Carnegie Learning: An ESSA Evidence-Based Approach

Many companies create products and then look to research to validate those products. Carnegie Learning started with the research and then built and improved our products from that research.

Founded by cognitive and computer scientists from Carnegie Mellon University, we have been deeply immersed in research from the start. Our work is guided by more than 20 years of scientific research into how students learn math. Everything we do is driven by our research background and commitment to using learning science to make math learning better for students and teachers.

ESSA LEVELS OF EVIDENCE

The Every Student Succeeds Act (ESSA), the primary federal law governing K–12 education, requires many funding recipients to demonstrate that their programs are "evidence-based." Some state funding programs also follow these federal requirements. Different funding sources may require different levels of evidence, but generally a program will count as being evidence-based if it either:

1. Shows evidence of effectiveness, as demonstrated by research findings categorized below:
   - **Tier 1: Strong** – there is at least one strong experimental study
   - **Tier 2: Moderate** – there is at least one quasi-experimental study
   - **Tier 3: Promising** – there is correlational evidence that the program has positive effects
2. Demonstrates a rationale based on high-quality research findings that it is likely to improve student outcomes and includes ongoing efforts to examine its effects (Tier 4).

Here is the evidence to support the use of Carnegie Learning's blended math solution, artificial intelligence-powered software and professional learning, which can be referenced in ESSA funding applications.

EVIDENCE FOR CARNEGIE LEARNING'S BLENDED SOLUTIONS

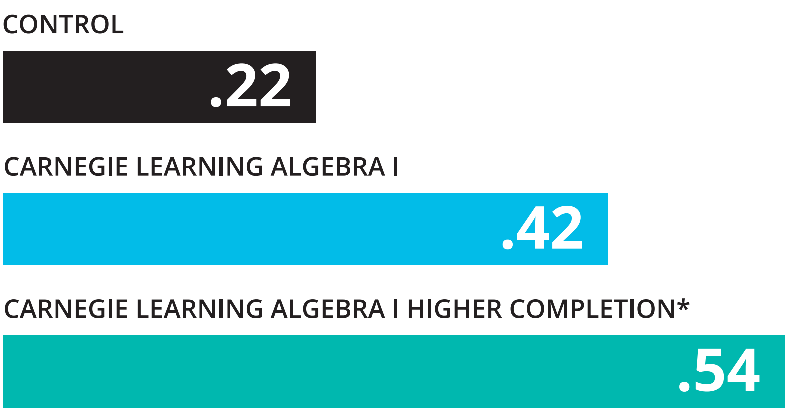
Carnegie Learning's blended solutions are among the most carefully studied mathematics curricula and meet Tier 1 "Strong" evidence standards. The Carnegie Learning Middle School and High School Math Solutions combine consumable textbooks with artificial intelligence-powered software.

On the basis of prior results, the RAND Corporation was awarded a grant from the U.S. Department of Education to evaluate the effectiveness of Carnegie Learning's blended Algebra I course in one of the largest randomized control trials of its kind. The trial spanned seven states, 147 middle and high schools and over 19,000 students, representing a wide range of demographic characteristics (Pane, Griffin, McCaffrey & Karam, 2014). This study is especially noteworthy not only for its large-scale and gold-standard design but also for its two-year approach, which allowed it to determine the curriculum's effectiveness after the school had a year to adjust to the new educational approach.

The study compared schools that were given Carnegie Learning's blended solution for Algebra I to schools that continued to use the curriculum they had in place (in almost all cases, one of the top 4 Algebra textbooks). Results for high schools and middle schools were considered separately since the middle school students were typically more advanced (taking Algebra I in 8th grade). Each school in the study participated in the experiment for two years, leading to two cohorts of Algebra I students. The McGraw-Hill Acuity® series was used for pre- and post-tests, and results are based on scores on that exam.

In the first (transitional) year, there was no significant difference found. **In the second year of the study, both groups using the Carnegie Learning Middle School and High School Math Solutions out-performed the control groups.** The effect was statistically significant for the high school cohort (p<.05; d=0.20) but not for the middle school cohort (p=0.16; d=0.19). Although the effect sizes were similar for both cohorts, the middle schools were smaller, resulting in less power to show a statistically significant result.

The effect is large enough to be educationally meaningful. Lipsey et al. (2012) found that 8th or 9th grade students typically improve about 0.22 standard deviations (SDs) on standardized math tests in one year. This is the improvement that was experienced by the control group. In the study, the Carnegie Learning Blended Solution group outscored the control group by 0.20 SDs, for a total improvement of 0.42 SDs, as compared to the 0.22 SD growth in a typical year. **In other words, growth in knowledge by students using the Carnegie Learning Blended Solution nearly doubled that of the control group.** The effect is also equivalent to moving a student from the 50th percentile for the Algebra I standardized test outcome to the 58th percentile.



STANDARD DEVIATIONS
*Students completing more than 27 sections in Carnegie Learning Software

Sales and Pane (2015) used principal stratification to explore the effect of the curriculum on students who used the software to a reasonable extent. They found that students who used the software to complete more than the median number of topics scored 0.32 standard deviations higher than matched students in the control group, equivalent to moving this group to the 62nd percentile.

A second randomized field trial demonstrated significant improvement in course grades. The Moore Oklahoma Independent School District conducted a randomized field trial comparing instruction based on the Carnegie Learning's blended Algebra I curriculum to control classes using the McDougal-Littell Heath Algebra I textbook (Ritter, Kulikowich, Lei, McGuire and Morgan, 2007). In three schools, teachers taught some of their classes using blended curriculum and some using the textbook. Results favored the classes using Carnegie Learning's curriculum as measured by first semester grades (p=.002, d=.42), final grades (p=.007, d=.36) and scores on the ETS Algebra I end-of-course exam (p=.091, d=.38). The U.S. Department of Education's Institute of Education Sciences' (IES) What Works Clearinghouse recognizes this study as meeting their evidence standards without reservations.

References
Lipsey, M.W., Puzio, K., Yun, C., Hebert, M.A., Steinka-Fry, K., Cole, M.W., Roberts, M., Anthony, K.S., Busick, M.D. (2012). Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms. (NCSER 2013-3000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.

Pane, J.F., Griffin, B.A., McCaffrey, D.F., Karam, R. (2014). Effectiveness of Cognitive Tutor Algebra I at Scale. *Educational Evaluation and Policy Analysis*, 36(2), 127-144.

Ritter, S., Kulikowich, J., Lei, P., McGuire, C.L. & Morgan, P. (2007). What evidence matters? A randomized field trial of Cognitive Tutor Algebra I. In T. Hirashima, U. Hoppe & S. S. Young (Eds.), *Supporting Learning Flow through Integrative Technologies* (Vol. 162, pp. 13-20). Amsterdam: IOS Press.

Sales, A. C. and Pane, J. F. (2015). Exploring causal mechanisms in a randomized effectiveness trial of the cognitive tutor. In *Proceedings of the 8th International Conference on Educational Data Mining. International Educational Data Mining Society*.

EVIDENCE FOR MATHIA® SOFTWARE

The Theory Behind MATHia

MATHia has its basis in the ACT-R theory of knowledge and performance (Anderson et. al., 2004; Anderson, 2007). Following ACT-R, MATHia treats complex problem solving as the coordination and strengthening of a large number of relatively simple "knowledge components," which represent strategies and concepts required to master a domain. These knowledge components are collected into a "cognitive model" which allows the software to follow individual students' solution strategies and track the growth of knowledge for each student over time.

MATHia employs mastery learning. Students need to demonstrate mastery on each knowledge component underlying a particular topic before they can proceed to the next topic. In this way, students set their own pace through the curriculum. Mastery learning has a long history of support. Kulik et al. (1990) concluded, after reviewing 108 studies, that mastery learning programs had strong impact (average effect size of 0.5).

MATHia selects problems for each student in order to maximize the amount of time that students spend on knowledge components that they have not yet mastered and minimize the amount of time that they spend on components that they have already mastered. The tutor is able to follow individual solution strategies and provide students with hints that are relevant to their individual approach. In addition to correct solution strategies, the cognitive model also includes information about common misconceptions and incorrect strategies and presents students with immediate feedback if they make common errors.

MATHia Evidence

Evidence for MATHia software used outside of the blended curriculum meets ESSA standards for Tier 2 "Moderate" evidence.

Researchers have conducted dozens of small-scale evaluations of individual units of instruction within the MATHia software. For example, Aleven and Koedinger (2002) compared two versions of instruction focused on reasoning about angle measures in a diagram. They found that students who were asked to articulate the reasons for geometric theorems relating angles in a diagram outscored those who were asked simply to provide the angle measures. Subsequent to these experiments, the MATHia software was revised to include student use of geometric reasons as a fundamental part of the task. Butcher and Aleven (2008) conducted several experiments that showed that closely integrating visual reasoning about geometric diagrams with the numeric and symbolic reasoning required to determine angle measures leads to more robust learning. The most recent version of the MATHia software incorporates these changes as well. Cen, Koedinger and Junker (2007), in a quasi-experimental design, used data from prior-year implementations to fit various parameters controlling the learning rate within the the MATHia software and found that, with properly fit parameters, students could reach the same level of performance in 12% less time.

Fancsali et al. (2018) provide correlational evidence that use of MATHia software is associated with test score outcomes. The study used data from over 23,000 students in Miami-Dade County Public Schools. Results showed that a statistical model based on process variables including the rate at which content was mastered, the number of errors and hints made in the software, the number of problems, skills and topics mastered and the amount of time taken was able to predict outcomes on the Florida Standards Assessment (FSA) and Florida Comprehensive Assessment Test (FCAT). The model controlled for prior ability (as measured by previous year test score) and student demographics and was able to predict data from three grade levels (6, 7 and 8) and three school years. Similar predictive models have shown to predict results in Virginia (Ritter et al., 2013) and West Virginia (Joshi et al., 2014).

References

Aleven, V., & Koedinger, K. R. (2002). An effective meta-cognitive strategy: learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science*, 26(2), 147-179.

Anderson, J. R. (2007) How Can the Human Mind Occur in the Physical Universe? New York: Oxford University Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y . (2004). An integrated theory of the mind. *Psychological Review* 111, (4). 1036-1060.

Cen, H., Koedinger, K. R., & Junker, B. (2006). Learning Factors Analysis - A general method for cognitive model evaluation and improvement. Proceedings of the 8th International Conference on Intelligent Tutoring Systems (pp. 164-175): Springer Berlin / Heidelberg.

Joshi, A., Fancsali, S.E., Ritter, S., Nixon T. (2014). Generalizing and Extending a Predictive Model for Standardized Test Scores Based On Cognitive Tutor Interactions. In J. Stamper, Z. Pardos, M. Mavrikis, B.M. McLaren (Eds.). Proceedings of the 7th International Conference on Educational Data Mining (pp. 369-370).

Kulik, C.-L., Kulik, J.A., & Bangert-Drowns, R.L. (1990). Effectiveness of mastery learning programs: A meta-analysis. Review of Educational Research, 60, 265-299.

Fancsali, S.E., Zheng, G., Tan, Y., Ritter, S., Berman, S.R., Galyardt, A. (2018). Using Embedded Formative Assessment to Predict State Summative Test Scores. In Proceedings of the International Conference on Learning Analytics and Knowledge (pp. 161-170): ACM, New York. DOI: 10.1145/3170358.3170392

Ritter, S., Joshi, A., Fancsali, S.E., Nixon, T. (2013). “Predicting Standardized Test Scores from Cognitive Tutor Interactions.” In Proceedings of the 6th International Conference on Educational Data Mining (pp. 169−176).

EVIDENCE FOR CARNEGIE LEARNING’S PROFESSIONAL LEARNING

Carnegie Learning’s Professional Learning services satisfy ESSA’s evidence standards in demonstrating a rationale that they are based on practices likely to be effective in increasing teachers’ mathematics content and pedagogical content knowledge.

These professional learning programs are focused on increasing teachers’ depth of understanding of mathematics and enabling teachers to leverage this content knowledge to support their students’ learning. The effect of such pedagogical content knowledge on student learning has been established in many studies (c.f. Hill, Rowan and Ball, 2005), demonstrating that interventions focused on pedagogical content knowledge are likely to be effective.

Carnegie Learning has evaluated the effectiveness of our professional learning using the University of Michigan’s Learning Mathematics for Teaching (LMT) assessment (2011). In one study (Fancsali, 2017), teachers were administered the LMT assessment prior and subsequent to a four-day summer Mathematics Academy. Analysis of LMT scores showed significant and substantial improvement over this period, $t(17)=3.93$, $p=.001$. The effect size was large, $d=0.93$.

A study by the University of Louisville (Jones and Bush, 2009) found substantial increases in teachers' content knowledge (as measured by the Algebra Assessment of the Diagnostic Teacher Assessments in Mathematics and Science) following three years of summer mathematics academies for middle school teachers in central Kentucky. Results showed significant and substantial increases in Algebra performance, $t(71)=13.13$, $p<.001$. The effect size was again large, $d=1.03$.

A study by the state of West Virginia (Stohr, 2013) looked at the impact of a 5-day Carnegie Learning Summer Academy. Participants were special education teachers in grades 5–12. This study found a significant increase in these teachers' knowledge of proportional reasoning as assessed by the LMT, $t(41)=2.05$, $p<.05$; $d=0.4$.

Carnegie Learning continues to study the effects of its professional learning programs.

References

Fancsali, S. (2017). *Preliminary Research Analysis: Carnegie Learning Professional Learning*. Internal Carnegie Learning report: unpublished.

Hill, H.C., Rowan, B., & Ball, D.L. (2005) Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42, 371-406.

Jones, J. H. and Bush, W. S. (2009). *K8 Mathematics Alliance Teachers' Beliefs and Attitudes toward Mathematics Algebra Content Knowledge Summary 2006-2008*. University of Louisville report: unpublished.

Learning Mathematics for Teaching. (2011). Measuring the Mathematical Quality of Mathematics teaching. *Journal for Mathematics Teacher Education 14(1)*, 25-47.

Stohr A. D. (2013). *Mathematics academies: Cohort 1 evaluation study*. Charleston, WV: West Virginia Department of Education, Division of Teaching and Learning, Office of Research.