Amira's Impact for Middle School Students

Kelly Boden, PhD

Amira Learning | Every Child Deserves the Chance to Become a Reader
5214F Diamond Heights Blvd, #3255, San Francisco, CA 94131
wwww.AmiraLearning.com | 866-883-7323 | info@amiralearning.com

Kelly Boden, PhD

Amira's Impact for Middle School Students

Key Findings:

1. **+6 to +7 Percentile Gains with Consistent Use:** Sixth-grade students using Amira for 20–30+ minutes per week gained 6–7 percentile points more than low-usage peers across two consecutive years, even after controlling for starting reading levels.
2. **Reliable, Year-over-Year Impact:** The positive effects of Amira were replicated in both the 2023–2024 and 2024–2025 school years, with effect sizes of 0.35–0.40—comparable to high-quality one-on-one tutoring and well within the "Zone of Desired Effects" for educational interventions.
3. **Research-Aligned and Nationally Validated:** Results align with Science of Reading principles and mirror similar findings in state-level studies (e.g., North Dakota, Texas, Utah), confirming Amira's broad, evidence-based effectiveness across diverse middle school settings.

Overview

Amira was created for one purpose – to couple the Science Of Reading with AI, giving every child a pathway to the power of reading. Amira was born in academia, is delivering accelerated growth by reflecting the guidance of leading Reading Scientists, and is rapidly evolving to reflect every key element of Science of Reading Research.

This analysis examines the effects of the Amira reading tutoring platform on middle-school student reading achievement over the 2023–2024 and 2024–2025 school years. Specifically, this analysis examined the impact of reaching the recommended amount of weekly tutor usage (20-30 mins) on students' reading ability. Reading ability was assessed at beginning-of-year (BOY) and end-of-year (EOY) using the ARM measure.

ARM is a composite score derived from Amira's curriculum based tasks as part of the Benchmark Assessment/Progress Monitoring/Dyslexia Screener. During the screener process, students attend to 5 to 10 different tasks, and the ARM Score is a synthesis of a student's demonstrated ability across the tasks – every task that the student does, from the Oral Reading Fluency passage to Reading Comprehension to the Rapid Automatized Naming task, will factor into their ARM score. The analysis focuses on ARM PR (percentile rank) for each student. The ARM PR situates a student's reading level relative to grade-level norms (a score of 50 indicates median performance).

Methodology

Analytical Sample

The analytic samples for both school years were selected by examining middle school districts with the highest average weekly Amira tutor usage, specifically for 6th grade, and selecting the top 10 from both school years respectively. This was done to ensure an adequate sample size of students who met the recommended weekly usage time.

The 23-24 school year analysis includes 10 school districts across 7 states (PA, IL, TN, LA, ND, NY, IN, MI). Across these districts, the overall demographic breakdown according to NCES data (2024) shows approximately 55% White students, 19% Black students, 12% Hispanic students, 6% Asian students, and 4% Other students. The White student population ranged from 1% to 95% across individual districts, while Black student enrollment varied from

1% to 85%, Hispanic enrollment from 2% to 29%, Asian enrollment from 0% to 14%, and Other enrollment from 2% to 6%.

The 24-25 school year analysis includes 10 school districts across 8 states (PA, IL, LA, NY, TX, NM, MI, KY). The overall demographic breakdown according to NCES data (2024) shows approximately 46% White students, 23% Black students, 21% Hispanic students, 6% Asian students, and 3% Other students. Individual district variations were significant, with White enrollment ranging from 5% to 97%, Black enrollment from 0% to 85%, Hispanic enrollment from 1% to 50%, Asian enrollment from 0% to 15%, and Other enrollment from 1% to 5%.

Analytical Strategy

The analysis employed Hierarchical Linear Modeling (HLM) to evaluate the impact of Amira usage on student reading performance, while controlling for ARM performance at BOY and school-level clustering. HLM was chosen specifically to account for the nested structure of the data, with students clustered within schools, ensuring that variability between schools did not bias the results.

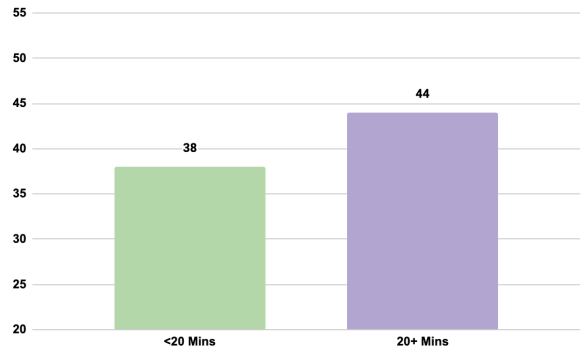
The primary comparison examined EOY ARM PR among students who averaged 20 or more minutes (averaged across 25 weeks) of Amira tutor use per week versus those with less usage.

23-24 Results

Sixth-grade students made substantial reading progress during the 2023–2024 school year, with those using Amira regularly showing the greatest gains. At the start of Fall 2023, students' reading levels (as measured by ARM percentile ranks) spanned a wide range, with the grade's average performance modestly below national norms. Over the course of the year, all students improved, as expected.

However, students who met the recommended weekly usage of 20 or minutes with the Amira platform exhibited accelerated growth relative to their peers. By Spring 2024, the recommended-usage group demonstrated considerably greater increases in ARM percentile ranks compared to their lower-using peers—exceeding low-usage students by over 6 percentile ranks, even after controlling for Beginning-of-Year (BOY) reading proficiency (see Figure 1).

Figure 1. Average end-of-year ARM PR for students reaching 20+ mins vs. those <20 minutes in 2023-24.



The difference in gains between high and low users was statistically significant ($p < .001$). High-usage students demonstrated approximately twice the reading proficiency growth of low-usage students during 2023–2024, a pattern consistent with findings from larger-scale Amira studies, including North Dakota's statewide data indicating Amira users achieved roughly two times faster growth than non-users. The estimated effect size of the usage differential in the sixth-grade data was approximately 0.35, representing a large effect by educational standards, calculated after controlling for students' BOY reading proficiency.

The data indicates that the Amira platform had a substantial impact on accelerating reading growth for sixth graders. In summary, during 2023–2024, students who used Amira consistently significantly outperformed their peers in reading growth, providing initial evidence of the platform's efficacy at the middle school level.

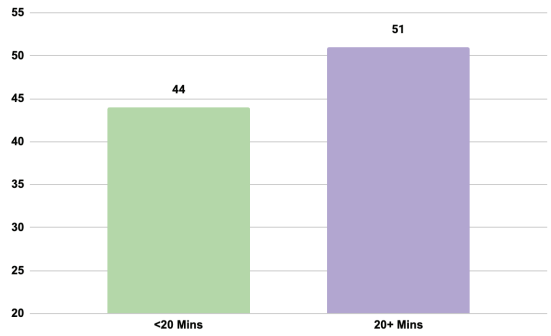
24-25 Results

The 2024–2025 data continued to demonstrate strong reading growth among sixth graders using Amira, confirming and extending the previous year's findings. Teachers integrated Amira into literacy instruction, encouraging students to practice reading with the AI tutor multiple times per week. Over the course of the year, student reading scores improved overall, but students with high Amira usage showed markedly greater gains by Spring 2025, underscoring the platform's sustained impact on reading growth at the middle school level.

By the end of the 2024–2025 academic year, sixth graders who met or exceeded 20 minutes of Amira usage per week demonstrated, on average, significantly larger increases in their ARM percentile scores compared to those who used Amira minimally or not at all. The high-usage group exceeded their lower-using peers by approximately 7 percentile ranks, even after controlling for Beginning-of-Year (BOY) reading proficiency to ensure differences were attributable to Amira

usage rather than starting ability (see Figure 2).

Figure 2. Average end-of-year ARM PR for students reaching 20+ mins vs. those <20 minutes in 2024-25.



The difference in growth between the usage groups was statistically significant ($p < .001$). The estimated effect size for the 2024–2025 usage impact was approximately 0.40, closely aligned and even exceeding the prior year's effect, reinforcing that this represents a replicable finding across years rather than a one-time anomaly.

The replication of a strong effect in 2024–2025 adds confidence that Amira's efficacy is reliable and sustained. The results corroborate the prior year's findings, demonstrating that Amira usage is associated with significantly accelerated reading growth for sixth graders. The consistency across two different cohorts, with different teachers and students, provides compelling evidence that the platform can repeatedly produce positive literacy outcomes in a middle school setting.

Comparison to Broader Research

The positive effects observed in our sixth-grade analysis align with broader research findings on Amira's efficacy. A statewide study in North Dakota found that students who used Amira regularly demonstrated significantly larger gains on the 2024 state reading assessment (NDSA) than non-users. Third graders engaging with Amira >20 minutes per week scored 15 points higher on average than low-usage peers, while 4th graders showed a 17-point advantage and 5th graders a 10-point advantage. These improvements represent substantially faster growth—roughly twice the progress of non-users over the same period.

Independent evaluations across multiple states have reported comparable results. In Texas, Amira users gained 36 STAAR scale score points more than non-users, translating to approximately a 9 percentile rank improvement with an effect size of 0.45. Utah's state education department reported effect sizes exceeding 0.4, placing Amira within John Hattie's "Zone of Desired Effects." Louisiana found positive

reading score accelerations for Amira users across all tested grades. These consistent findings across varied contexts demonstrate that our sixth-grade results represent part of a growing body of evidence supporting AI-driven tutoring's impact on reading growth.

Amira's effectiveness is grounded in Science of Reading principles, emphasizing guided oral reading practice with immediate feedback for developing fluency and comprehension. Developed in collaboration with leading reading researchers like Dr. Nell K. Duke, the platform provides structured practice that reinforces classroom instruction and accelerates skill development.

Conclusion: Key Findings and Implications

Consistent Efficacy Across Two Years: Data from both academic years demonstrate that sixth graders using Amira regularly experienced significantly greater reading proficiency growth than low-usage peers. This replication indicates reliable, sustained impact rather than a one-time occurrence.

High Usage Yields Substantially Greater Growth: Students meeting recommended usage achieved approximately 70% more growth than minimal users. High-usage students consistently demonstrated roughly double the percentile rank improvement, with effect sizes of 0.3 to 0.4, representing moderate to large educational impact.

Research-Based and Generalizable: Amira's success stems from established Science of Reading principles—explicit phonics, guided oral reading, and adaptive practice. The platform's methodology delivers evidence-based instruction that has produced measurable, significant growth across diverse real-world contexts and multiple states, confirming its broad applicability for literacy improvement.

References

1. **Amira Learning (2024).** *Student Reading Growth: The Science of Reading + AI = Student Reading Growth*. Research summary showing Amira's impact in ND, LA, UT, TX (effect sizes ~0.4 and doubled growth for high users) amiralearning.comamiralearning.com.
2. **Angel, M.** (2024). *Amira's Merger with Istation – Q&A*. Amira Learning Blog. (Describes Amira's design with Nell Duke and reports ~0.45 effect sizes and RCT results outperforming human tutors) amiralearning.comamiralearning.com.
3. **National Center for Education Statistics.** (2024). School & district navigator [Data set]. U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/ccd/districtsearch/
4. **National Reading Panel (2000).** *Teaching Children to Read* – Report summary (Reading Rockets) on evidence-based reading instruction. (Finds guided oral reading with feedback significantly improves fluency and comprehension) readingrockets.org.
5. **North Dakota Dept. of Public Instruction (2024).** *Amira Learning* – ND state analysis highlights significant gains for regular Amira users (15+ point NDSA improvements) nd.gov.
6. **Texas Education Agency Efficacy Study (2023).** Analysis of STAAR reading score gains with Amira (Amira users +36 scale points, ~9 PR gain vs. peers; effect size ~0.45) amiralearning.com.
7. **Utah State Board of Education (2022).** Independent evaluation of early literacy software. (Amira usage at recommended levels yielded >0.4 effect size, within Hattie's desired impact range) amiralearning.com.