Waterford.org and ESSA | Evidence Base & Efficacy

Waterford supports states and districts as they implement ESSA, sharing an evidence base that meets even the highest standards of ESSA's four tiers of evidence.

As a not-for-profit research organization, Waterford Research Institute, LLC, has always grown innovation from the fertile soil of evidence-based efficacy research. As such, Waterford resources and blended learning designs meet the highest standards of evidence-based frameworks, including the What Works Clearinghouse and the four-tiered evidence model outlined in the Every Child Succeeds Act (ESSA).[1]

In fact, the Waterford Early Reading Program (part of Waterford Reading Academy) is one of the few digital programs for early literacy with study results that meet What Works Clearinghouse evidence standards and shows positive or potentially positive findings.[2]

This document provides a summary table of research studies aligning to each ESSA evidence tier. Waterford supports states and districts as they implement ESSA and harness the power of personalized learning and data transparency to close the achievement gap for their youngest and most vulnerable learners.

Waterford Proven Efficacy

STRONG, Tier 1 evidence in 9 random control trial studies, conducted as part of a U.S. Department of Education Investing in Innovation grant. These studies show the substantively important effects of Waterford programs on at-risk learners, as well as the impact of innovative assessment strategies.

MODERATE, Tier 2 evidence in 10 quasi-experimental studies showing substantively important effects of Waterford technologies on foundational reading skills for preschool and kindergarten learners.

PROMISING, Tier 3 evidence in 17 correlation studies, showing substantively important effects of Waterford Early Learning (Reading, Math & Science) on key learning domains in a variety of instructional settings, locations, and populations.

DEMONSTRATES A RATIONALE, Tier 4 evidence in 34 case studies, often conducted in partnership with a district or state agency that chooses to use historical controls in a study designed to serve as many students as possible.

1. See ESSA Section 8101(21)(A)
2. What Works Clearinghouse. (2007, July 16). WWC Intervention Report: Waterford Early Reading Program. Retrieved from What Works Clearinghouse: ies.ed.gov/ncee/wwc/Docs/InterventionReports/WWC_Waterford_071607.pdf

Waterford.org

This table provides a summary of research studies aligning to each ESSA evidence tier.

1—STRONG
Meets WWC Evidence Standards with Reservations*

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Hobbs, L. J., & Overby, M. (2019a). Impact of UPSTART Reading Participation for Rural Students While Controlling for Prior Reading Achievement and School District.	Waterford Early Learning (UPSTART)	Randomized Controlled Trial	Brigance Inventory of Early Development III Preschool Early Literacy Indicators (PELI)	Students served: 491 Pre K Students Analytic Matched Sample—Brigance & PELI: • Treatment (students assigned to UPSTART Reading), n=252 • Control (students assigned to UPSTART Math/Science), n=239	The Rural UPSTART program prepares children for success upon entering kindergarten by providing computer-adaptive reading curriculum to pre-kindergarteners (Hobbs & Overby, 2019). This RCT study sampled 491 preschoolers from 13 of the most rural school districts in Utah during the 2014-2015 school year. Students were randomly assigned to receive either the UPSTART Reading program (the treatment group) or the UPSTART Math/Science program (the control group). The treatment group significantly outperformed the control group on six (identifying uppercase letters, reciting the alphabet, phonological awareness, phoneme manipulation, word recognition, and reading words from common signs) of the eight subtests of the Brigance and the initial word sounds subtest on the PELI. The UPSTART Reading program improved foundational literacy skills in treatment students, with meaningful effect sizes for phonological awareness ($d = 0.30$ to 0.32), letter knowledge ($d = 0.21$ to 0.51), and decoding ($d = 0.22$ to 0.49).
Shamir, H., Miner, C., Izzo, A., Feehan, K., Yoder, E., & Pocklington, D. (2019). Improving early literacy skills using technology at home. International Journal of Learning and Teaching, 5(3), 191-197.	Waterford Early Reading	Randomized Controlled Trial	Waterford Assessments of Core Skills (WACS)	523 four-year old PreK students from 13 rural Utah districts • Treatment: Pre-kindergarten students randomly assigned to use ERP (n=273) • Control: Pre-kindergarten students randomly assigned to use EMS (n=250)	Analysis of Overall WACS end of year scores, while covarying for beginning of year scores, revealed a statistically significant and positive effect of Waterford Early Reading on students in the Waterford treatment group. Experimental group students significantly outperformed control group students across demographics, including gender, experiencing poverty, special education services, and whether students attended another preschool. The interaction between race/ethnicity and treatment was not significant, which indicates that the Waterford reading program had a similar impact on WACS scores for Caucasian/White and Latino/a students.
Shamir, H., Yoder, E., Feehan, K., & Pocklington, D. (2019). Randomized controlled trial of kindergarten students using literacy technology. In Smart Education and eLearning 2019 (pp. 243-252). Springer, Singapore.	Waterford Early Learning	Randomized Controlled Trial	Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP)	• Treatment: Kindergarten students randomly assigned to use WEL (n=217) • Control: Kindergarten students randomly assigned to receive traditional literacy instruction (n=213)	In this randomized controlled trial, kindergarten classes were randomly assigned to either use WEL or receive the same amount of traditional, teacher-directed literacy instruction. Students who used WEL during their kindergarten year outperformed their control counterparts on all end of year literacy strands. Across demographics, experimental group students outperformed control students on all MAP strands and TRC across race/ethnicity and lunch status. Additional analysis demonstrated that African American/Black students in the experimental group significantly outperformed their control group counterparts on MAP subskills for Reading RIT Score, Reading Foundations, Reading Literature & Nonfiction, and Reading Vocabulary, with medium effects ranging from 0.58 to 0.68.
Evaluation and Training Institute (2023a). Improving early math skills: UPSTART program effects from pre-kindergarten to kindergarten.	Waterford Early Learning (UPSTART)	Randomized Controlled Trial	Kaufman Test of Educational Achievement (KTEA-3) and Children's Math Concepts and Applications test	367 PreK Students Treatment sample: n=176; English = 84 and DLL = 92 Control sample: n=191; English = 101 and DLL = 90	UPSTART Math had a significant impact on children's early math skill development, with treatment students outperforming the control group on math concepts and applications in the initial posttest (effect size, $g = 0.48$). Children in the treatment condition continued to outperform their control group peers on math concepts and applications at the end of kindergarten, nearly one year following the intervention (effect size, $g = 0.35$). At both posttests, there was no significant interaction effect between the treatment condition and whether or not students were multi-lingual learners, indicating that both mono- and multi-lingual learners benefited from the use of the program.
Evaluation and Training Institute (2023b). Preparing Students for Kindergarten: Waterford Upstart Summer Learning Path Program Effects.	Waterford Upstart Summer Learning Path (SLP) Program	Randomized Controlled Trial	Kaufman Test of Educational Achievement (KTEA-3)	329 PreK Students. 166 randomly assigned to the SLP reading program. 163 randomly assigned to the SLP math program.	Students who were assigned to the SLP math program scored significantly higher than the students who were assigned to the SLP reading program on measures of early math skills at the end of the program, with effect sizes ranging from $g = 0.25$ (Concepts and Application) to $g = 0.37$ (Core Composite). Math treatment students who began the program with lower math scores (low performance) had a larger average gain in math learning by the end of the program.

Waterford.org

This table provides a summary of research studies aligning to each ESSA evidence tier.

1—STRONG
Meets WWC Evidence Standards with Reservations*

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Evaluation and Training Institute (2020d). EIR Year 1 summary outcomes and fidelity results	Waterford Early Learning (UPSTART)	Randomized Controlled Trial	Kaufman Test of Educational Achievement Third Edition (KTEA-3) Social Skills Improvement System Rating Scales (SSIS RS)	Experimental group: assigned reading (n=134). Control group: assigned math (n=145).	Treatment children outperformed control children on the literacy subscales of the KTEA with effect sizes ranging from ES = 0.32 (Reading Comprehension) to ES = 0.56 (Letter-Word Recognition), while control children outperformed their treatment counterparts on the KTEA subscale of Math Concepts and Applications (ES = 0.55).
Evaluation and Training Institute (2021). EIR Great Plains Task Force Grant Year 2 Summary Report	Waterford Early Learning (UPSTART)	Randomized Controlled Trial	Waterford Early Learning (UPSTART)	Experimental group: assigned reading (n=319). Control group: assigned math (n=332).	Treatment children outperformed control children on the literacy subscales of the KTEA with effect sizes ranging from ES = 0.33 (Reading Comprehension) to ES = 0.39 (Letter and Word Recognition), while control children outperformed their treatment counterparts on the KTEA subscale of Math Concepts and Applications (ES = 0.33).
Evaluation and Training Institute (2022a). EIR Year 3 Summary: Implementation and Preliminary Outcomes	Waterford Early Learning (UPSTART)	Randomized Controlled Trial	Kaufman Test of Educational Achievement Third Edition (KTEA-3) Social Skills Improvement System Rating Scales (SSIS RS)	Experimental group: assigned reading (n=168). Control group: assigned math (n=184).	Treatment children outperformed control children on the literacy subscales of the KTEA with effect sizes ranging from ES = 0.27 (Reading Comprehension) to ES= 0.42 (Phonological Processing) while control children outperformed their treatment counterparts on the KTEA subscale of Math Concepts and Applications (ES = 0.41).
Hobbs, L. J., & Coordt, A. (2021). Impact of Upstart participation on Nevada pre-kindergarteners' school readiness.	Waterford Early Learning (UPSTART)	Randomized Controlled Trial	Kaufman Test of Educational Achievement Third Edition (KTEA-3) Social Skills Improvement System Rating Scales (SSIS RS)	Experimental group: assigned reading (n=272) of which 130 were multi-lingual learners. Control group: assigned math (n=278) of which 135 were multi-lingual learners	Treatment children outperformed control children on the literacy subscales of the KTEA with effect sizes ranging from g = 0.26 (Phonological Processing) to g = 0.44 (Reading Comprehension), while control children outperformed their treatment counterparts on the KTEA subscale of Math Concepts and Applications (effect size, g = 0.44). The two groups scored equally well on assessment measures of social-emotional development, and no significant interactions between MLL and non-MLL treatment children were found, indicating that all students that used the program benefitted from it in comparable terms. In addition, after posttest a total of 55% of at-risk students in the treatment group moved out of the risk category compared to 38% of the control group.

This table provides a summary of research studies aligning to each ESSA evidence tier.

2—MODERATE
Meets WWC Evidence Standards with Reservations*

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Hecht, S. A. & Close, L. (2002). Emergent literacy skills and training time uniquely predict variability in responses to phonemic awareness training in disadvantaged kindergartners. Journal of Experimental Child Psychology, 82, 93-115	Waterford Early Reading	Quasi-Experimental Study	• Wide Range Achievement Test • Stanford-Binet • Stones—Concepts About Print Test • Woodcock-Johnson Tests of Achievement (Form B) • Comprehensive Test of Phonological Processing"	• Treatment: Used Waterford Early Reading), n=42; control (did not use Waterford Early Reading), n=34	Analysis of pre-literacy gains over the course of their kindergarten year showed that the Waterford treatment group significantly outperformed the comparison group in Phonemic Awareness, Invented Spelling, and Word Reading.
Evaluation and Training Institute (2016). UPSTART Program Evaluation: Year 6 Program Results.	Waterford Early Learning (UPSTART)	Quasi-Experimental Study	• Brigance Inventory of Early Development III (Brigance) • Bader Reading and Language Inventory (Bader)	Students served: 5,091 PreK Students Analytic Matched Sample—Brigance & Bader: • Treatment: Students enrolled in UPSTART (n=138) • Control: Students who did not enroll in UPSTART, (n=138)	Combined post-test results showed that UPSTART participation using Waterford Early Learning had a large impact on students' early literacy skills. Large effect sizes (Bader = 0.95; Brigance = 0.81) were shown favoring UPSTART students as measured by the total Bader and Brigance composite scores. Favoring the UPSTART treatment group, differences in growth rates between the UPSTART treatment and control group were significantly different for the overall Brigance and for five of the Brigance subtests and for the Total Bader and all three Bader subtests. Children participating in UPSTART demonstrated improvement in word decoding and phonological awareness skills. Medium effect sizes were observed for Survival Sight Words (0.45), Rhyme Recognition (0.44). Large effect sizes were found for Pre-Primer Vocabulary (1.10), Phonemic Blending (0.99), and Phoneme Segmenting (0.85). Children participating in UPSTART also demonstrated greater gains on Pre-Primer Vocabulary, Survival Sight Words, and all Phonological Awareness subscales than control students. Participation in UPSTART was associated with improvement in all phonological awareness strands of the Bader, including Rhyme Recognition, Phoneme Blending, and Phoneme Segmenting. Children participating in UPSTART had a 36 point advantage on Brigance post-test scores compared to non-participating children.
Evaluation and Training Institute (2017). UPSTART Program Evaluation: Year 7 Program Results.	Waterford Early Learning (UPSTART)	Quasi-Experimental Study	• Brigance Inventory of Early Development III (Brigance) • Bader Reading and Language Inventory (Bader)	Students served: 6,639 PreK Students Analytic Matched Sample—Brigance & Bader: • Treatment: Students enrolled in UPSTART (n=208) • Control: Students who did not enroll in UPSTART (n=208)	Combined post-test results showed that UPSTART participation had a meaningful impact on students' early literacy skill development: Children enrolled in UPSTART produced significant positive effects (ES = 0.52) compared to control children on the Brigance composite. Similarly, UPSTART participants experienced significant positive effects (ES = 0.62) on the Bader composite. Children participating in UPSTART demonstrated significant improvement in word decoding and phonological awareness skills. Medium effect sizes were observed for Pre-Primer Vocabulary (0.74), Phoneme Segmenting (0.64), and Phoneme Blending (0.63). Children participating in UPSTART also demonstrated greater gains on Pre-Primer Vocabulary, Survival Sight Words, and all Phonological Awareness subtests than control students. Participation in UPSTART was associated with improvement in all phonological awareness strands of the Bader, including Rhyme Recognition, Phoneme Blending, and Phoneme Segmenting. Children participating in UPSTART had a 21-point advantage on Brigance post-test scores compared to non-participating children.

Waterford.org

This table provides a summary of research studies aligning to each ESSA evidence tier.

2—MODERATE
Meets WWC Evidence Standards with Reservations*

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Evaluation and Training Institute (2018a). UPSTART Program Evaluation: Year 8 Program Results.	Waterford Early Learning (UPSTART)	Quasi-Experimental Study	• Brigance Inventory of Early Development III (Brigance) • Bader Reading and Language Inventory (Bader) • Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Next	• Students served: 10,745 PreK Students Analytic Matched Sample - Brigance & Bader: • Treatment (students enrolled in UPSTART), n=245 • Control (students who did not enroll in UPSTART), n=245 Analytic Matched Sample - DIBELS: • Treatment (1st grade students who were enrolled in UPSTART during PreK), n=2,701 • Control (1st grade students who did not enroll in UPSTART during PreK), n=2,701	Combined post-test results showed that UPSTART participation had a meaningful impact on students' early literacy skill development. In the matched post-test sample, UPSTART produced strong to medium effects: Children enrolled in UPSTART produced significant positive effects (ES = 0.50) compared to control children on the Brigance composite. Similarly, UPSTART participants experienced significant positive effects (ES = 0.81) on the Bader composite. Children participating in UPSTART demonstrated significant improvement in word decoding and phonological awareness skills. Medium effect sizes were observed for Phonemic Blending (0.78), Phoneme Segmenting (0.64), and Pre-Primer Vocabulary (0.60). Children participating in UPSTART also demonstrated greater gains on both Phonemic Blending and Phoneme Segmenting subscales than control students. Participation in UPSTART was associated with significant improvement on both of the phonological awareness strands of the Bader assessed, Phoneme Blending and Phoneme Segmenting. Children participating in UPSTART had a significantly higher growth rate, with a 21-point advantage on Brigance post-test scores compared to non-participating children. First grade students who were enrolled in UPSTART during pre-kindergarten significantly outperformed control students on beginning of year DIBELS scores. The average beginning of year DIBELS composite score was 7.91 points higher for students who were enrolled in UPSTART compared to students who were not (ES = 0.18).
Evaluation and Training Institute (2018b). Utah High-Quality School Readiness Expansion (HQSR-E) Program Evaluation.	Waterford Early Learning (UPSTART)	Quasi-Experimental Study	• Brigance Inventory of Early Development III (Brigance) • Bader Reading and Language Inventory (Bader) • PELI	Cohort 1 • 75 public, 58 private, 93 UPSTART, and 134 students not enrolled in a high-quality school readiness program Cohort 2 • 99 public, 65 private, 101 UPSTART, and 112 students not enrolled in a high-quality school readiness program	77% of UPSTART children had post-test literacy quotients of average or above average, representing a greater level of school readiness than was achieved by either the other intervention groups or children not participating in high-quality school readiness programs. UPSTART children outperformed children not participating in high-quality school readiness programs on Overall Literacy test scores, and subtest scores for UPSTART children were significantly higher in Letter Knowledge, Listening Comprehension, and Phonological Awareness. By the end of the program year, social emotional development (SED) was similar for all three treatment groups.

2—MODERATE

Meets WWC Evidence Standards with Reservations*

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Evaluation and Training Institute (2019a). UPSTART Program Evaluation: Year 9 Program Results.	Waterford Early Learning (UPSTART)	Quasi-Experimental Study	• Brigance Inventory of Early Development III (Brigance) • Bader Reading and Language Inventory (Bader) • Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Next	• Students served: 14,278 Analytic Matched Sample - Brigance • Treatment (students enrolled in UPSTART), n=248 • Control (students who did not enroll in UPSTART), n=248 Analytic Matched Sample - Bader: • Treatment (students enrolled in UPSTART), n=215 • Control (students who did not enroll in UPSTART), n=215 Analytic Matched Sample - DIBELS: • Treatment (1st grade students who were enrolled in UPSTART during Pre-K), n=3,503 • Control (1st grade students who did not enroll in UPSTART during Pre-K), n=3,503	Combined post-test results showed that UPSTART participation had a meaningful impact on students' early literacy skill development. In the matched post-test sample, UPSTART produced strong to medium effects: Children enrolled in UPSTART produced significant positive effects (ES = 0.53) compared to control children on the Brigance composite. Similarly, UPSTART participants experienced significant positive effects (ES = 0.56) on the Bader composite. Children participating in UPSTART demonstrated significant improvement in word decoding and phonological awareness skills. Medium effect sizes were observed for Phonemic Blending (0.71), Pre-Primer Vocabulary (0.67), and Phoneme Segmenting (0.48). Participation in UPSTART was associated with significant improvement on both of the phonological awareness strands of the Bader assessed, Phoneme Blending and Phoneme Segmenting. Children participating in UPSTART had a significantly higher growth rate, with a 23-point advantage on Brigance post-test scores compared to non-participating children. First grade students who were enrolled in UPSTART during pre-kindergarten significantly outperformed control students on beginning of year DIBELS scores. The average beginning of year DIBELS composite score was 7.97 points higher for students who were enrolled in UPSTART compared to students who were not (ES = 0.17).
Evaluation and Training Institute (2019b). Utah High-Quality School Readiness Expansion (HQSR-E) Program Evaluation.	Waterford Early Learning (UPSTART)	Quasi-Experimental Study	• Brigance Inventory of Early Development III (Brigance) • Bader Reading and Language Inventory (Bader) • PELI • Acadience Reading	134 UPSTART students, 77 Public Preschool, 77 Private Preschool, and 107 students not enrolled in a high-quality school readiness program Longitudinal: Treatment • Students who participated in UPSTART during Pre-School (n=1598), assessing literacy during kindergarten and first grade Control • Students who did not participate in UPSTART during Pre-School (n=1598), assessing literacy during kindergarten and first grade	UPSTART students were the only ones to significantly outperform control students on the overall Brigance Global Literacy Composite (ES =.383), and the only group to outperform control students on phonological awareness (ES = 0.462), and letter knowledge (ES = 0.270), as measured by the Bader and Brigance, respectively. UPSTART students also had the highest percentage of students with PELI listening comprehension scores at or above benchmark (79%) and were the only group to outperform the comparison group. By the end of the program year, social emotional development (SED) was similar for all three treatment groups. Students who participated in UPSTART during preschool had significantly higher literacy scores at the beginning, middle, and end of kindergarten, and at the beginning of first grade, compared to students who did not participate in UPSTART. This demonstrated a long-term effect of the Waterford UPSTART program.

Waterford.org

This table provides a summary of research studies aligning to each ESSA evidence tier.

2—MODERATE

Meets WWC Evidence Standards with Reservations*

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Evaluation and Training Institute (2020a). UPSTART Program Evaluation: Year 10 Program Results.	Waterford Early Learning (UPSTART)	Quasi-Experimental Study	• Brigance Inventory of Early Development III (Brigance) • Bader Reading and Language Inventory (Bader) • Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Next	• Students served: 14,125 Analytic Matched Sample - Brigance • Treatment (students enrolled in UPSTART), n=225 • Control (students who did not enroll in UPSTART), n=225 Analytic Matched Sample - Bader: • Treatment (students enrolled in UPSTART), n=237 • Control (students who did not enroll in UPSTART), n=237 Analytic Matched Sample - DIBELS: • Treatment (1st grade students who were enrolled in UPSTART during Pre-K), n=4,012 • Control (1st grade students who did not enroll in UPSTART during Pre-K), n=4,012	Combined post-test results showed that UPSTART participation had a meaningful impact on students' early literacy skill development. In the matched post-test sample, UPSTART produced strong to medium effects: Children enrolled in UPSTART produced significant positive effects (ES = 0.53) compared to control children on the Brigance composite. Similarly, UPSTART participants experienced significant positive effects (ES = 0.35) on the Bader composite. Children participating in UPSTART demonstrated significant improvement in word decoding and phonological awareness skills. Medium effect sizes were observed for Phoneme Segmenting (0.71), Phonemic Blending (0.68), and Pre-Primer Vocabulary (0.59). Participation in UPSTART was associated with significant improvement on both of the phonological awareness strands of the Bader assessed, Phoneme Blending and Phoneme Segmenting. Children participating in UPSTART had a significantly higher growth rate, with a 21-point advantage on Brigance post-test scores compared to non-participating children. First grade students who were enrolled in UPSTART during pre-kindergarten significantly outperformed control students on beginning of year DIBELS scores. The average beginning of year DIBELS composite score was 10.67 points higher for students who were enrolled in UPSTART compared to students who were not (ES = 0.22).
Evaluation and Training Institute (2020b). Indiana UPSTART Evaluation Preschool Impact Study	Waterford Early Learning (UPSTART)	Quasi-Experimental Study	• Brigance Inventory of Early Development III (Brigance) • Social Skills Improvement System (SSIS)	Analytic Matched Sample— Brigance & SSIS: • Treatment: Students enrolled in UPSTART (n=51) • Control: Students who did not enroll in UPSTART (n=73)	Treatment group students significantly outperformed control group students across literacy strands, with effect sizes ranging from medium (Visual Discrimination, ES = 0.29) to large (Word Recognition, ES = 0.86), demonstrating that the UPSTART program can positively impact students' literacy skills before kindergarten entry. While treatment group students did not significantly outperform their control counterparts on the Brigance Math Composite, treatment and control group students performed similarly on social skill scale scores, indicating that students enrolled in UPSTART and their control counterparts experienced similar growth in their social skills over the course of the program.
Hobbs, L. J., & Overby, M. (2019b). Impact of the UPSTART Program on Forestalling Summer Learning Loss. In ERIC. https://eric.ed.gov/?id=ED605302	Waterford Upstart Summer Learning Path (SLP) Program	Quasi-Experimental Study	• Acadience Reading Test	Cohort 1- Treatment: (n=1,611), Control: (n=1,611) Cohort 2- Treatment: (n=898), Control: (n=898) Cohort 3-Treatment (n=429), Control: (n=429)	Regression analysis revealed that Upstart Summer program (SLP) participation was a significant predictor in reducing overall literacy learning loss (p = 0.003) for students in Cohort 1. Treatment students had a higher average increase than the control students on Reading Composite Scale (g = 0.22), NWF-CL (g = 0.32), and LNF (g = 0.17) test scores. Participation in the Upstart Summer program for Cohort 2 and Cohort 3 was not a significant predictor of forestalling learning loss.

This table provides a summary of research studies aligning to each ESSA evidence tier.

3—PROMISING

Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Powers, S., & Price-Johnson, C. (2006). Evaluation of the Waterford Early Reading Program in Kindergarten, 2005-06. Online Submission.	Waterford Early Reading	Correlational Study (with statistical controls for selection bias)	• Dynamic Indicators of Basic Early Literacy Skills (DIBELS) • Core Curriculum Standard Assessment (CCSA)	1838 kindergarten children in 15 classes • Treatment: Students who used Waterford (n=358) • Control: Students who did not use Waterford) (n=1480)	The students who used Waterford Early Learning significantly outperformed control students on both the DIBELS (effect size = 0.42) and CCSA (effect size = 0.28) tests. MLL students in the treatment group demonstrated greater gains than the non-MLL group in the comparison schools (F[1, 1045] = 8.62, p = .003).
Shamir, H., Feehan, K., & Yoder, E. (2016). Using technology to improve reading and math scores for the digital native. In Proceedings of EdMedia 2016—World Conference on Educational Media and Technology (pp. 1405-1412). Vancouver, BC, Canada: Association for the Advancement of Computing in Education (AACE).	• Waterford Early Reading • Waterford Early Math and Science	Correlational Study (with statistical controls for selection bias)	Florida Voluntary Pre-Kindergarten (VPK) Assessment	Early Reading Program • Treatment: Usage over 900 minutes throughout the school year (n=653) • Control: Usage less than 300 minutes throughout the school year (n=67) Early Math and Science • Treatment: Usage over 1,000 minutes throughout the school year (n=183) • Control: Usage less than 300 minutes throughout the school year (n=372)	Treatment students significantly outperformed control students on end of year scores while covarying for beginning of year scores on all strands of the VPK. Effect sizes are substantively important for each strand, including Oral Language Vocabulary (0.72), Phonological Awareness (1.32), Print Knowledge (1.12), and Math (0.77). Students with MLL and non-MLL status, as well as African American/Black and Latino/a students in the experimental group, significantly outperformed their control group counterparts across all strands. The interaction between special education services and treatment was not significant across all strands, which indicates that WEL had a similar impact on all strands for students with and without special education services.
Shamir, H. & Goethe, R. (2015). Common Core State Standards: Is computer assisted instruction up for the challenge? In Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2015 (pp. 220-227). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE).	• Waterford Early Reading	Correlational Study (with statistical controls for selection bias)	AIMSweb	Kindergarten • Treatment: Used Waterford Early Reading for more than 700 minutes during the school year; had assessment scores in at least 2/3 administrations (n=41) • Control: Did not use Waterford Early Reading; had assessment scores in at least 2/3 administrations (n=77) First Grade • Treatment: Used Waterford Early Reading for more than 700 minutes during the school year; had assessment scores in at least 2/3 administrations (n=44) • Control: Did not use Waterford Early Reading; had assessment scores in at least 2/3 administrations (n=50)	Students that used Waterford Early Reading significantly outperformed the comparison group on end of year scores covarying for beginning or middle of year scores (depending on the skill) on three of the four sub-strands: Letter Sound Fluency (LSF), Letter Name Fluency (LNF), and Nonsense Word Fluency (NWF).

This table provides a summary of research studies aligning to each ESSA evidence tier.

3—PROMISING

Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Feehan, K., & Yoder, E. (2017b). Does CAI improve early math skills? In Proceedings of the 9th International Conference on Computer Supported Education (CSEDU 2017) - Volume 2 (pp. 285-292). Porto, Portugal: Institute for Systems and Technologies of Information, Control and Communication.	Waterford Early Math and Science	Correlational Study (with statistical controls for selection bias)	mClass: Math	Kindergarten • Treatment: Used Early Math and Science (n=114) • Control: Did not use Early Math and Science (n=58) First Grade • Treatment: Used Early Math and Science (n=68) • Control: Did not use Early Math and Science (n=255)	Students in the Kindergarten treatment group significantly outperformed control group students in end of year scores covarying for beginning of year scores with substantively important effect sizes in Number Identification (effect size = 0.33) and Quantity Discrimination (effect size = 0.29). For first grade, effect sizes were substantively important in Number Facts, Quantity Discrimination, Missing Number, and Next Number strands. For Kindergarten and first grade, students in the treatment group outperformed students in the control group across demographics including gender, free/reduced lunch status, and special education services.
Evaluation and Training Institute (2012). Utah UPSTART Education Program Evaluation: Program Impacts on Early Literacy. Cohort 2 Results: Technical Report.	Waterford Early Learning (UPSTART)	Correlational Study (with statistical controls for selection bias)	• Brigance Inventory of Early Development • Bader Reading and Language Inventory	Students served: 1,018 PreK Students Analytic Matched Sample—Brigance: • Treatment: Students enrolled in UPSTART (n=77) • Control: Students who did not enroll in UPSTART (n=82) Analytic Matched Sample—Bader: • Treatment: Students enrolled in UPSTART (n=76) • Control: Students who did not enroll in UPSTART (n=82)	The UPSTART treatment group performed significantly better than the control group on the Total Brigance post-test, with an average difference of 7.9 points. The UPSTART treatment group also performed significantly better than the control group on the overall Bader and the Bader Phoneme Blending subtest. Growth rates between the treatment group and the control group were significantly different at the 99% Confidence Interval (CI) for the Total Brigance and the Total Bader. Children participating in UPSTART demonstrated moderately strong improvements in literacy skills on the Brigance and small improvements in literacy skills on the Bader compared to control children on the Brigance. Children participating in UPSTART scored higher on strands for Lower Case Letters and Lower Case Letter Sounds even when taking initial literacy skills into account. Improvements on the Bader were driven by performance on the Phoneme Blending subtest.
Evaluation and Training Institute (2013). Utah UPSTART Program Evaluation: Program Impacts on Early Literacy. Third Year Results: Cohort 3 Technical Report.	Waterford Early Learning (UPSTART)	Correlational Study (with statistical controls for selection bias)	• Brigance Inventory of Early Development • Bader Reading and Language Inventory	Students served: 1,168 PreK Students Analytic Matched Sample—Brigance: • Treatment: Students enrolled in UPSTART (n=129) • Control: Students who did not enroll in UPSTART (n=130) Analytic Matched Sample—Bader: • Treatment: Students enrolled in UPSTART (n=112) • Control: Students who did not enroll in UPSTART (n=120)	The UPSTART treatment group significantly outperformed the control group on both the Total Brigance and Total Bader. Effect sizes range from 0.33 to 0.85. Favoring the UPSTART treatment group, growth rates between the treatment group and the control group were significantly different at the 99% Confidence Interval (CI) for the Total Brigance and the Total Bader. Consistent with prior cohorts, children participating in UPSTART demonstrated moderately strong improvements in measures of phonics skills: At the beginning of kindergarten, improvement was observed across all but one strand of the Brigance. Participation in UPSTART was associated with improvement in all phonological awareness strands of the Bader, including Rhyme Recognition, Phoneme Blending, and Phoneme Segmenting. Children participating in UPSTART had a 28 point advantage on Brigance post-test scores compared to non-participating children.

This table provides a summary of research studies aligning to each ESSA evidence tier.

3—PROMISING

Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Evaluation and Training Institute (2014). Utah UPSTART Program Evaluation: Program Impacts on Early Literacy. Fourth Year Results: Cohort 4 Technical Report.	Waterford Early Learning (UPSTART)	Correlational Study (with statistical controls for selection bias)	• Brigance Inventory of Early Development • Bader Reading and Language Inventory	Students served: 1,250 PreK Students Analytic Matched Sample—Brigance: • Treatment: Students enrolled in UPSTART (N=101 • Control: Students who did not enroll in UPSTART) (N=102 Analytic Matched Sample—Bader: • Treatment: Students enrolled in UPSTART (n=79) • Control: Students who did not enroll in UPSTART (n=93)	The UPSTART treatment group performed significantly better than the control group on both the Total Brigance and Total Bader. Effect sizes range from 0.34 to 0.59. Favoring the UPSTART treatment group, growth rates between the treatment group and the control group were significantly different at the 99% Confidence Interval (CI) for the Total Brigance and the Total Bader. Preschool children participating in UPSTART improved phonics skills with small to medium effect sizes, as demonstrated on the Overall Brigance Composite as well as on Visual Discrimination, Letter Sounds, and Basic Pre-Primer Vocabulary subtests. Children participating in UPSTART had a 29 point advantage on Brigance post-test scores compared to non-participating children.
Evaluation and Training Institute (2015). Utah UPSTART Program Evaluation: Program Impacts on Early Literacy. Year 5 Results: Cohort 5 Technical Report.	Waterford Early Learning (UPSTART)	Correlational Study (with statistical controls for selection bias)	• Brigance Inventory of Early Development • Bader Reading and Language Inventory	Students served: 1,577 PreK Students Analytic Matched Sample—Brigance: • Treatment • Students enrolled in UPSTART (n=94) • Control • Students who did not enroll in UPSTART (n=100) Analytic Matched Sample—Bader: • Treatment: Students enrolled in UPSTART (n=89) • Control: Students who did not enroll in UPSTART (n=100)	The UPSTART treatment group performed significantly better than the control group on the Total Brigance and Total Bader. Effect sizes range from 0.27 to 0.85. Favoring the UPSTART treatment group, growth rates between the UPSTART treatment group and the control group were significantly different at the 99% Confidence Interval (CI) for three of the Brigance subtests, and for the Total Bader and for two of the Bader subtests: Phoneme Blending and Phoneme Segmenting. Children participating in UPSTART demonstrated small overall improvements in phonics skills when assessed on the Brigance, indicated by a 12 point advantage on the Overall Brigance Composite compared to control children. Additionally, large effects were found for Pre-Primer Vocabulary. Participation in UPSTART was associated with significant improvement in two of the three phonological awareness strands of the Bader, including Phoneme Blending and Phoneme Segmenting.
Shamir, H., Pocklington, D., Feehan, K., & Yoder, E. (2018). Educational equity using computer-assisted instruction. In E. Langran & J. Borup (Eds.), Proceedings of Society for Information Technology & Teacher Education International Conference (pp. 717-722). Washington, D.C.: Association for the Advancement of Computing in Education (AACE).	Waterford Early Learning	Correlational Study (with statistical controls for selection bias)	Developmental Reading Assessment (DRA)	Kindergarten • Treatment: Used WEL for more than 2,000 minutes (n=699) • Control Used WEL for less than 500 minutes (n=40) First Grade • Treatment: Used WEL for more than 2,000 minutes (n=777) • Control: Used WEL for less than 500 minutes (n=61)	Significant, positive results were found in both kindergarten and first grade. Significant differences in usage of WEL were found between races/ethnicities despite equal access in the school district: African American/Black and Latino/a students in the experimental group scored higher than those in the control group; however, the difference was significant for only African American/Black students, who used the program significantly more than Latino/a students. Examination of available demographics indicated that students of all genders and races/ethnicities benefited from WEL.

This table provides a summary of research studies aligning to each ESSA evidence tier.

3—PROMISING
Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Pocklington, D., Feehan, K., & Yoder, E. (2019c). Game-based learning for young learners. International Journal of Learning and Teaching, 5(3), 206-212.	Waterford Early Learning	Correlational Study (with statistical controls for selection bias)	District-Administered Literacy Assessment	Kindergarten • Used WEL for more than 2,000 minutes (n=967) • Control: Used WEL for less than 500 minutes (n=723) First grade • Treatment: Used WEL (n=4,032) • Did not use WEL (n=1,680) Second grade • Treatment: used WEL (n=4,018) • Control: Did not use WEL(n=2,887)	This study supports the hypothesis that the use of WEL in a classroom setting can have a positive effect on learning in elementary school students. Students in kindergarten, first grade, and second grade who used WEL as part of their curriculum had higher literacy end of year scores, while covarying for beginning of year scores, compared to students who were exposed only to traditional classroom instruction. Experimental students outperformed control counterparts across genders and races/ethnicities as well.
Shamir, H., Pocklington, D., Feehan, K., & Yoder, E. (2019a). Bridging the achievement gap for low-performing students using computer-adaptive instruction. International Journal of Information and Education Technology, 9(3), 196-200.	Waterford Early Learning	Correlational Study (with statistical controls for selection bias)	Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP)	Kindergarten • Treatment: Used WEL (n=88) • Control: Did not use WEL (n=100) First grade • Treatment: Used WEL (n=78) • Control: Did not use WEL (n=79)	Kindergarten and first graders who used WEL scored significantly higher on end of year NWEA MAP compared to Kindergarteners and first graders who did not use WEL when covarying for beginning of year scores. Students in experimental group scored lower than the control group on the BOY assessment but surpassed the control group on the EOY assessment.
Shamir, H., Pocklington, D., Feehan, K., & Yoder, E. (2020). Using CAI with Fidelity: Impacts on Literacy Skills of Kindergarten Students Across Demographics. In D. Schmidt-Crawford (Ed.), Proceedings of Society for Information Technology & Teacher Education International Conference (pp. 832-837). Online: Association for the Advancement of Computing in Education (AACE).	Waterford Early Learning	Correlational Study (with statistical controls for selection bias)	Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP)	Kindergarten • Treatment: Used WEL for more than 1,000 minutes (n=101) • Control: Used WEL for less than 300 minutes (n=41)	Students in the experimental group scored significantly higher than the control group on Reading RIT (ES = 0.68), Reading Foundations (ES = 0.80), Reading Literature and Nonfiction (ES = 0.69), Reading Vocabulary, (ES = 0.63) and Writing (ES = 0.65) strand scores. Across all strands, for students with special education services and no special education services, as well as free/reduced lunch status and paid lunch status, students in the experimental group outperformed students in the control group.

This table provides a summary of research studies aligning to each ESSA evidence tier.

3—PROMISING
Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Pocklington, D., Yoder, E., & Feehan, K. (2019). Reinforcing second grade literacy skills using a computer-adaptive reading program. In EdMedia+ Innovate Learning (pp. 1412-1417). Association for the Advancement of Computing in Education (AACE).	Waterford Early Learning	Correlational Study (with statistical controls for selection bias)	Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP)	Second grade • Treatment: Used WEL for more than 2,000 minutes (n=778) • Control: Used WEL for less than 300 minutes (n=784)	Second grade students with high usage of WEL outperformed their control counterparts on the MAP literacy assessment. After using WEL for only thirty minutes per day, five days per week, students who used WEL to fidelity outperformed their control counterparts across all literacy strands.
Shamir, H., Yoder, E., Pocklington, D., Wang, C., & Greene, E. (2023). Using technology to teach young learners math skills. In EDULEARN23 Proceedings (pp. 1693-1700). IATED.	Waterford Reading Academy (WRA)	Correlational Study (with statistical controls for selection bias)	Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP) Growth Math	559 Kindergarten Students. Students who used the software for at least 1,500 minutes during the school year were assigned to the experimental group (n=459) Students who used the software for less than 800 minutes were assigned to the control group (n=100)	Students in the experimental group scored significantly higher than the control group on Overall RIT Math (d = 0.38), Operations and Algebraic Thinking (d = 0.25), Numbers and Operations (d = 0.30), Measurement and Data (d = 0.37), and Geometry (d = 0.35). Scores for students with special education services were significantly higher in the experimental group than the control group for all subskills.
Shamir, H., Pocklington, D., & Yoder, E. (2023). Using computer-assisted instruction to fidelity to combat learning loss in early learners. American Educational Research Association (AERA) Conference Proceedings 2023	Waterford Upstart Summer Learning Path (SLP) Program	Correlational Study (with statistical controls for selection bias)	Waterford Assessments of Core Skills (WACS)	Initial sample, PreK students that used Waterford, n=3,082, Treatment (used Waterford for at least 1300 minutes) n=323, Control (used Waterford for less than 750 minutes) n=338	There was a significant difference in end of program scores between groups due to higher end of program scores made by experimental students than by control students (d = 0.31). For all comparisons between the experimental and control group, across all demographics, students in the experimental group consistently outperformed students in the control group.
Shamir, H., Ortiz-Wood, C., Pocklington, D. & Yoder, E. (2022). Can Children Achieve Literacy Gains During the Summer?. In T. Bastiaens (Ed.), Proceedings of EdMedia + Innovate Learning (pp. 711-718). Online: Association for the Advancement of Computing in Education (AACE).	Waterford Upstart Summer Learning Path (SLP) Program	Correlational Study (with statistical controls for selection bias)	Waterford Assessment of Core Skills (WACS) Waterford Reading Academy (WRA)	PreK students that used Waterford, n=3,082	Students achieved significantly higher WACS scores after their participation in the SLP program (d = 1.25). These positive results were seen across all demographics, including race/ethnicity, experience of poverty, gender, preschool attendance, MLL status, and location.

Waterford.org

This table provides a summary of research studies aligning to each ESSA evidence tier.

3—PROMISING
Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Yoder, E., & Pocklington, D. (2023a). Using CAI to Provide Early Literacy Instruction for All Learners. In: Uskov, V.L., Howlett, R.J., Jain, L.C. (eds) Smart Education and e-Learning—Smart University. KES-SEEL 2023. Smart Innovation, Systems and Technologies, vol 355. Springer, Singapore.	Waterford Reading Academy (WRA)	Correlational Study (with statistical controls for selection bias)	Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP) Growth Reading	Experimental group (CAI for more than 1,500 minutes): (n=540) Control group (CAI for less than 800 minutes): (n=66)	Treatment children scored significantly higher on literacy scores compared to the control group when assessed at the end of the school year with effect sizes ranging from ES = 0.50 (Literature and Informational Text) to ES = 0.61 (Overall RIT Reading). Students of all races/ethnicities and special education services in the experimental group scored on average higher on the end of year literacy assessment than their control group counterparts.

Waterford.org

This table provides a summary of research studies aligning to each ESSA evidence tier.

4—DEMONSTRATES A RATIONALE
Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Waterford Institute (1996). Preliminary research: Waterford Institute's Early Research in Utah and New York schools. Research Compendium, 1998.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation	Waterford Early Reading Instrument		In every case, classes that used the software made greater gains in pre-literacy skills than comparison classrooms. Waterford students at one elementary school improved reading test average scores from 50% to 91.8% over the course of the year—compared with score averages of 55% (pretest) to 73% (posttest) among the control group. In two schools, MLL classes performed better on posttests than the non-MLL classes in the control group (60% vs. 47% on the WERI in PS 43 and 85% vs. 68% in PS 1).
Waterford Institute (1997). A preliminary report of the 1996-97 test results from elementary schools in the Dallas ISD on the effectiveness of the Waterford Early Reading Program. Research Compendium, 1998.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation		668 Kindergarten children	After a one-year trial, researchers noted highly significant differences (p < 0.01) between students using the Waterford program and control classrooms.
Research, Assessment, & Measurement, Inc. (1999). Evaluation of Waterford Early Reading Program: Hacienda la Puente and Whittier School Districts. Los Angeles, California. Research Compendium, 1999.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation	Emergent Literacy Development Inventory (ELDI)	• Treatment: Used Waterford for approximately six months (n=381) • Control: Did not use Waterford (n=177)	The average growth scores for Waterford students were significantly higher than those of comparison classes (p < .001).
Hecht, S. A (2000). Waterford Early Reading Program in Ohio: An Evaluation. Research Compendium, 2000.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation			This study found significant gains among students using the program in comparison with the control group (p < .05) for skills including Letter-Word Identification, Spelling, and Phonological Awareness.
Young, J. W., & Tracey, D. H. (1999). An evaluation of the Waterford Early Reading Program: Newark, New Jersey. Research Compendium, 1998.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation	TERA-2	(n=265)	The Waterford software student group over the control group (p < .02) on the TERA-2 standardized assessment.
Reynolds, C. (2000). An evaluation of the Waterford Early Reading Program in the Decatur, Ill., school district: Analysis of impact on vocabulary development. Research Compendium, 2000.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation	Iowa Test of Basic Skills (ITBS)	• Treatment: Kindergarten and first grade students during two successive years (n=700)	First grade students who used the software significantly outperformed control group students on the Iowa Basic Test of Skills for reading (p = .003).

This table provides a summary of research studies aligning to each ESSA evidence tier.

4—DEMONSTRATES A RATIONALE
Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Walberg, H. J. (2001). Final evaluation of the reading initiative. Research Compendium, 2002.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation		• Treatment: Kindergarten students using Waterford (n=2414)	After a statewide implementation of the reading software in Idaho kindergartens, evaluators working in connection with the Albertson Foundation reported the effect size for students who had originally tested in the lowest third on standardized reading measures was 1.14, and the overall effect size for students who completed the program was 0.52.
Electronic Education (2002). Los Angeles Unified school district - Waterford Early Reading Program initial implementation findings. Research Compendium, 2002.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation	• Waterford Comprehensive Reading Inventory (WCRI) • Waterford Computer Adaptive Reading Test (WCART)	Matched sample: • Treatment: Used Waterford (n=33) • Control: Did not use Waterford) (N=27)	In a matched sample comparison, in which the majority of students were multi-lingual learners, students who used Waterford software achieved gains almost twice as high as students who did not use Waterford.
Research, Assessment, & Measurement, Inc. (1998). A study of the effectiveness of the Waterford program at Glenridge elementary school. Research Compendium, 1998.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation			The use of Waterford software increased multi-lingual learners' scores more than 600% (as compared to 283% for the control group).
Cassady, J. C., & Smith, L. L. (2003). The impact of a reading-focused integrated learning system on phonological awareness in kindergarten. Journal of Literacy Research, 35(4), 947-964.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation	Phonological Abilities Test (PAT)	Kindergarten • Treatment: Used Waterford (n=26) • Control: Did not using Waterford (n=62)	Despite no significant differences in pretest scores, students using Waterford software experienced a faster acquisition of phonological awareness skills than students who had not used the program, $F(2, 85) = 3.05$, $p < .05$, $\eta 2 = .07$.
Cassady, J. C. & Smith, L. L. (2005). The impact of a structured integrated learning system on first grade students' reading gains. Reading and Writing Quarterly, 21, 361-376.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation	CTBS Terra Nova	First grade • Treatment: Used Waterford in fall 2001 (n=46) • Control: Did not use Waterford in fall 2000 (n=47)	Students who used Waterford experienced significantly greater reading skill gains on the CTBS Terra Nova than the comparison group, $F(1, 91) = 10.61$, $p < .01$, $\eta 2 = .10$. The lowest performing students in the treatment group outperformed the low-performing comparison group ($F[1, 21] = 15.67$, $p < .01$, $\eta 2 = .43$). By the end of the first grade year, test scores among this at-risk group were equivalent to those of the moderate-performing students in the comparison group.

This table provides a summary of research studies aligning to each ESSA evidence tier.

4—DEMONSTRATES A RATIONALE

Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Feehan, K., & Yoder, E. (2017a). Computer-assisted instruction in early literacy for African American, economically disadvantaged children. Journal of Educational Multimedia and Hypermedia, 26(2), 179-192.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation	Standardized Test for the Assessment of Early Reading (STAR) Dynamic Indicators of Early Literacy Skills (DIBELS)	First sample group: (First grade) • Treatment: Waterford usage over 600 minutes (n=78) • Control: Did not use Waterford (n=28) Second sample group (K) • Treatment: Waterford usage over 1,000 minutes (n=105) • Control: Did not use Waterford (n=128) Third sample group (K) • Treatment: Waterford usage over 1,000 minutes (n=189 • Control: Did not use Waterford (n=94)	On all measured skills, students who used ERP outperformed the control groups. In the first sample group, the treatment group significantly outperformed the control group on six of the ten sub-strands; in the second sample group, the treatment group significantly outperformed the control group on seven of the ten sub-strands; and in the third sample group, the treatment group significantly outperformed the control group on two of the three sub-strands.
Shamir, H. & Goethe, R. (2015). Common Core State Standards: Is computer assisted instruction up for the challenge? In Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2015 (pp. 220-227). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE).	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation	Developmental Reading Assessment (DRA)	Kindergarten • Treatment: Used Waterford more than 470 minutes (n=12) • Control: Used an earlier version of Waterford (n=61)	The treatment students using ERP Cloud Version significantly outperformed the comparison group on three of the four sub-strands: Initial Sounds, Capital Letters, and Segmenting Phonemes.
Shamir, H., Feehan, K., & Yoder, E. (2017c). Effects of personalized learning on kindergarten and first grade students' early literacy skills. In Proceedings of the 9th International Conference on Computer Supported Education (CSEDU 2017) - Volume 2 (pp. 273-279). Porto, Portugal: Institute for Systems and Technologies of Information, Control and Communication.	Waterford Early Reading	Provides a well-specified logical model informed by research or evaluation	Developmental Reading Assessment (DRA)	Kindergarten • Treatment: Used Waterford for more than 1000 minutes throughout the school year (n=1004) • Control: Used Waterford for less than 500 minutes throughout the school yea (n=28) First Grade • Treatment: Used Waterford for more than 1000 minutes throughout the school year (n=1064) • Control: Used Waterford for less than 500 minutes throughout the school year (n=52)	Analysis of end of year scores revealed a significant difference between groups due to higher end of year scores made by treatment students than by control students in both kindergarten and first grade. Further analysis was conducted to examine the effects of gender and subsidized lunch on end of year scores, covarying for beginning of year scores, revealing that kindergarten and first grade treatment group students outperformed control group students across demographics. ANCOVAs showed that the interaction between lunch status and treatment was not significant across all strands, which indicates that WEL had a similar impact on end of year scores for students with free, reduced, and paid lunch status.

Waterford.org

This table provides a summary of research studies aligning to each ESSA evidence tier.

4—DEMONSTRATES A RATIONALE
Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Feehan, K., & Yoder, E. (2017d). Literacy improvement in early learners using technology. In Proceedings of the 11th International Technology, Education and Development Conference (pp. 7889-7896). Valencia, Spain: International Academy of Technology, Education and Development.	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	Volusia Literacy Test (VLT)	Kindergarten • Treatment: Waterford usage more than 1000 minutes (n=1287) • Control: Usage less than 500 minutes (n=43) First grade • Treatment: Waterford usage more than 1000 minutes (n=1892) • Control: Usage less than 300 minutes (n=34) Second grade • Treatment: Waterford usage more than 1200 minutes (n=2150) • Control: Usage less than 500 minutes (n=109)	Independent sample t-tests showed statistically significant positive effects for high use of WEL on the end of year VLT scores of Kindergarten students, $t(1, 1328) = -1.97$, $p < .05$, first grade, $t(1, 1924) = -3.14$, $p < .01$, and second grade, $t(1, 2257) = -2.57$, $p < .05$. The interaction between treatment and race/ethnicity, multi-lingual learners, and special education services was not significant, which shows that the Waterford reading program had a similar impact on VLT scores for African American/Black, Caucasian/White and Latino/a students, as well as MLL students, non-MLL students, special education services, and no special education services.
Shamir, H., Feehan, K., & Yoder, E. (2017e). Using computer-assisted instruction to improve early literacy skills for struggling students. In Proceedings of the 11th International Technology, Education and Development Conference (pp. 7897-7903). Valencia, Spain: International Academy of Technology, Education and Development.	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	Developmental Reading Assessment (DRA)	First Grade • Treatment: Used Waterford (n=103) • Control: Did not use Waterford (n=534) Second Grade • Treatment: Used Waterford (n=70) • Control: Did not use Waterford (n=407)	Analysis of percent gains between beginning of year and end of year scores revealed a significant difference between groups due to higher percent gains made by first and second grade students who used WEL than by control students. First and second grade treatment group students outperformed control group students in gender, free/reduced lunch status, and special education services.
Shamir, H., Feehan, K., Yoder, E., & Pocklington, D. (2018b). Can CAI improve reading achievement in young ELL students? In J. Vopava (Ed.), Proceedings of the Multidisciplinary Academic Conference on Education, Teaching, and E-Learning (MAC-ETEL 2018) (pp. 229-237). Prague, Czech Republic: Academic Conferences Association.	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	• The Texas Primary Reading Inventory (TPRI) • VLT	District 1 Kindergarten • Treatment: Students who used WEL (n=212) • Control: Students who did not use WEL (n=1484) Second Grade • Treatment: Students who used WEL (n=138) • Control: Students who did not use WEL, (n=1492) District 2 Kindergarten • Treatment: Students who used WEL for more than 1,000 minutes (n=1287) • Control: Students who used WEL for less than 500 minutes (n=43) Second grade • Treatment: Students who used WEL for more than 1,200 minutes (n=2150) • Control: Students who used WEL for less than 500 minutes (n=109)	Multi-lingual learners demonstrated dramatic improvements in learning skills following use of WEL enhanced curriculum: Use of WEL consistently increased performance on reading metrics.

This table provides a summary of research studies aligning to each ESSA evidence tier.

4—DEMONSTRATES A RATIONALE
Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Feehan, K., Yoder, E., & Pocklington, D. (2018a). Can CAI improve literacy and mathematics achievement in young low socioeconomic status students? In E. Langran & J. Borup (Eds.), Proceedings of Society for Information Technology & Teacher Education International Conference (pp. 993-998). Washington, D.C.: Association for the Advancement of Computing in Education (AACE).	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	• Developmental Reading Assessment (DRA • Mobile Classroom: The Dynamic Indicators of Basic Early Literacy Skills (mCLASS: DIBELS Next)	District 1 Kindergarten • Treatment: Students who used WEL for more than 1,000 minutes (n=1,004) • Control: Students who used WEL for less than 500 minutes (n=28) First grade • Treatment: Students who used WEL for more than 1,000 minutes (n=1,064) • Control: Students who used WEL for less than 500 minutes (n=52) District 2 Kindergarten • Treatment (reading): Students who used WEL for more than 1,000 minutes (n=108) • Control (reading): Students who used WEL for less than 400 minutes(n=30) • Treatment (math): Students who used WEL for more than 1,000 minutes (n=114) • Control (math): Students who used WEL for less than 400 minutes (n=58) First grade • Treatment (math): Students who used WEL for more than 1,000 minutes(n=255 • Control (math): Students who did not use WEL (n=68)	Students who used WEL throughout the school year in addition to traditional classroom learning scored consistently higher on reading and math assessments than their peers who did not use WEL. ANCOVAs showed that the interaction between lunch status and treatment was not significant across all strands, which indicates that WEL had a similar impact on end of year scores for students with free and paid lunch status, across both districts and grades.

This table provides a summary of research studies aligning to each ESSA evidence tier.

4—DEMONSTRATES A RATIONALE
Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Yoder, E., Pocklington, D., & Feehan, K. (2018b). Using CAI for improving academic skills of students with special needs. In M. Carmo (Ed.), Proceedings of the International Conference on Education and New Developments (END 2018) (pp. 557-561). Budapest, Hungary: World Institute for Advanced Research and Science.	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	• Mobile Classroom: The Dynamic Indicators of Basic Early Literacy Skills (mCLASS: DIBELS Next) • Mobile Classroom: Math (mCLASS: Math) • The Texas Primary Reading Inventory (TPRI)	District 1 • Treatment (reading): Students with special education services who used WEL for more than 1,000 minutes (n=28) • Control (reading): Students with special education services who used WEL for less than 400 minutes (n=8) • Control (reading): Students with no special education services who used WEL for less than 400 minutes (n=12) • Treatment (math): Students with special education services who used WEL for more than 1,000 minutes (n=33) • Control (math): Students with special education services who used WEL for less than 400 minutes (n=21) • Control (math): Students with no special education services who used WEL for less than 400 minutes (n=36) District 2 • Treatment (reading): Students with special education services who used WEL (n=7) • Control (reading): Students with special education services who did not use WEL (n=30) • Control (reading): Students with no special education services who did not use WEL (n=257)	Students with special education services who used WEL showed consistent improvement in reading and math. In all cases, students who used WEL not only outperformed the control group with students with special education services, but also the control group with students without special education services.
Shamir, H., Yoder, E., Pocklington, D., & Feehan, K. (2018a). Using adaptive CAI to supplement literacy development in early learners. Journal of Educational Multimedia and Hypermedia, 27(3), 367-389.	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	Texas Primary Reading Inventory (TPRI)	Kindergarten • Treatment: Used Waterford (n=212) • Control: Did not use Waterford (n=1484) First Grade • Treatment: Used Waterford (n=160) • Control: Did not use Waterford (n=1391)	Analysis of Kindergarten TPRI gains by sub-strand shows substantively important effect sizes (>0.25) for 6 of the 11 strands, including Letter Name Identification (0.74), Letter to Sound Linking (0.51), Inferring Word Meaning (0.34), Linking Details (0.49), Recalling Details (0.30), and Listening Comprehension Total Score (0.58). First Grade TPRI gains by sub-strand show substantively important effect sizes (>0.25) for eight of the nine substrands, including Blending Phonemes (0.85), Blending Word Parts (0.37), Blends in Final Position (0.71), Deleting Initial Sounds (0.68), Final Consonant Substitution (0.48), Initial Blending Substitution (0.51), Initial Consonant Substitution (0.39), and Middle Vowel Substitution (0.97).

This table provides a summary of research studies aligning to each ESSA evidence tier.

4—DEMONSTRATES A RATIONALE
Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Feehan, K., Pocklington, D., & Yoder, E. (2019a). Dosage effects of CAI on literacy skills. In Smart Education and e-Learning 2019 (pp. 149-158). Springer, Singapore.	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	• Standardized Test for the Assessment of Reading (STAR) • Idaho Reading Indicator (IRI)	STAR Kindergarten • Treatment: Randomly assigned to use WEL, who used WEL for at least 1,500 minutes (n=142) • Control: Randomly assigned to receive traditional literacy instruction (n=289) IRI Kindergarten • Treatment: Randomly assigned to use WEL, who used WEL for at least 1,500 minutes (n=145) • Control: Randomly assigned to receive traditional literacy instruction (n=314)	Students who used WEL for at least 1,500 minutes over the course of their Kindergarten school year had significantly higher gains from beginning of year to end of year across literacy strands on the STAR assessment. Additionally, students who met the recommended usage of WEL had higher gains than their control counterparts on the IRI Letter Naming Fluency strand and significantly higher gains on the IRI Letter Sound Fluency strand.
Shamir, H., Yoder, E., Pocklington, D., & Feehan, K. (2019b). Technology improving literacy skills for all students: Findings from three districts. International Journal of Information and Education Technology, 9(4), 280-285.	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	• The Texas Primary Reading Inventory (TPRI) • Developmental Reading Assessment (DRA) • VLT	District 1 Kindergarten • Treatment: K: Students who used WEL (n=212) • Control: K: Students who did not use WEL (n=1484) First grade • Treatment: Students who used WEL (n=160) • Control: Students who did not use WEL (n=1391) District 2 Second grade • Treatment: Students who used WEL(n=70) • Control: Students who did not use WEL (n=407) District 3 Kindergarten • Treatment: Students who used WEL for more than 1,000 minutes (n=1287) • Control: Students who used WEL for less than 500 minutes (n=43) First grade • Treatment: Students who used WEL for more than 1,000 minutes (n=1892) • Control: Students who used WEL for less than 300 minutes (n=34)	For all three districts, students who used WEL in addition to traditional classroom instruction had higher gains, percent gains, or end of year scores on reading assessments than students who did not use WEL. Results from District 3 demonstrated that students who had higher usage of WEL had higher literacy scores than students with less usage of WEL.

This table provides a summary of research studies aligning to each ESSA evidence tier.

4—DEMONSTRATES A RATIONALE
Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Yoder, E., Pocklington, D., & Feehan, K. (2019a). Computer-assisted instruction: Long-term effects on early literacy skills of low socioeconomic status students. International Journal for Information and Education Technology, 9(4), 263-267.	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	Kindergarten Readiness Test (KRT)	Kindergarten • Treatment : Used WEL in pre-kindergarten for more than 1,250 minutes (n=266) • Control: Did not use WEL (n=9,435)	In this longitudinal study, the students who used WEL outperformed the control group on the overall assessment score and on each of the substrands tested, including the Overall score, Letter ID and Sounds, Concepts of Print, and Phonemic Awareness. Students with MLL status and free lunch status, as well as Latino/a students in the experimental group, significantly outperformed their control group counterparts across all strands. African American/Black students in the experimental group also significantly outperformed African American/Black students in the control group on all strands except for Concepts of Print.
Shamir, H., Pocklington, D., Feehan, K., & Yoder, E. (2019b). Evidence for dosage and long-term effects of computer-assisted instruction. International Journal of Learning and Teaching, 5(3), 220-226.	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	Developmental Reading Assessment (DRA)	Cohort 1 First grade • Kindergarten and first grade usage (n=1,416) • Kindergarten only usage (n=26) • No usage (n=31) 2nd grade • First grade only usage (n=1,529) • No usage (n=323 Cohort 2 First grade • Kindergarten and first grade usage (n=1,381) • Kindergarten only usage (n=39) • No usage (n=32) 2nd grade • Kindergarten and first grade usage (n=1,235) • Kindergarten only usage (n=49) • First grade only usage (n=237) • No usage (n=281)	This longitudinal study found that students who used WEL only in kindergarten outperformed students in the control group on first and second grade literacy scores. The findings also offer support for a dosage effect of WEL: While analysis showed all experimental students outperformed their control counterparts, the largest effect sizes were found in the analysis including students with high usage of WEL.
Shamir, H., Feehan, K., Pocklington, D., & Yoder, E. (2019b). Effects of long-term and early use of CAI on students' literacy skills. In EdMedia+ Innovate Learning (pp. 1406-1411). Association for the Advancement of Computing in Education (AACE).	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP)	Treatment • Students who used WEL for three years for more than 100 minutes (n=2,067) • Students who used WEL for two years for more than 100 minutes, in K and first grade (n=136) Control • No usage (n=75) • Students who used WEL for two years for more than 100 minutes, in first and second grade (n=677)	Students who used WEL had higher literacy test scores than their control counterparts, demonstrating an overall effect for the use of WEL. Students who used, and then stopped using, WEL still outperformed their control counterparts one year after they stopped using WEL, demonstrating a long-term effect for the use of WEL. Additionally, students who started using WEL early in their academic careers outperformed their counterparts who started later, demonstrating an effect for the early use of WEL.

This table provides a summary of research studies aligning to each ESSA evidence tier.

4—DEMONSTRATES A RATIONALE
Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Pocklington, D., Yoder, E., & Feehan, K. (2020). Long-Term Effects of CAI in Early Education: Evidence from Three Districts. In D. Schmidt-Crawford (Ed.), Proceedings of Society for Information Technology & Teacher Education International Conference (pp. 826-831). Online: Association for the Advancement of Computing in Education (AACE).	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	• District-Administered Literacy Assessment • Developmental Reading Assessment (DRA)	District 1 • 3 Years of Usage Experimental (n=2,108) • K & 1st Grade Usage Experimental (n=288) • No Usage Control (n=593) • No K Usage Control (n=729) District 2 • 3 Years of Usage Experimental (n=585) • K & 1st Grade Usage Experimental (n=478) • No Usage Control (n=142) • No K Usage Control (n=137) District 3 • K & 1st Grade Usage Experimental (n=1,288) • No Usage Control (n=287) • No K Usage Control (n=278)	All three hypotheses were supported across the different districts, demonstrating overall, long-term, and early effects of using WEL. For the analysis of overall effects, in the two districts where overall effects were assessed, students who used WEL for all three years significantly outperformed control students, who did not use WEL, on all strands assessed. For the analysis of long-term effects, students who used WEL prior to second grade outperformed control students at the end of second grade on all strands assessed and in all three districts. For the analysis of early effects, across all districts, students who started using WEL during kindergarten, and consistently used for two or three years, significantly outperformed students who started using the program one year later on all strands assessed.
Evaluation and Training Institute (2020c). Mississippi UPSTART and Head Start Evaluation Kindergarten Impact Study.	Waterford Early Learning (UPSTART)	Provides a well-specified logical model informed by research or evaluation	Kaufman Test of Educational Achievement (KTEA-3)	Kindergarten • Treatment: Assigned to use Reading during the previous year through UPSTART (n=50) • Control: Assigned to use Math during the previous year through UPSTART (n=62)	Students who were assigned to use the reading software during the previous year significantly outperformed students assigned to use math software when assessed on literacy skills at the end of kindergarten, including letter-word recognition (ES = 0.41), reading comprehension (ES = 0.39), and phonological processing (ES = 0.38). Additionally, students who used the math software during the previous year significantly outperformed students who had reading usage when assessed on the math concepts and applications subskill (ES = 0.50). The majority of students (96%) in this longitudinal study were African American/Black.
Evaluation and Training Institute (2011). Kindergarten Outcomes: Program Impacts on Reading Proficiency. First Year Results: Utah UPSTART Education Program Evaluation Technical Report.	Waterford Early Learning (UPSTART)	Provides a well-specified logical model informed by research or evaluation	Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Next	Students served: 1,248 Pre-K Students • Treatment: Students enrolled in the UPSTART program (n=137) • Control: Students not enrolled in the UPSTART program (n=247)	Findings revealed that children who had participated in UPSTART during preschool scored almost 18 points higher in reading proficiency on the DIBELS Next (DN) Composite compared to beginning kindergarten children who did not participate in UPSTART prior to enrolling in public school. Additionally, middle kindergarten children who had participated in the UPSTART preschool program scored approximately 19 points higher in reading proficiency on the DN Composite compared to middle kindergarten children who did not participate in UPSTART prior to enrolling in public school. When assessed at the beginning of kindergarten by the DIBELS Next Beginning Kindergarten Composite, children participating in UPSTART demonstrated moderately strong improvements in reading proficiency compared to children who did not participate in UPSTART and demonstrated higher gains when assessed on the DIBELS Next Middle Kindergarten Composite.

This table provides a summary of research studies aligning to each ESSA evidence tier.

4—DEMONSTRATES A RATIONALE

Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Miner, C., Izzo, A., Pocklington, D., Feehan, K., & Yoder, E. (2018). Preparing students for kindergarten using UPSTART at home. Journal of Educational Multimedia and Hypermedia, 27(2), 209-230.	Waterford Early Learning (UPSTART)	Provides a well-specified logical model informed by research or evaluation	Waterford Assessments of Core Skills (WACS)	• Year 1 participants consisted of 384 children, including 137 in the treatment group and 247 in the control group. • Year 2 participants consisted of 190 children, including 95 in the treatment group and 95 in the control group. • Year 3 participants consisted of 305 children, including 151 in the treatment group and 154 in the control group. • Year 4 participants consisted of 220 children, including 117 in the treatment group and 103 in the control group. • Year 5 participants consisted of 271 children, including 109 in the treatment group and 162 in the control group. • Year 6 participants consisted of 529 children, including 200 in the treatment group and 329 in the control group.	Using a pre- and posttest study design, UPSTART participants completed the program at the WACS Kindergarten Advanced level on average. Students made WACS reading score gains across demographics including race/ethnicity, other preschool attendance, and experiencing poverty.
Suddreth, D., Throndsen, J., & Wiebke, S. (2016). UPSTART Program: Report of FY 2016. Utah State Office of Education.	Waterford Early Learning (UPSTART)	Provides a well-specified logical model informed by research or evaluation	• Dynamic Indicators of Basic Early Literacy Skills (DIBELS) • Student Assessment of Growth and Excellence (SAGE)	Students served: 5,091 PreK Students Analytic Matched Sample: • Treatment (students enrolled in UPSTART), n=138 • Control (students who did not enroll in UPSTART), n=138	This study shows that students who participated in the UPSTART program as preschoolers maintained their gains longitudinally on state testing, outscoring non-UPSTART students on state testing in Grades 1-4.
Shamir, H., Pocklington, D., Yoder, E., Feehan, K., & Ortiz-Wood, C. (2022). Long-Term and Early Effects of Computer-Assisted Instruction in Low Socioeconomic Status Students. In Smart Education and e-Learning-Smart Pedagogy (pp. 77-86). Singapore: Springer Nature Singapore	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	District-Administered Literacy Assessment	1,796 2nd grade students. The three years of usage experimental group included 585 students who used WEL for more than 100 minutes during kindergarten (2016-2017), first grade (2017-2018), and second grade (2018-2019). The usage only in kindergarten and first grade experimental group included 478 students who used WEL for more than 100 minutes during kindergarten (2016-2017) and first grade (2017-2018) but used WEL for less than 100 minutes while in second grade (2018-2019). The no usage control group included 142 students who used WEL for less than 100 minutes during all three school years. The usage only in first grade and second grade control group included 137 students who used WEL for more than 100 minutes during first grade (2017-2018) and second grade (2018-2019) but used WEL for less than 100 minutes while in kindergarten (2016-2017).	For three years of usage compared to no usage: Across most demographics, students in the experimental group significantly outperformed students in the control group at the end of second grade (d = 0.30). For kindergarten and first grade usage compared to no usage: Across all demographics, students in the experimental group outperformed students in the control group (d = 0.48). For three years of usage compared to first grade and second grade usage: Across all demographics, students in the experimental group outperformed students in the control group (d = 0.26).

This table provides a summary of research studies aligning to each ESSA evidence tier.

4—DEMONSTRATES A RATIONALE

Does Not Meet WWC Evidence Standards

Evidence	Program	Study Design	Instruments Used	# of Children	Outcomes
Shamir, H., Yoder, E., Pocklington, D., Feehan, K., & Ortiz-Wood, C. (2022). Long-Term and Early Effects of Computer-Assisted Instruction in Early Childhood. In Smart Education and e-Learning-Smart Pedagogy (pp. 109-117). Singapore: Springer Nature Singapore	Waterford Early Learning (WEL)	Provides a well-specified logical model informed by research or evaluation	District-Administered Literacy Assessment	5,002 second grade students. The three years of usage experimental group included 2,108 students who used WEL for more than 100 minutes during kindergarten (2015-2016), first grade (2016-2017), and second grade (2017-2018). The kindergarten and first grade usage experimental group included 288 students who used WEL for more than 100 minutes during kindergarten (2015-2016) and first grade (2016-2017) but used the program for less than 100 minutes during second grade (2017-2018). The no usage control group included 593 students who used WEL for less than 100 minutes during all three school years. The first grade and second grade usage control group included 729 students who used WEL for more than 100 minutes during first grade (2016-2017) and second grade (2017-2018) but used the program for less than 100 minutes during kindergarten (2015-2016).	For three years of usage compared to no usage: End of second grade scores revealed a significant difference between groups due to higher end of second grade scores made by experimental students for both subskills, Dictation Sounds ($d = 0.19$) and Dictation Words ($d = 0.15$). For kindergarten and first grade usage compared to no usage: Analysis of end of second grade scores revealed a significant difference between groups due to higher end of second grade scores made by experimental students than by control students for Dictation Sounds ($d = 0.17$). Diction Word scores made by experimental students were higher than scores made by the control students, but the difference was not significant. For three years of usage compared to first grade and second grade usage: Analysis of the end of second grade scores revealed a significant difference between groups due to higher end of second grade scores made by experimental students than by control students for both subskills, Dictation Sounds ($d = 0.15$) and Dictation Words ($d = 0.09$).
Evaluation and Training Institute (2022b). UPSTART Program Evaluation: Year 13 Program Results	Waterford Early Learning (UPSTART)	Provides a well-specified logical model informed by research or evaluation	Utah's Pre-Kindergarten Entry and Exit Profile (PEEP) Entry Utah's Kindergarten Entry and Exit Profile (KEEP) Entry	Students served: 13,404 • Tier 1: n=454 given UPSTART software only • Tier 2: n=411 received text message support • Tier 3: n=12,535 full support	Average usage of UPSTART was around 36 hours with program graduates having an average usage of 45 hours. Students in Tier 1 spent less time with the UPSTART program compared to students in Tier 3; $t(12987) = 3.5$, $p = 0.000$. Similarly, students in Tier 3 spent more time with the program than students in Tier 2; $t(12944) = 3.0$, $p = 0.001$. Students in Tiers 2 and 3 had higher percentages of students reaching Level 3 on the PEEP entry in literacy and numeracy (87% and 81%, respectively). Of the students who classified as in need of support for literacy at the start of the program, 75% of them achieved Level 3 at the end of the program and of those in need of support for numeracy, 65% of them achieved Level 3 at the end of the program.
Shamir, H., Yoder, E., & Pocklington, D. (2023b). Usage matters: Longitudinal benefits of computer-assisted instruction for early learners. In EDULEARN23 Proceedings (pp. 7775-7782). IATED.	Waterford Reading Academy (WRA)	Provides a well-specified logical model informed by research or evaluation	Texas Primary Reading Inventory (TPRI) exam	Experimental group: students with high usage (at least 1,000 minutes) during kindergarten in 2019-2020 (n=534) and 2021-2022 (n=592). Control group: Students with low usage (less than 500 minutes) during kindergarten in 2019-2020 (n=115) and 2021-2022 (n=127).	Examination of end of first grade scores between experimental and control groups revealed that the experimental group significantly outperformed the control group on TPRI Word Reading for the 2020-2021 cohort ($d = 0.59$), the 2021-2022 cohort ($d = 0.65$), and the combined sample ($d = 0.62$). Results indicated that all students across race/ethnicity, special education services, experiencing poverty, and multi-lingual learner status, benefited from the use of WRA at the end of first grade, one year after they had stopped using the software.

Waterford.org