2022

Utah’s Reading Intervention Program Efficacy Research

The Science of Reading + AI = Growth

Evidence from Utah

Abstract

The Utah State Board of Educaton approves and funds several intervention programs for use by Utah Districts. Employing usage data at the student level, the State then commissions independent research to analyze the impact and effect size of these programs.

In 2022, for students using at the recommended dosage, the State's data showed that Amira's effect size was over .4, in Hattie's Zone of Desired Effects.

The study shows gains on the Utah State Summative Assessment for frequent users of Amira far in excess of the average for Utah students employing other interventions.

What Does This Research Prove About Amira's Efficacy?

From the "Evaluation & Training Institute" Report:

Mean score differences and effect sizes of matched program students who met Amira's recommendations across both average weekly
minutes and total weeks. As shown below, all treatment students in grades 1st - 3rd exhibited higher predicted mean scores than their matched control counterparts. Second grade students exhibited the highest mean score differences, with treatment students scoring 21 points higher than their control counterparts, on average. Among this highest use analytic sample, first and second graders had effect sizes within the large effect size range (g= 0.41 and g=0.45, respectively), while third grade students had an effect size within the medium range (g=0.25).

Who Conducted and Published the Research?

The research was sponsored by the Utah State Board of Education. The study was conducted by The Evaluation and Training Institute, a non-profit, nonpartisan research organization, is dedicated to improving academic excellence and social-emotional well being from preschool to adulthood.

Utah students who worked at recommended dosage saw remarkable gains. As shown from the research below, student gains exceed .40 effect size.

Background (From Utah's Report)

"The Early Intervention Software Program (EISP) was designed to increase the literacy skills of all students in K-3 through adaptive computer-based literacy software.

The program provided Utah's Local Education Agencies (LEAs) with an option to select among four adaptive computer-based programs. State-wide program implementation provided the opportunity for large numbers of students to receive program benefits. "

MRU Sample Predicted End-of-Year Acadience Reading Composite Mean Scores

Grade	Ctrl		Tr		Dif.	ES
	Mean	SE	Mean	SE		
Kindergarten	ISS					
First Grade	76.37	0.44	87.16	1.94	10.79	**0.41**
Second Grade	284.92	0.71	305.89	3.27	20.97	**0.45**
Third Grade	384.06	1.03	403.87	4.77	19.81	*0.25*

Even students who worked at only 80% of the recommended dosage saw meaningful improvements.

MRU80 Sample Predicted End-of-Year Acadience Reading Composite Mean Scores

Grade	Ctrl		Tr		Dif.	ES
	Mean	SE	Mean	SE		
Kindergarten	ISS					
First Grade	75.21	0.40	79.42	0.99	4.21	*0.15*
Second Grade	270.00	0.69	287.05	1.74	17.05	**0.30**
Third Grade	380.88	0.90	403.32	2.29	22.44	**0.34**

2022-2023 Program Enrollment by Grade

Kindergarten	First Grade	Second Grade	Third Grade
179	7,922	8,010	8,016

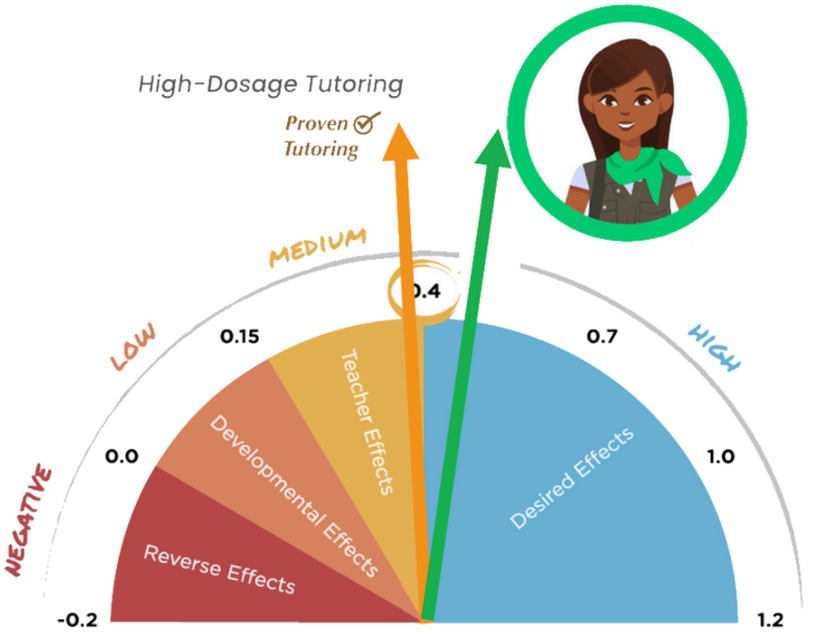
How Was the Study Conducted?

Statistical Modeling of Program Impacts on Acadience Test Scores. Ordinary least squares (OLS) regression models were computed for each analytic sample. The OLS models predicted the differences in treatment and control groups' end-of-year group mean scores, while controlling for students' beginning-of-year (BOY) reading scores and key demographics, gender, race, ELL status, SPED designation, and poverty status. We examined treatment effects for each analytic sample based on their usage and grade.
1 Students in kindergarten, 2nd and 3rd grade were matched on reading composite scores (BOY Comp) and students in 1st grade were matched on nonsense word fluency, correct letter sounds (NWF-CLS) scores.

Amira showed significant treatment effects for students in first through third grade among students who met Amira's usage requirements, and even demonstrated benefits for those meeting a more lenient usage criteria.

The following results are broken up into two different usage groups of K-3rd grade students and their matched control counterparts, (1) students who met Amira's recommended weeks and average minutes, and (2) students who met 80% of recommended weeks and average minutes. This section is focused on participants who engaged with the Amira program most closely aligned to the recommendation. Results for the third usage group (ITT), which included the students whose time with the program fell far below the recommend levels, can be found in Appendix B. To determine if the mean score differences could be interpreted as meaningful, we examined their. effect sizes. Effect sizes show the magnitude of the difference between two groups on an outcome and are often interpreted as meaningful if they reach a certain minimum threshold. We adapted a. set of effect size benchmarks based on categories from Kraft (2020) that were adjusted for early. literacy outcome measures: less than 0.10 is small, 0.10 to less than .30 is medium, and .30 or. greater is large (M. Kraft, personal communication, October 13, 2023).

1.5 Years Of Growth In 22/23 SY.



	Other Edtech in 2022	Amira in 2023
1st Grade	.09	.41
2nd Grade	-.03	.45
3rd Grade	.09	.25

Where Can One Access the Study?

Background is available on the Amira Learning website.

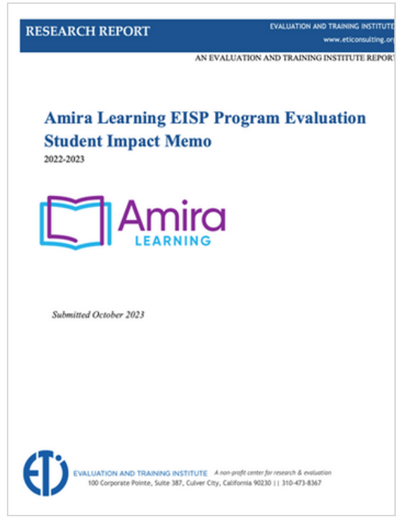
Related Studies

Several other studies support Amira's impact on literacy growth:

Evaluation and Training Institute (2023): Found Amira significantly outperformed a control group in Utah's Early Intervention Reading Software Program **(Read Report)**.

Mostow et al. (CMU): Demonstrated that AI-powered reading tutors can be as effective as human tutors **(Read Study)**.

Poulsen et al. (UBC): Showed Amira's effectiveness for bilingual students **(Read Study)**


RESEARCH REPORT

EVALUATION AND TRAINING INSTITUTE
www.eticonsulting.org

AN EVALUATION AND TRAINING INSTITUTE REPORT

Amira Learning EISP Program Evaluation
Student Impact Memo
2022-2023

Amira LEARNING

Submitted October 2023

EVALUATION AND TRAINING INSTITUTE *A non-profit center for research & evaluation*
100 Corporate Pointe, Suite 387, Culver City, California 90230 | | 310-473-8367

Summary & Conclusion

The research shows that Amira delivers effect size in Hattie's Zone of Desired Effect when a student works with the program more than 30 minutes a week.

The researchers concluded:
"Students served by Amira outperformed the students who were not. Further, the students who were able to use with the software as it was intended by Amira also showed. greater end-of-year literacy scores relative to those participating below the recommended usage levels in the program."