Science Achievement Gains From A Computer-Based Learning Platform: A Multi-Site Quasi-Experimental Study

Steven L Miller*
Elite Performance Solutions
Pacifica, CA, USA
steve@eliteperform.org

***Abstract*—This quasi-experimental, multi-site study examined the impact of *Penda Learning*, a computer-based science instructional and intervention platform, on student science achievement across 46 schools, involving 19,500 students in Grades 3 through 12. The intervention group—classrooms implementing *Penda Learning*—was compared to a control group receiving business-as-usual science instruction in Grades 3 through 8 (n = 18,738). The platform was aligned with state science standards, curriculum, and instruction. Student performance was assessed using the NWEA Measures of Academic Progress (MAP) Science Assessment at both pretest (Fall) and posttest (Spring). A repeated-measures MANCOVA was conducted using pretest scores as a covariate. Results revealed statistically significant improvements in science achievement for students in the intervention group across all grade levels and demographic subgroups, with a dose-response relationship observed—greater usage was associated with higher learning gains. The average effect size, calculated using Hedges' *g* under a fixed-effects model, was 0.56 (95% CI [0.53, 0.59]), reflecting an estimated 21% gain. Students in the intervention group showed 20% higher increases in estimated science proficiency rates compared to their peers in the control group. These findings suggest that computer-based instruction using *Penda Learning* produced meaningful gains in both psychometric growth and science proficiency, with implications for future educational equity, STEM workforce development, and economic mobility. This study meets Tier II ("moderate evidence") standards of the Every Student Succeeds Act (ESSA) based on the U.S. Department of Education guidelines.**

Keywords—Science Intervention, K12 Science, State Science Proficiency, NWEA MAP Science, Closing Science Gaps

I. Introduction

Education is widely recognized as a foundational driver of socio-economic development. As the educational attainment of a population increases, economic sectors tend to expand, resulting in greater employment opportunities and contributing to national growth in gross domestic product (GDP) [1]. Within this framework, science and mathematics proficiency—routinely assessed in large-scale international evaluations such as the Trends in International Mathematics and Science Study (TIMSS) and the National Assessment of Educational Progress (NAEP)—has emerged as a critical predictor of both academic achievement and long-term economic outcomes [2].

Enhancing scientific knowledge plays a central role in improving human capital, boosting labor productivity, and fostering innovation, all of which are deeply interconnected mechanisms for advancing economic prosperity [3]. The development of scientific competencies, particularly at the primary and secondary levels, equips individuals with the cognitive tools necessary to participate in increasingly knowledge-based economies.

A. The Economic Impact of STEM Education

Even modest increases in STEM proficiency at the local level can yield significant economic returns. For instance, improving science outcomes for just 100 students in a single community has the potential to raise local economic productivity, elevate future earnings, and reduce public expenditures [3]. These improvements not only benefit individuals through increased wages but also generate broader fiscal advantages, including higher tax revenues and decreased reliance on public assistance programs.

Hanushek and Woessmann [3] found that improving workforce cognitive skills by just 0.25 standard deviations could substantially enhance national economic well-being. Their longitudinal economic modeling demonstrates that it is the *quality* of learning—measured by outcomes in math and science—not the *quantity* of schooling, that drives growth. According to their projections, raising average PISA scores by 25 points over two decades could increase the cumulative GDP of OECD countries by approximately $115 trillion over the lifetime of the cohort born in 2010. More ambitious reforms could yield economic gains exceeding $260 trillion, a topic to be addressed in future analyses.

B. Research Objectives and Methodology

The primary objective of this study is to evaluate the effectiveness of a computer-based science instructional platform (*Penda Learning*) in improving science knowledge and proficiency among students within a single academic year. To examine this effect with both breadth and rigor, the study employs a quasi-experimental design across multiple school sites, incorporating an active control group receiving standard science instruction. This design enables the assessment of program efficacy under real-world educational conditions, increasing the external validity of the findings.
Student performance in science was measured using the NWEA Measures of Academic Progress (MAP) Growth Science Assessment, a nationally normed and

Funding provided by Learning 2020 Inc.

*Corresponding Author

psychometrically validated instrument. The MAP Science assessment is recognized for its high reliability, strong construct validity, and capacity to predict science proficiency across grade levels [4]. These psychometric properties support its use for both student-level diagnostic feedback and school- or district-level evaluation [2].

The primary hypothesis tested is that students receiving instruction through *Penda Learning* will exhibit significantly greater gains in science knowledge, as measured by changes in MAP Growth Science scores, than students in the business-as-usual control condition. Secondary analyses investigate whether the program's impact varies by grade level, demographic characteristics, and intensity of software engagement, including the presence of a dose-response relationship. The results will report key indicators of success, such as effect sizes and percentile gains, to anticipate potential proficiency improvements based on NWEA MAP scores..

II. Data Collection and Processing

A. Research Design

This study employed a two-group, quasi-experimental pretest–posttest design to assess the impact of computer-based science instruction using the *Penda Learning* platform. Baseline science achievement data were collected in the Fall using the NWEA MAP Growth Science Assessment, and post-intervention data were collected in the Spring of the same academic year. The comparison group received standard science instruction, while the intervention group received instruction supplemented by *Penda Learning*.

Demographic and grade-level data were obtained from participating schools to support subgroup analyses. To evaluate the intervention's impact, Multivariate Analysis of Covariance (MANCOVA) procedures were used, with pretest scores entered as a covariate to adjust for initial differences in performance and reduce potential selection bias between groups.

B. Instruction and Assessments

The study was conducted across 46 Charter Schools USA (CSUSA) sites in Florida, encompassing Grades 3 through 12. All students in both groups received science instruction as part of their core curriculum. CSUSA implements an educational model informed by Robert J. Marzano's "What Works in Schools" framework and guided by a Guaranteed and Viable Curriculum (GVC), which emphasizes evidence-based instructional strategies and alignment with state standards.

During the 2021–2022 academic year, a total of 6,097 CSUSA students in Grades 3–12 used *Penda Learning* as a supplementary science instruction and intervention tool. Student engagement varied from minimal use to completion of over 300 activities. A "mastered activity" was defined as one in which a student achieved a score of 80% or higher on embedded formative assessments.

The NWEA MAP Growth Science Assessment served as the primary outcome measure. This assessment is nationally normed, computer-adaptive, and designed to track student science achievement and growth over time with strong psychometric properties [4].

C. Subjects

The study included 19,591 students in Grades 3 through 12 across 46 schools. Because over 95% of participants were enrolled in Grades 4 through 8, all grade-level analyses were restricted to that subgroup.

Table 1 presents the ethnic distribution of the sample. The majority of students identified as White (54%), Black or African American (29%), or Hispanic or Latino (8%). The gender distribution was approximately balanced, with a female-to-male ratio of 1.02:1.

Table 1. The study's distribution by ethnicity.

Ethnicity	Sample %
American Indian or Alaska Native	<1%
Asian	2%
Black or African American	29%
Hispanic or Latino	8%
Multi-ethnic	4%
Non-specified or Other	3%
White	54%

D. Instructions for the Intervention and Control Conditions

Students in the experimental group (n = 6097) participated in *Penda Learning* as a supplemental instructional and intervention tool. After completing the Fall pre-assessment, teachers were granted access to the platform and instructed to integrate Penda activities into science instruction at least three times per week for 20–30 minutes per session.

Teachers in the intervention group received structured professional development (PD) before and during the intervention period. PD consisted of a five-day Summer Institute, which provided training in curriculum integration, platform navigation, and instructional alignment. Additionally, Penda Learning content specialists conducted ongoing classroom visits and facilitated professional learning communities (PLCs) to promote the implementation of fidelity and shared instructional practices.

The comparison group consisted of 12,597 students who received traditional, business-as-usual science instruction designed to meet state science standards. These students were taught by educators who did not participate in *Penda Learning* professional development. This condition is best characterized as an active control, in which students received conventional science instruction without exposure to the intervention platform.

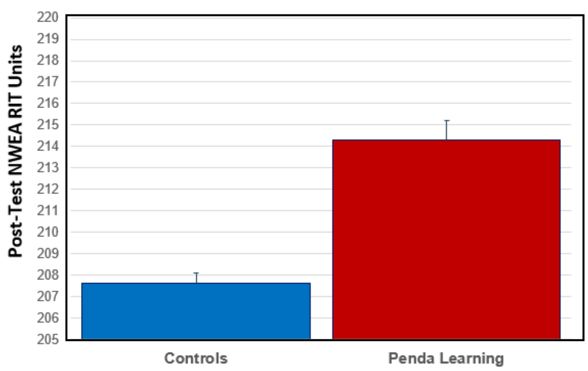
III. RESULTS

A. *Science Knowledge Results*

The data were subjected to a repeated measures multivariate analysis of covariance using SPSS. The factors in the model are Assessment Time (2), Group (2), Gender (2), and Ethnicity (7). The MANCOVA used the Pre-Test score (Fall NWEA MAP) as the covariate to control for any correlated differences in pre-training Science performance.

The results from the analysis show a main effect for the Group variable. Variable Spring NWEA, $F(1, 28) = 41.2$, $p<.001$. Figure 1 shows that Science knowledge gains are significantly larger for students receiving Penda Learning compared to standard science curriculum instruction ($p<.001$). Note: The pre-test Science assessment performance was used as a covariate for the analyses.

Figure 1. Main Effect of Group for Penda Users and Active Controls on the NWEA MAP Growth Science assessment.



B. *Science Knowledge by Demographic Groups*

Additionally, the interactions between Group by Gender and Group by Ethnicity were statistically significant. In both cases, all groups made substantial gains in Science performance from the Fall to Spring testing. However, gains on the Penda learning platform were larger than those for standard classroom science instruction, with the effect being larger for girls than for boys, as shown in Figure 2. Science knowledge improved more for the Penda Learning group versus the Control group, and these differences interacted across the Penda Learning and Control groups. Ethnicity $F(6,28)= 45.12$, $p<05$. and significant interactions for Group by Gender $F(1,28) = 3.98$, $p<.05$, and Group by Ethnicity $F(6,28) = 2.27$, $p<.05$. At pre-test (baseline), group differences were present for both control and penda learning groups. At the post-test, the Penda learning group outperformed the controls for each reported ethnic group; however, the size of the differences also varied by group. See Figure 4 for an illustration of these group differences.

Research consistently shows that interest in science, technology, engineering, and mathematics (STEM) among girls and boys is highest in early education but declines further as students' progress from elementary school through high school. By fourth grade, national data suggest that while achievement gaps in math and science between boys and girls are relatively small, disparities in interest and self-efficacy are already evident [5].

Figure 2. Science knowledge gains for males and females. Penda Learning subjects and females benefited significantly more than comparison participants (p<.05).

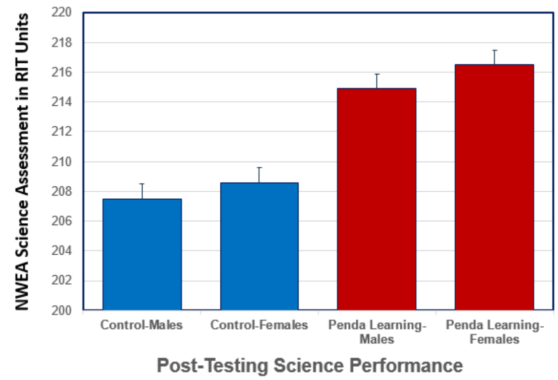
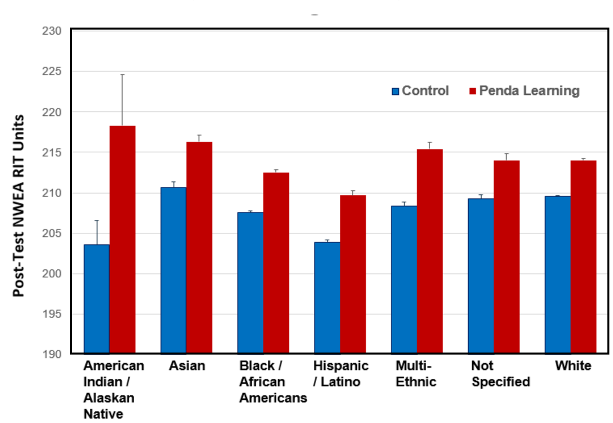


Figure 2 shows that students receiving Science instruction and intervention resources score higher on an independent, nationally normed measure (NWEA MAP Science Assessment) compared to students receiving "business as usual" classroom instruction.

Middle school serves as a critical inflection point. During grades six through eight, girls' enthusiasm for STEM subjects often decreases more rapidly than boys'. Contributing factors include entrenched societal stereotypes, limited visibility of female role models in technical fields, and social pressures that subtly discourage girls from pursuing or expressing interest in STEM [6]. Addressing these disparities is essential not only for gender equity but also for the future competitiveness and innovation capacity of the U.S. workforce.

Figure 3. Science knowledge gains were significantly larger for Penda Learning students. Gains varied by ethnicity for the Penda Learning and the Control group. (p<.05



In Figure 3, we observe differences at both the pre-test and post-test for both the control and Penda learning groups. At the post-test, the Penda learning group outperformed the controls for each reported ethnicity; however, the size of the improvements varied by group. A more detailed discussion of the gains by ethnicity will be discussed in a separate article.

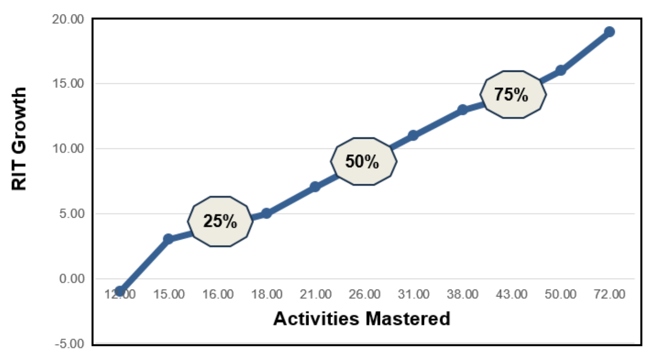
The improvements in science knowledge for each group by ethnicity are very encouraging. Despite growing national investments in STEM education, persistent racial and ethnic disparities remain in STEM interest, access, and achievement. Students from historically underrepresented groups, including Black, Hispanic/Latino, and Indigenous populations, consistently demonstrate lower rates of participation in advanced STEM coursework and careers compared to their White and Asian peers. These disparities emerge early and intensify through the K–12 and postsecondary pipelines [6]. These results suggest an opportunity to reverse the flow and grow the pipeline with these historically underrepresented populations. Importantly, these results show that all ethnicities represented in the study showed significant growth. Some of the initial differences persisted, albeit at a higher level of scientific knowledge. Overall, considerable progress was made in reducing these pre-existing differences within the Penda Learning group. We will discuss these issues more directly in a future article.

It is critical to recognize that addressing ethnic and racial disparities in STEM requires a multifaceted approach. Evidence-based strategies include providing early STEM enrichment programs in underserved communities, investing in culturally responsive curriculum and instruction, increasing representation among STEM educators and mentors, and expanding access to rigorous coursework and technology resources. Empowering historically marginalized students in STEM is crucial not only to improve equity among learners but also to foster national innovation and economic competitiveness.

C. Dose Response Effects on Science Knowledge

Figure 4 illustrates an interesting correlation between activities completed on the Penda Learning platform and growth on the NWEA MAP Science assessment. The distribution of student performance using the Penda Learning platform is identified with the corresponding growth in RIT.

Figure 4. A dose-response relationship between activities mastered and MAP Science assessment growth.

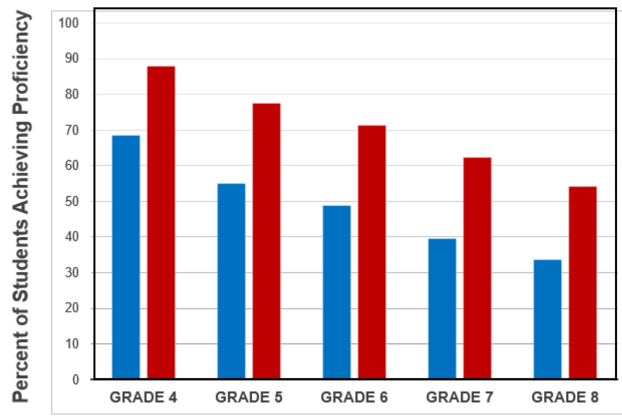


First, the relationship appears to be dose-related; the more activities mastered, the greater the academic growth in science. There is no evidence in the levels recorded to suggest that the top of the curve is flat at higher levels. This is important to note and to organize further research for a better understanding of faster learning. It is important to note that this is correlational data, and a more comprehensive understanding of this relationship should be replicated and extended in future research.

D. Achieving Grade Level Proficiency in Science Knowledge

In the NWEA MAP Growth assessment, a score of Level 3 is generally considered proficient for science. MAP Growth assessments are designed to measure student growth in various subjects, including science. Level 3 signifies a level of performance considered proficient for accountability purposes, marking the minimum level of achievement considered adequate according to NWEA [4]. In the following Figure 5 we see the percentage of students reaching level 3 (proficiency) for Penda Learning students and control students.

Figure 5. NWEA MAP Growth proficiency achievement for study participants. Penda Learning users outperform individuals in the active control group for each grade evaluated.



Proficiency scores are estimated for individual uses according to the NWEA Growth Assessment Manual and displayed in Figure 6. Penda Learning users outperform the percentage of control group students identified as proficient by approximately 20% for each assessed grade.

IV. Conclusions

A. Program Effectiveness and Group Differences

The statistically significant main effect of group was observed, indicating that students receiving science instruction through *Penda Learning* outperformed peers in the active control group on the NWEA MAP Science Assessment, a nationally normed and validated measure of science achievement. These results reflect robust gains across demographic subgroups and grade levels, controlling for pre-existing differences in baseline performance.

Prior evaluations of *Penda Learning* have demonstrated its effectiveness in improving student outcomes through standards-aligned content, structured independent practice, and scaffolding tools for English learners and students with disabilities [9-14]. The platform's support for differentiated instruction and teacher workflow automation makes it a valuable component of 21st-century science education.

B. *STEM Improvements and Economic Development*

The broader impacts of STEM education extend beyond test scores to economic mobility and workforce preparedness. High-quality science instruction requires not only research-based curricula but also sustained investment in professional learning for educators. Collaborative professional development programs enhance instructional effectiveness, increase student engagement, and support STEM retention. This is particularly important in middle school, a critical period for sustaining interest in science fields.

As technological advancements accelerate, equipping educators with up-to-date pedagogical strategies and tools ensures that students are prepared for a dynamic global economy. In this way, teacher training is both an instructional necessity and a workforce development strategy. This emphasis on educational development not only strengthens the quality of STEM instruction but also contributes to building a resilient workforce in teachers and students capable of adapting to an ever-evolving global economy.

C. *Summary Evaluation of Penda Learning*

This quasi-experimental study evaluated the impact of Penda Learning, a computer-based science intervention platform, compared to 'business-as-usual" controls on measures of science achievement across 46 schools and more than 18,000 students in Grades 3 through 8.

The Penda Learning platform is aligned with state science standards. Student science achievement was measured using a nationally normed independent assessment of science knowledge, the NWEA MAP Science Assessment. Testing was administered at two time points: the Fall (pre-test) and the Spring (post-test). A repeated-measures MANCOVA, using the pre-test as a covariate to control group baseline differences, found significantly greater gains among students in the Penda Learning group across all grades and demographic groups. It is noteworthy that a dose-response relationship was identified in the study results, indicating that higher levels of engagement and mastery with the platform corresponded to greater Science learning.

Effect sizes, calculated using Hedges' g, ranged from moderate (0.36) to large (0.82), with an overall average effect size of 0.56. A value that represents more than half a standard deviation of growth within a single academic year. Prior reviews of educational interventions have identified effect sizes of 0.20–0.25 SD or greater as the threshold for identifying educationally meaningful improvements used by various educational researchers and organizations [8].

D. *Alignment with ESSA Criteria*

Under the Every Student Succeeds Act (ESSA), educational programs must demonstrate effectiveness through one of four evidence tiers. These tiers help school and district leaders select interventions that are both impactful and aligned federal funding eligibility.

- <u>Evidence Tier 1: Strong Evidence</u> – Supported by a well-designed randomized controlled trial (RCT).
- <u>Evidence Tier 2: Moderate Evidence</u> – Based on a quasi-experimental study with statistical controls.
- <u>Evidence Tier 3: Promising Evidence</u> – Relies on correlational studies with controls for bias.
- <u>Evidence Tier 4: Demonstrates a Rationale</u> – Includes a logic model and plan for future evaluation.

ESSA requires that federally funded school improvement initiatives use interventions supported by evidence at least at the Tier 1, 2, or 3 level. Tier 4 may be used for innovative programs with a commitment to building future evidence.
For administrators, ESSA tiers provide a framework to justify program selection, ensure compliance, and maximize the impact on students. Prioritizing evidence-based solutions supports smarter budgeting, stronger accountability outcomes, and equitable learning gains across all student groups.

This study meets the criteria for Tier 2: Moderate Evidence under the Every Student Succeeds Act (ESSA), see Table 2. It employed a quasi-experimental design with baseline equivalence, used a psychometrically validated outcome measure (NWEA MAP), and included statistical controls for confounding variables. While use of covariates can slightly reduce power, this strengthens causal inference and internal validity.

With statistically significant subgroup effects on an independent and nationally normed science assessment and a clear dose-response relationship between platform use and science assessment performance, the findings satisfy this author's understanding of the federal guidelines to justify the program's consideration for program funding eligibility.

This study meets or exceeds the criteria for Tier 2 (Moderate Evidence) as defined by the Every Student Succeeds Act (ESSA) and used the U.S. Department of Education. The design included a large sample (>18,000 students), an active control group, a valid and reliable outcome measure (NWEA MAP Science), and statistical adjustment for baseline differences. While the use of covariates may slightly reduce statistical power, this approach strengthens internal validity and supports causal inference regarding program impact.

The average effect size, calculated using Hedges' g under a fixed-effects model, was 0.56 (95% CI [0.53, 0.59]), reflecting an average gain in performance of 21% for Penda Learning students. Adding to these finds are statistically significant demographic subgroup effects and a dose-response relationship, providing further evidence of the intervention's effectiveness across diverse educational contexts.

Table 2. The ESSA tiers of evidence framework aligned with the current study. [15].

Study Feature	Penda Learning vs Controls	Tier of Evidence
Study Design	Quasi-Experimental Design (QED)	Tier 2
Comparison Group	Yes, with baseline equivalence	Tier 2
Causal Inference	Yes, with pre-test covariates.	Tier 1
Statistical Significance Required	Yes, positive and for multiple subgroups (P<.05)	Tier 1
Sample Size and Setting	18,738 subjects across 46 schools	Tier 1
Measurement Validity and Reliability	NWEA MAP Science (valid, normed)	Tier 1
Examples of Acceptable Studies	Well-constructed QEDs with pre-post analysis	Tier 2
Purpose / Intended Use	Proven intervention ready for scale	Tier 1

E. *Summary*

This evaluation of the *Penda Learning Platform* provides rigorous evidence that technology-enhanced, standards-aligned science intervention can produce significant and educationally meaningful gains in student achievement. Students in the intervention group showed higher test scores, increased proficiency rates (a 21% gain), and broad effects across demographic categories.

Given the study's design, sample size, analytic rigor, and consistent outcomes, this intervention meets ESSA Tier 2 criteria. It contributes to the growing body of literature on effective STEM education resources and tools. As schools pursue both academic recovery and future readiness, investments in digital science instruction, accompanied by educator support and equity-focused strategies, represent a viable path to improved outcomes and sustained economic growth.

REFERENCES

[1] Hanushek, E. A., & Woessmann, L. (2010). The high cost of low educational performance: The long-run economic impact of improving PISA outcomes. OECD Publishing. https://doi.org/10.1787/9789264077485-en

[2] Jaffe K., Caicedo M., Manzanares M., Gil M., Rios A., Florez A., et al. (2013). Productivity in Physical and Chemical Science Predicts the Future Economic Growth of Developing Countries Better than other Popular Indices. *PloS One*, pmid:23776640

[3] Hanushek, E. A. and Woessmann, L. (2020) "Education, knowledge capital, and economic growth," in *The Economics of Education*, 2nd ed., S. Bradley and C. Green, Eds. New York, NY, USA : Academic, pp. 171–182, ch. 14

[4] NWEA. (2024). MAP Growth Science technical report. NWEA.Romer, P. M. (1990). *Endogenous technological change.* Journal of Political Economy, 98(5), S71-S102.

[5] Holian, L, Kelly, E. (2020) U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. STEM Occupational Intentions: Stability and Change Through High School. NCES 2020-167.

[6] American Association of University Women. (2015). Solving the equation: The variables for women's success in engineering and computing*. https://www.aauw.org/research/solving-the-equation/

[7] National Science Board, National Science Foundation. 2022. *Science and Engineering Indicators 2022: The State of U.S. Science and Engineering*. NSB-2022-1. Alexandria, VA. Available at https://ncses.nsf.gov/pubs/nsb20221

[8] Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, *49*(4), 241-253. https://doi.org/10.3102/0013189X20912798

[9] Penda Learning, (2023). Leveling Up Science Success: Transforming Challenges into Triumphs. Penda Science Helps Flagler County Schools Boost Scores on Florida State Assessments (SSA). Pp1-3.

[10] Penda Learning, (2018). Penda Learning Supports Hillsborough Middle School Science Classrooms with Varying Needs, pp1-3.

[11] Penda Learning (2024). Impact Report: How Penda Science Supports Improved Academic Outcomes in Sarasota County Schools, Florida. 2022-23 Academic School Year SSA and EOC Scores.

[12] Penda Learning, (2022) Penda Science Usage Increases Science Achievement in Charter Schools USA, pp. 1-13.

[13] Penda Learning (2022). Closing the Achievement Gap, Advancing Science Success. Penda Science's Impact on Clay County Schools' FL SSA and Biology EOC Scores. Pp 1-6.

[14] Conway, B. (2019) Engagement with Penda Learning Boosts Science Scores in Miami-Dade Public Schools. 1-7

[15] Regional Education Laboratory at American Institute for Research, (2025).ESSA Tiers of Evidence: What you need to know. https://ies.ed.gov/ies/2025/01/essa-tiers-evidence