKS LOGIC MODEL



Problem Statement: Students with dyslexia benefit most when instructional strategies emphasize visual-spatial learning skills. Unfortunately, today's reading instruction uses predominantly phonics strategies that are most effective for the neurotypical or non-dyslexic brain. BrainScience Works provides a structured program that systematically builds literacy skills by implementing visual-spatial instruction to support reading, spelling, comprehension, as well as math, through explicit instruction and tailored assessments to ensure mastery and success in reading, writing, and math.

BrainScience Works Logic Model



APPENDIX B. EXPLORATORY OUTCOMES

Beyond the primary and secondary outcomes, we applied the same pretest-adjusted model to a set of exploratory measures: the STAR percentile rank and five STAR skill-area domains — alphabetic sequence, alphabetic knowledge, letter matching, letter recognition, and sound recognition. These should be read as an exploratory profile rather than as separate confirmatory tests. With a small sample and multiple outcomes, individual results are best interpreted collectively.

Exploratory Outcome	Non-users Adjusted Mean	BSW users Adjusted Mean	Difference	<i>p</i> - value
STAR percentile	73.	86.3	+13.2	.073
Alphabetic knowledge	79.5	87.6	+8.1	.086
Letter matching	82.9	90.6	+7.7	.077
Letter recognition	82.1	89.9	+7.8	.079
Sound recognition	67.8	77.2	+9.5	.119
Alphabetic sequence	52.2	62.0	+9.8	.190

Across every exploratory measure, users posted a higher adjusted mean than non-users. The differences ranged from roughly 8 to 13 points, and several sat just above the conventional significance threshold (*p* between .07 and .12) — notable in a sample this size, though none should be read as a confirmed effect. The most striking feature is the consistency: the same direction, favoring users, on every domain, mirroring the pattern on the primary STAR outcome.

One interpretive note: the STAR skill-area scores are estimated from a student's overall STAR scaled score and aligned to early-literacy standards, rather than measured independently. Because they share that common source, the domains are correlated with one another and with the primary outcome, so their consistent direction reflects a shared basis and should not be read as independent replication of the effect.

APPENDIX C. DESCRIPTIVE FINDINGS FOR THE SPANISH-LANGUAGE GROUP

Four Spanish-speaking students received a Spanish-language adaptation of the *BrainScience Works* program delivered via recorded video. Because this group had no comparison condition and includes only four students, results are presented descriptively and should not be interpreted as evidence of program effect. To protect student privacy given the very small group, only aggregate pre- and post-test means are reported.

Measure	Pretest mean	Posttest mean
STAR scaled score	666	674
STAR percentile	64.5	60.5
BSW letter check total (0–52)	20.0	32.8
— lowercase (0–26)	10.2	16.0
— uppercase (0–26)	9.8	16.8

These students began well below their English-speaking peers on every measure, which left substantial room to grow on the program-aligned letter check. Over the course of the pilot, the group's average letter check total rose from 20.0 to 32.8 out of 52, its largest movement on any measure, while STAR results were flat to mixed. With only four students and no comparison group, these patterns are illustrative only, but they suggest the Spanish-language adaptation is worth evaluating with a properly designed comparison study in future work.