Game-Based Learning Platform to Enhance Student Science Outcomes

Table of Contents

Game-Based Learning Platform to Enhance Student Science Outcomes

Impact Florida, Legends of Learning, and WestEd propose a mid-phase EIR project to enhance, scale, and perform a randomized controlled trial (RCT) of Legends of Learning, a game-based learning platform designed to improve science outcomes, particularly for high-need students. These students include historically marginalized groups in STEM (e.g., Black students), students in the lowest achievement quartile, and students from low-income backgrounds.

Impact Florida is a non-profit organization that partners with districts across Florida to connect practitioners, advocates, and leaders with experts, resources, and data, in order to contribute to a more knowledgeable, collaborative, and innovative field of educators who are supported in focusing on the right things for students. Impact Florida believes it can use its established network and experience in Florida to bring this project to life. Particularly with the support of Miami-Dade County Public Schools (see Appendix C), Impact Florida will have a strong network to begin scaling and support of the execution of the RCT.

This project addresses Absolute Priority 1, Absolute Priority 3 (STEM, Science), and Competitive Preference Priority 1 (Promoting Equity with Partners). To address the Competitive Preference Priority, Legends of Learning and WestEd will each support 3 years of internships from the minority-serving institutions with whom Impact Florida has partnered. The internships will allow students to get valuable experience with education and research organizations, which can lead to important connections and interest in related fields for their career.

The study will engage 280 teachers and 6,800 students in at least 80 elementary schools in Florida during the formative studies and the RCT. It will support the improvement and implementation of Legends of Learning, which will increase students' engagement in and attitudes towards science, opportunities to learn, and ultimately end-of-year assessment scores.

Section A: Significance

Despite heroic efforts by educators during the pandemic, many teachers struggled to hold their students' attention in remote and hybrid learning settings. This led to dramatic decreases in NAEP scores for both math and reading (NAEP 2022). With districts focusing on math and ELA to address COVID-related learning loss, subjects not yet tested by NAEP post-pandemic, such as science, were also hit hard. For example, post-pandemic scores for Florida fifth grade science dropped from 55% of students reaching proficiency in 2018 to 47% in 2021 (Florida Department of Education, 2018 and 2021). Fifth grade is an especially critical time for science education as adolescence frequently brings a drop in science interest (Bathgate & Schunn 2016, Bryan et al. 2011; Gottfried et al. 2001; Hawkey and Clay 1998; Osborne et al. 2003; Simpson and Oliver 1990). Student engagement with the subject material is one of the keys to increasing test scores (Christenson et al 2012).

Instead of schoolwork, students shifted their focus to other activities during the pandemic, specifically additional screen time. Between 2019 and 2021, the total amount of daily screen time went from 4 hours and 44 minutes to 5 hours and 33 minutes among 8–12 year olds, and from 7 hours and 22 minutes to 8 hours and 39 minutes among 13–18 year olds (Robb, 2022). Much of that screen time was spent on video games targeted to school-aged children such as Roblox, Minecraft, and Fortnite. For example, Roblox daily active users in the US and Canada jumped from 6.3 million in Q4 2019 to 11.4 million in Q2 2020 (Roblox, 2023). This has not abated since the pandemic, with Roblox daily active users reaching 13.3 million in Q4 2022.

Teachers recognize that engaging students is a problem as the "competition" for student attention from video games increases (Papadakis, 2018), and teachers are proactively choosing materials that are gamified or use games to teach, often without the knowledge or support of

their district administration (Turkay et al., 2014). Using their proprietary access to school district internet traffic data, Learn Platform found that 6 out of the top 10 content sites accessed most by teachers and students in the 2021–22 school year were gamified or used games to teach (the four others were YouTube, Wikipedia, Encyclopedia Britannica, and Nearpod) (Instructure 2022).

Investigation into the use of games for learning has grown from a small niche area to a major focus of research over the past 15 years (Gee, 2007; Tobias & Fletcher, 2011). Studies have demonstrated the potential of digital games to support learning in terms of conceptual understanding (Barab et al., 2007; Klopfer et al., 2009); process skills and practices (Kafai et al., 2010; Steinkuehler & Duncan, 2008); epistemological understanding (Squire & Jan, 2007; Squire & Klopfer, 2007); 21st century skills (Qian & Clark, 2016); and attitudes, identity, and engagement (Barab et al., 2009; Dieterle, 2009; Ketelhut, 2007). Reports by the National Research Council (NRC) and others (Martinez-Garza et al., 2013; NRC, 2011; Young et al., 2012) have also acknowledged this potential.

The same interactivity and hand-crafted engaging designs that confer the benefits of game-based learning (GBL) also lead to high development costs. Such high costs associated with developing games means covering large portions of the curriculum remains cost-prohibitive, and the academic literature has focused on small-scale studies of individual games in controlled environments. In a highly cited meta-analysis of 69 high-quality GBL studies from 2000–2012, while overall results from media comparisons indicated that digital games significantly enhanced student learning relative to non-game conditions ($g = 0.33$, 95% confidence interval [0.19, 0.48], $k = 57$, $n = 209$), each study had an average of only 100 participants (Clark et al., 2016). These small-scale studies demonstrate the potential for game-based learning, but they have not yet convinced district administrators to widely adopt GBL as a reliable supplement to textbook or

lecture-based instruction. Furthermore, uncertainty around how well GBL can align with specific curricula and instructional materials have inhibited widespread administrator buy-in, creating a barrier to scaled implementation (Vu and Feinstein 2017).

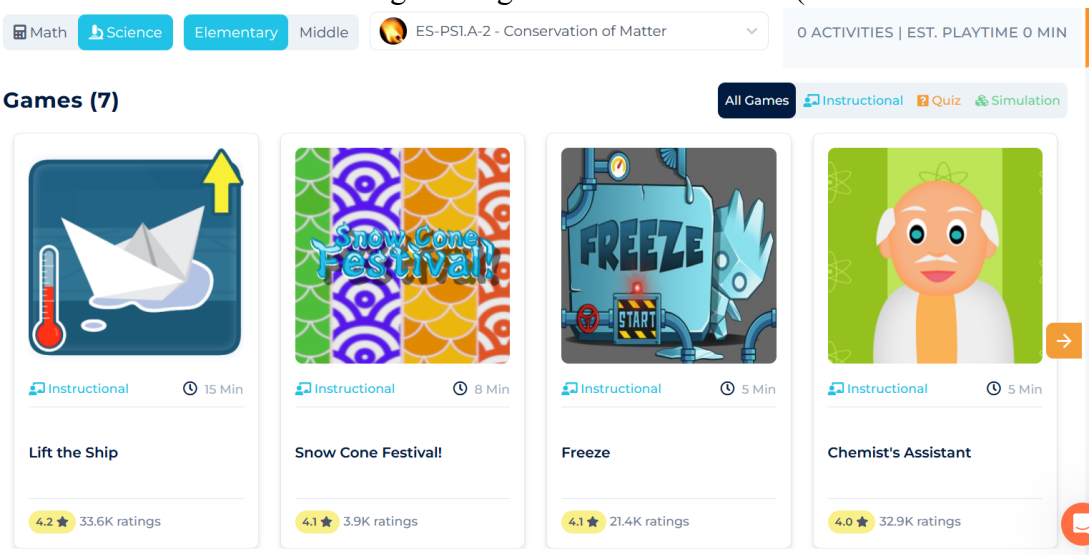
Given the extraordinary focus of teachers on GBL as evidenced by their usage of GBL tools, and the remarkable desire by students for games as evidenced by their own screen time choices, it would be of great significance to understand whether a full year of a GBL supplement can compete with a standard, non-GBL curriculum, especially in science, which lends itself well to GBL as it deals with natural phenomena that can explored through simulation-like games. For example, **there are no year-long studies on game-based learning in science in the WWC**, and only two on gamified math tools (Dreambox and ST-Math). The recent addition to the WWC of a Tier 1 study of Mission Hydrosci was for a 6 to 8 day game-based curriculum that teaches very specific water systems science and scientific argumentation lessons (and cost was listed as a major barrier) (Reeves et al 2020). It may be hard for a district administrator to generalize the benefits of GBL from a 6 to 8 day curriculum, or from single games such as Dreambox and ST-Math, which are built by specific teams with a particular game design philosophy.

But administrators and practitioners know that they cannot use the same game all year round. An expensively built, rich, immersive, hours-long game like Mission Hydrosci is a treat for a special occasion to teach a handful of standards, but it is too expensive to build into year-long instruction, and most of the time teachers need shorter, simpler modules that can be easily slotted into any lesson. **To allow educators to feel comfortable about shifting to GBL-supported curricula, it would be useful to have research on a general purpose GBL supplemental curriculum that spans the full year scope and sequence.** Such a study should evaluate the GBL-supplemented curriculum relative to regular instruction across a whole year,

and reaches a sophistication and scale that teachers expect from the tools they use on Learn Platform's list. To date, such a broad GBL curriculum has been cost prohibitive and therefore not explored as possible best practices by Impact Florida's various school and district initiatives.

Recently, Legends of Learning has been able to overcome the issue of high production costs for games that cover large swaths of the curriculum through an innovative game development model. Legends of Learning's "Learning Universe" product is a curated platform of over 2,000 learning games built by over 800 different professional game studios around the world. Each game covers exactly one curriculum standard: a single Next Generation Science Standards' (NGSS) Disciplinary Core Idea (DCI) bullet point (also incorporating Crosscutting Concepts and Science and Engineering Practices). Each standard contains 3–10 games for teachers to choose from (see Figure 1). No one studio bears the cost of developing games across a large curriculum, and studios compete to make the best game in a marketplace model. Studios share in the revenue Legends of Learning generates based on the playtime their game receives. The games are also aligned to the state's curriculum, so in Florida the NGSS aligned games must be placed in the correct Next Generation Sunshine State Standard (NGSSS).

Figure 1. A teacher's account showing 4 of 7 games for ES-PS1.A.2 (Conservation of Matter).



Looking at the marketplace from the teacher's perspective, **games can be assembled into assignments alongside assessment items, videos, and simulations in the Learning Universe product**. Teachers typically use Learning Universe as whole-class instruction, assigning a playlist of games to the full class as a supplement to the standards being covered in the regular curriculum. For example, if a teacher is covering ES-PS1.A.2 (Conservation of Matter) this week, they can create an assignment with the following structure (taking games from the figure above): [assignment start] [3 question pre-assessment] → ["Lift the Ship" game] → ["Chemist's Assistant" game] → [3 question post-assessment] [assignment end]. This assignment should take each person in the class approximately 30 minutes (15 minutes for Lift the Ship, 5 minutes for Chemist's Assistant, and 5 minutes each for the pre- and post-assessment). Over a full school year, these assignments can then cover the whole scope and sequence for a district, allowing teachers to incorporate the games into their regular instructional materials. In addition to the Learning Universe marketplace, teachers can also send students into the "Awakening" product, which offers gamified assessments as students explore an open world and collect and battle pets called Beasties. Teachers can leverage the assessment data collected by Awakening to then launch Learning Universe assignments that cover the gaps identified during Awakening gameplay. This approach has allowed the Legends of Learning product to achieve significant scale among teachers looking for a rigorous GBL product to engage their students, and overall, 2.7 million students across the United States used Legends of Learning in the 2022–2023 school year. **See Appendix J.1 for screenshots of the features** described above.

Unlike games like Dreambox, ST-Math, and Mission Hydrosci, the games are not built by a single team with a single design philosophy. Other than the requirements that games 1) must fit a single class period, 2) must work on a browser on a low-cost device (i.e. chromebook, tablet, or

computer), and 3) must teach exactly one standard, there is no single game design philosophy to affect academic results (positively or negatively). As the wide variety of game types built by a large number of professional game studios are deployed, tested in real-world classrooms, and rated by teachers, higher quality games rise to the top. As a result, Legends of Learning has the capacity to demonstrate how GBL can be an enhancing supplement embedded into regular instruction, which would be significant for researchers and practitioners who want to extract learnings about GBL overall rather than any specific game.

Studies have shown that GBL in general can be effective (Clark et al., 2016), and Legends of Learning specifically can be effective in raising test scores (Clark et al., 2018). In the 2018 study, published in the Journal of Learning Sciences, Legends of Learning games were found to be effective at helping students learn curricular material as compared to control groups taught by the same teacher. The quasi-experimental design study enrolled more than 1,000 students across six states. Half were taught the standard curriculum and half received games as a substantial part of their curriculum. The post-test revealed a statistically significant increase in test scores by the game-based learning students, and the teacher survey revealed increases in student engagement (movement from 3.23 to 4.31 out of 5, where 5 is “strongly agree”) and high scores for focusing high-needs students on the learning task (4.23 out of 5, where 5 is “strongly agree”) and desire to continue using GBL (4.46 out of 5, where 5 is “strongly agree”).

A more recent correlational study conducted by WestEd researchers found students who used Legends of Learning in four large, diverse Southeastern US districts received higher scores on their end of year standardized state science assessment than those who did not use this product (McKinney et al., 2023). During the 2020–2021 school year, nearly 10,000 students utilized Legends of Learning in these four districts, compared to 14,000 students who did

not. The positive effect was dose-dependent, so the more students used Legends, the higher their scores were. Because a relatively small number of teachers utilized the platform in its intended two activities per week, further study is needed to determine the effect of dosage on student science achievement. However, the correlational study revealed promise of this relationship, as intended usage was correlated with a 23 percentile increase in test scores. A follow-up quasi-experimental or randomized control trial design is warranted to see whether positive effects on the state science assessment would occur with a more rigorous sampling design.

Unfortunately, insufficient usage by teachers arises directly from the marketplace model inherent in Legends of Learning, where it is the teacher's responsibility to find the proper games, assemble playlists, and assign them to their classes. The marketplace model offers more freedom and a much wider variety of GBL curriculum supplements, but it also imposes a large responsibility on the teacher. Some teachers find the process enjoyable and create many assignments, but many are confused or overwhelmed, and fall back on their standard materials.

With intense interest in GBL from students, teachers, and researchers, it would be significant to study how GBL-supplementation can affect student test scores over a full year of instruction. Legends of Learning's innovative marketplace development model has overcome the historical cost issues associated with creating wide GBL curriculum coverage, but the lack of a well-designed implementation model, combined with a lack of easy mechanisms to integrate the learning games into district curricula, has led to insufficient usage by teachers as evidenced by prior WestEd research (McKinney et al, 2023). Legends of Learning does offer **professional learning** (PL) for partner districts, and this PL involves training teachers on how to use the Learning Universe platform in virtual sessions and optional self-paced training (described in Appendix J.2), but further support may be needed. This proposal aims to deliver on the promise

of GBL **by building appropriate district supports and lowering the burden on teachers to help overcome the barriers to scaling identified in prior studies. Each strategy will be rigorously tested, culminating with a rigorous randomized control trial of a full year GBL science supplemental curriculum**. Scaling research-backed effective teaching practices and materials such as GBL aligns with the mission of Impact Florida, and we are excited to expand this impact through this project.

Section B: Strategy to Scale

B.1 Barriers and Strategies

Barrier 1: While game-based activities and resources are aligned with multiple state standards, districts do not have an easy mechanism to integrate them into their curricula.

A primary barrier to scaling is that educators have historically been skeptical about how well GBL solutions align with and can support the scope and sequence of their particular curriculum (e.g. Kaimara et al., 2021). Most GBL products cannot be adapted to meet individual districts' needs, including supporting their adopted curricula and the implementation models of their courses. Customized curricular alignment will ensure that Legends of Learning can consistently support implementation of standards-based curricula across multiple districts and states. To overcome this barrier, we propose the following strategies:

Strategy 1: Currently, the games are aligned using an overly simplistic algorithm that is generic enough to align games to all relevant state outcomes. We propose to perform a comprehensive audit of the existing games to determine how well the algorithm worked for their alignment specifically with Florida's NGSSS. Following the audit, we will establish stricter, state-specific criteria for alignment for all existing and future games submitted to the Legends of Learning platform. This enhanced alignment will ensure that districts in Florida, and eventually other

states, will be able to more easily access a collection of games that meet their state standards and are easily integrated into their chosen curriculum (see Strategy 2).

Strategy 2: We propose to **develop a "district-endorsed game" feature** where district personnel can recommend or endorse the use of specific games to support particular lessons or units of instruction, specific for their curriculum. For example, in preparation for teaching photosynthesis, Legends of Learning offers teachers a collection of games to choose from as they plan their lessons. While options are a benefit, many teachers do not have the time to make the most informed selection. The new feature will allow districts to create templated lessons using the "endorsed games" which can be accessed and used by all teachers in the district.

Additionally, research on high-quality instructional materials recommends that decisions about supplementation of adopted standards-based curricular materials be made collectively at the school or district level, rather than by individual teachers (Polikoff 2019). This saves teachers time, leverages district staff expertise, and supports more equitable learning experiences by reducing variations in students' learning experiences from classroom to classroom.

Barrier 2: Scalable implementation models have not previously been tested and evaluated.

The second barrier to scaling, identified in WestEd's research with four large Southeastern districts (McKinney et al., 2023), is inconsistent or unintended usage of the Legends of Learning platform, which may prevent optimal efficacy. Based on teacher feedback, the issue stems from challenges integrating GBL into day-to-day curriculum, a lack of customized usage guidance, and a lack of a research-informed implementation model for teachers. As a result, teachers have been left to figure out how, when, and how often to use games in their lessons. Easing the burden on teachers by testing implementation models is needed to establish an evidence-based scalable implementation plan for Legends of Learning that will support customizable, standards-aligned

game-based learning experiences. To overcome this barrier, we propose the following strategies:

Strategy 3: To ease the burden on teachers, we propose to **develop a personalized learning mode** for students in Awakening. Currently, the Awakening product assesses student content knowledge in a gamified world, and surfaces that assessment data to teachers, who must then create an assignment to address any learning gaps using the Learning Universe marketplace. The proposed personalized learning mode would immediately offer a top rated learning game within Learning Universe to the student once a gap was identified, allowing the student to progress through content aligned to the state standards at their own pace. In schematic form, this would be an ongoing repetition of [Assess] → [Identify Learning Gap] → [Offer Instructional Game] → [Assess]. Combined with Strategy 2, this would reduce the burden on teachers who may not have sufficient time or expertise to create their own GBL assignments, while also allowing additional uses of Legends of Learning, including support for students who may be behind. This feature, which is embedded in the gamified assessment world of Awakening, will require significant game design and game quality assurance testing and optimization, which may be facilitated by a multi-year internship program with partner minority-serving institutions for students who want game design or software development experience. Development and use of this feature will be tested in the formative evaluation phases (see D.4).

Strategy 4: We propose to develop an implementation model through iterative design, testing, and development with feedback from teachers and districts. During the first year and a half, WestEd will complete usability and feasibility studies to inform the design of 3–4 potential high impact implementation models, which will then be tested in a full-year implementation study (described in D.4). Evidence will be used to finalize a scalable implementation plan for Legends of Learning which will be tested during the impact study (see D.1).

Strategy 5: We propose to provide comprehensive, ongoing support for districts and teachers to improve implementation fidelity. To ensure all teachers know how to both incorporate the games into their teaching and draw on resulting data to inform their instruction, the implementation plan will include a PL component for both teachers and districts. The associated PL will provide initial support for districts with making customized recommendations to teachers regarding the best games to support their curriculum and support for teachers around understanding the platform and incorporating games into their teaching. The plan will also include ongoing support with best practices and tips for intentionally and effectively utilizing both the games and their resulting data to improve student outcomes.

B.2 Management Plan

Impact Florida (IF) will lead the proposed project, collaborating closely with Legends of Learning (LL) and WestEd (WE). Appendix J.3 provides an organizational chart, which illustrates the responsibilities of each staff member and organization. Figure 2 provides a high-level project timeline, while Table 1 lists the project objectives by year and responsible organization(s). The bolded font for an organization indicates that they are the lead on the milestone. **Appendix J.4** provides a detailed timeline, with objectives and responsibilities designated for each month of the five-year project.

Figure 2. High-Level Project Timeline, by Objective

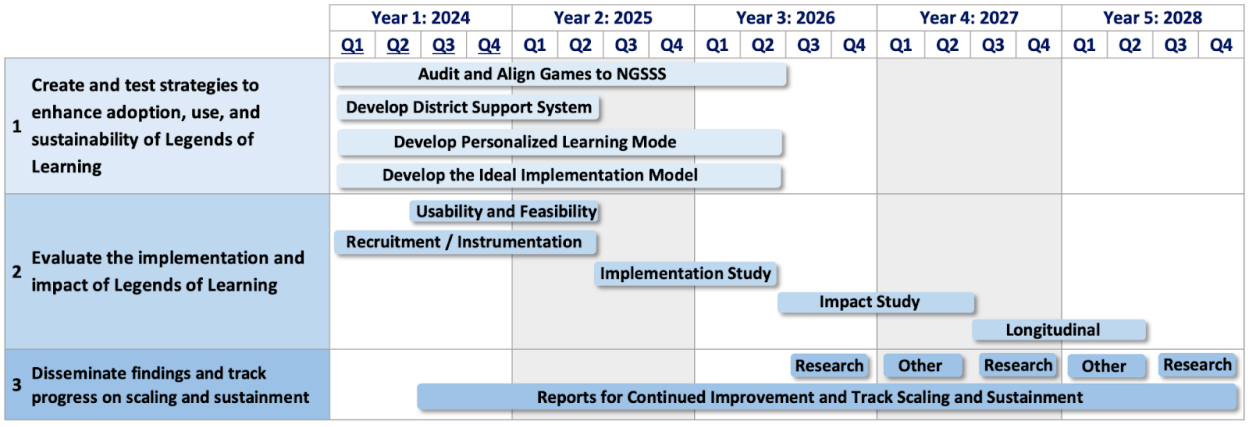


Table 1. Project Objectives and Responsibilities, by Organization and Project Year

Milestone	Y1	Y2	Y3	Y4	Y5	Responsible Organization
Objective 1. Create and test strategies to enhance use and sustainability of Legends of Learning						
1.1 Audit existing NGSSS alignment and update alignment criteria.	X	X				**LL**
1.2 Develop and implement a "district-endorsed game" feature.	X	X	X			**LL**, WE
1.3 Improve and implement the "personalized learning mode" feature.	X	X	X			**LL,** WE
1.4 Explore, test, and finalize a scalable implementation model.	X	X	X	X		**WE**, LL
Objective 2. Evaluate the implementation and impact of Legends of Learning						
2.1 Prepare data collection instruments and procedures and train staff.	X					**WE**
2.2 Identify and randomly assign to treatment and control conditions.	X					**WE**
2.3 Measure and assess implementation fidelity of Legends of Learning.	X	X	X	X		**WE**
2.4 Measure and assess the impact of Legends of Learning.		X	X	X	X	**WE**
Objective 3. Disseminate findings and track progress on scaling and sustainment						
3.1 Disseminate findings to research audiences.			X	X	X	**WE**, IF
3.2 Disseminate findings to public, teacher, and policy audiences.				X	X	**WE**, IF, LL
3.3 Support adoption and track cost, scaling, and sustainment.		X	X	X	X	**WE**, IF, LL

B.3 Capacity to Bring the Project to Scale

Impact Florida, Legends of Learning, and WestEd are capable of bringing the proposed project to scale. If funded, Impact Florida will allocate a part-time project coordinator to oversee the identification and participation of schools and teachers and to lead internal project management. Senior leadership will help the coordinator by using their deep connections with districts and schools in which they have worked. Impact Florida will then use a project assistant to coordinate with IT and school personnel and to assist with support tasks as needed. They have expertise in carrying out large scale projects, including reporting and accountability. Impact Florida has already secured a commitment from Miami-Dade County Public Schools, who is eager to evaluate GBL as a full-year science supplement, as a study district (see Appendix C).

WestEd is a preeminent educational research, development, and service organization with over 900 employees and 13 offices nationwide. WestEd has been a leader in moving research

into practice by conducting research and development (R&D) programs, projects, and evaluations. Over the past five years, WestEd has carried out almost 2,500 successful projects representing major contributions to the nation's R&D resources. At any given time, the agency has between 450 and 700 active contracts and grants. WestEd staff on this project will include senior personnel with expertise in science education, online and game-based learning, formative and summative evaluation, advanced quantitative methods and statistics, and research implementation. These staff have a proven track record of successfully executing projects of similar complexity, scope, and focus, including multiple large-scale, multi-site RCTs.

Legends of Learning is the industry leader in GBL technology, and has ample technology and education specialists to help achieve its responsibilities on the project. They will work closely with Impact Florida and WestEd as they develop each new feature to ensure they are properly tested and iteratively improved to have maximum impact for the teachers and students in the study population. They operate in over 10,000 schools in 49 out of 50 states and have resources and processes to onboard and sufficiently support the number of schools in the study. They have PL staff, but may need to expand capacity to support any new coaching or on-going support models that emerge from WestEd's formative research (see D.4).

Table 2 provides details on the capabilities of the key personnel and their project roles. Impact Florida will be responsible for managing the project and ensuring progress on meeting objectives. They will also share responsibility for recruiting, onboarding, and implementation monitoring. Legends of Learning will lead development while WestEd will lead a series of formative evaluation studies to improve new features and potential administration models. Legends of Learning will implement all training and support of its product. WestEd will be

responsible for leading the independent impact evaluation and leading implementation fidelity analyses. All parties will contribute to dissemination of study activities and findings.

Table 2. Roles, Responsibilities, and Relevant Experience of Key Project Staff

Staff and Role	Primary Responsibilities and Relevant Experience
Impact Florida Key Personnel	
[redacted], President, leads and manages all aspects of the organization's strategy and relationships. She is a passionate advocate for providing equitable educational opportunities for all students in Florida. [redacted], Chief Operating Officer, leads the organization's operations and activities. As former President of the Jacksonville Public Education Fund, he has a history of contributing to smart improvements in schools that help all students succeed. Together they will: serve as the key point of contact for the funder, LL, and WE; oversee the IF staff and partner relationships; support the identification and onboarding of schools; and contribute to dissemination of results including annual reports.	
WestEd Key Personnel	
[redacted] is the Senior Director of STEM Research and Entrepreneurship at WestEd, and brings decades of experience managing large-scale research projects and center grants. [redacted] will serve as WestEd's Principal Investigator, key point of contact, and manage the project and partner relationships. [redacted] is a Senior Research Associate II and evaluation lead in two EIR grants. He will assist with project management and lead the development of measures, collection processes, and analyses of implementation data. [redacted], Director of STEM Networking and Partnerships, will be the lead support for Impact Florida and school liaison for WestEd. [redacted] is a Senior Research Associate in WestEd's Learning and Technology team and brings content expertise around online learning and programs. [redacted] is a Research Associate II in Science Education and brings expertise in elementary science learning and evaluation. [redacted] is a Senior Research Associate at WestEd. He brings expertise and extensive experience in applied statistics and psychometrics and has led many research and measurement projects funded by the USDOE and NSF. [redacted] will lead the quantitative data analysis in all stages of the study.	
Legends of Learning Key Personnel	
[redacted], CEO, has a strong history of developing successful and effective education products, and will serve as Legends' key point of contact, manage the project and partner relationships, and lead progress monitoring. [redacted], CRO / CAO, has 20 years of experience in leading curriculum design to support learning and effective pedagogy and will lead instructional design and education vision of the product and features. [redacted] Principal Engineer, has worked on dozens of project teams leading the digitization of products and information systems. He will lead the development of new features. [redacted] Director of Partner Success, specializes in implementing processes that lead to high customer retention. She will lead partner relationships between Legends and the districts/schools.	

B.4 Dissemination Strategies

To reach research, policy, and practitioner audiences interested in issues related to implementation and impact, WestEd plans to present research findings at national research conferences (e.g., AERA, SREE, NARST, etc.) and publish findings in peer-reviewed journals (e.g., Journal of Research on Educational Effectiveness, Educational Researcher, Science Education, Journal of Research in Science Teaching, etc.). In addition, WestEd will submit presentations with at least one science leader from Florida schools to national practitioner or policy annual conferences (e.g., ASTE, ISS, National Rural Education Association, etc.). WestEd will also highlight key findings through all of its communication networks (e.g. WestEd.org, Insights blog, R&D Alert Online, WestEd E-bulletin, etc.).

Legends of Learning will use the findings from this study to inform other educators interested in the program about the evidence from this study. They will present at national practitioner-oriented conferences facilitated by organizations such as NSTA, ISTE, FETC (Future of Educational Technology Conference), and NSELA (National Science Education Leadership Association). In Florida, Legends of Learning will present at FASS (Florida Association of Science Supervisors) and FAST (Florida Association of Science Teachers). Legends of Learning will keep full-time staff based in Florida who will attend statewide and local conferences, regularly visit schools, and provide unlimited, ongoing support to teachers and schools. It will disseminate findings on its website and through its communication channels to all existing and potential partner states.

Impact Florida will share learnings and promote findings that are useful to practitioners through a variety of dissemination channels, including its annual statewide Education Summit. Impact Florida will leverage its strong base of knowledge management expertise to design

products for dissemination, utilizing our website, social channels, email base of state education leaders and policymakers, and direct outreach through the strong network of relationships Impact Florida holds with a variety of education stakeholder organizations both within and beyond Florida.

B.5: Utility

The proposed project's resulting evidence-based implementation plan will support broad, consistent implementation of Legends of Learning to improve instruction, address district leader and practitioner concerns about the implementation and benefits of GBL, and contribute to the overall knowledge base regarding best practices in the use of GBL to improve student outcomes. The implementation and professional learning plans built during the study will guide expansion and effective impact in both new and existing sites (which span 49 states). Additionally, the new tools will make the implementation easier, particularly the district recommendation and lesson building supports, which will ease the burden on teachers and make the product more appealing. The addition of the personalized learning mode will not only enable effective use across a variety of settings, it will also ensure that it serves diverse learners within those settings. These features build upon a product which has shown to already have a positive relationship to student outcomes and could increase both uptake and impact of Legends of Learning.

Section C: Project Design

C.1 Conceptual Framework

Legends of Learning aims to create and support a lesson building and learning environment that is known to relate to positive student outcomes. This environment is labeled as Inputs in the Logic Model (see Appendix G). Legends of Learning has then generated a theory that identifies four high-leverage drivers that can contribute to increasing district and teacher

buy-in and use of their program, which can lead to increased engagement, opportunities to learn, conceptual understanding, and student standardized assessment outcomes. These drivers are PL supports, state specific alignment, district supports for integration, and teacher and student use. The logic model represents these four drivers as Output Activities, which lead to student engagement and participation and builds towards key outcomes (see Table 3).

Table 3. Key evidence in support of each driver.

Conceptual Framework for the Impact of Legends of Learning
Professional Learning Supports
These supports include a blend of virtual sessions, self-paced training, and coaching with on-going support. These sessions use active learning to engage teachers in the same discourse and activities their students will experience to get them comfortable with and exploring the student features (Darling-Hammond et al., 2017). As teachers better understand the conceptual underpinnings of what the approach, they are more likely to be comfortable fostering scientific discourse that encourages students to share ideas and discuss varying approaches to solving problems (Bishop, 2016; Hill et al., 2005). This comfort and improved use will also encourage teachers to implement the lessons with fidelity.
State Specific Alignment
Legends of Learning will perform a comprehensive audit of the existing games to determine how well the algorithm worked for their alignment specifically with Florida's NGSSS. The closer alignment will make games easier to find, will create more confidence in district leadership and teachers (Kaimara et al., 2021), and can improve impact (Squires, 2011). This will make the district support tools more effective and increase transparency around learning progressions when students are using Awakening, the personalized learning tool.
District Supports for Integration
These supports will allow districts to identify and recommend games specifically aligned with district-specific curriculum. These further the impact of the NGSSS alignment and gets district leadership involved, which can increase implementation support (Whitworth & Chiu, 2015).
Teacher and Student Activity
Ultimately, the impact comes from teachers using the lessons and students engaging with the materials and experiencing the personalized learning tool (as needed). The games and corresponding assessments, videos, and simulations can encourage student engagement and increase opportunities to learn, understanding of the concepts, and eventual standardized test outcome (e.g. Grabau & Ma, 2017; Lee et al, 2021).

C.2 Project Objectives, Outcomes, and Measures

We expect that the project will reach approximately 6,000 students in 5th grade across 80 schools in Florida. To achieve the goals and objectives described in Table 1 (**see B.2**), we propose to (1) audit and align games to NGSSS; (2) develop a district support system including recommended games for teachers; (3) develop a personalized learning mode; (4) develop an ideal implementation model; (5) implement and test the efficacy of Legends of Learning in our high-need sample; and (6) disseminate findings and track progress on cost and sustainability A table detailing a full set of tasks and timelines is seen in **Appendix J.4**.

C.3.a Population Description

For this study, we define high-need students as "students at risk of educational failure or otherwise in need of special assistance and support, such as students who are living in poverty, who attend high-minority schools . . . [or] who are far below grade level" (US DOE, 2012). Florida has nearly 3 million school students enrolled in over 4,000 schools. Over 200,000 students participate in 5th grade science each year. The state has a diverse population of students, including nearly 64 percent minority students (see Table 4).

Table 4. Florida Public Education Enrollment, by Race and Ethnicity, for 2021-2022

Race / Ethnicity	*AI / AN*	*As/AsPI*	*Bl*	*His*	*Haw / PI*	*Wh*	*Multi*
Percent of students	0.3%	2.8%	21.2%	35.5%	0.2%	36.1%	4.0%

Source: NAEP (2022).

Note: AI / AN = American Indian/Alaskan Native; As = Asian/Asian Pacific Islander; Bl = Black; His = Hispanic; Haw / PI=Hawaiian Native or Pacific Islander; Wh = White Non-Hispanic; Multi = Multi-Ethnicity or Unknown.

Florida has 52% economically disadvantaged students (Florida Department of Education [FLDOE] 2023). Additionally, 2022 Florida Department of Education data shows that 52 percent

of 5th grade students in Florida are performing below satisfactory on Florida's statewide standardized assessment (FLDOE, 2022).

C.3.b How Legends of Learning Will Address Students' Needs

As discussed in *Section A*, the existing Legends of Learning program has shown success in multiple other states, and will improve its tools to optimally address the needs of Florida's students. As discussed above, giving districts the power to review and recommend standards-aligned games to teachers based on the districts' chosen curriculum and particular population of students will allow for impactful and easy to implement lessons. These lessons will provide high needs students with access to engaging high-quality supplements to enhance learning. The personalized learning mode will also allow students to be recommended games that are specifically designed to address their knowledge gaps. Teachers can utilize this feature for students who may be struggling with certain areas of content or those that may need more time or extra support with a difficult topic. The implementation models that we will test and the one ultimately adopted will be tailored to have maximum impact on the students the districts serve. They will consider Florida state standards, the curricula in the districts, the amount of days and time that teachers perceive as appropriate for enacting GBL in the classroom, and any other contextual factors discovered during the formative stages of the project.

Section D: Project Evaluation

WestEd will conduct an independent evaluation of the implementation and impact of Legends of Learning on 5th grade science classes (see Table 5) with the impact study designed to meet WWC standards without reservation (see D.1). Following extensive formative evaluation (see D.4), the impact study will use research questions (RQ) 1–3 to study the impact of the intervention on teacher and student outcomes, including a WWC acceptable state standardized

test for student achievement. Questions 4 and 5 address implementation and are designed to provide both performance feedback during initial stages and guide measurement towards acceptable implementation thresholds, and to document key factors that should be considered during further replication and scaling of Legends of Learning. Questions 6–8 explore mediating and moderating effects, which will unpack how key project components and contextual factors can influence outcomes. Questions 9–10 are exploratory but will provide evidence of sustaining impact of access to Legends of Learning and could provide evidence of replicability.

Table 5. Evaluation Research Questions and Data Sources

Research Question	Primary Data Source(s)
Impact analyses	
1. What is the impact of Legends of Learning on students' science achievement?	Florida 5th Grade Statewide Science Assessment
2. What is the impact of Legends of Learning on students' opportunities to learn?	Opportunities to Learn scales (on a study-administered teacher log)
3. What is the impact of Legends of Learning on the nature of teachers' instructional activities?	Instructional Activities scales (on a study-administered teacher log)
Implementation analyses	
4. To what extent is Legends of Learning implemented with fidelity? How do teachers perceive the professional learning? How many activities are students completing compared to expectation?	Training sign-ins, Legends of Learning teacher and student data, teacher monthly log
5. How does the implementation of Legends of Learning differ across school contexts and teacher and classroom characteristics? What factors hinder or facilitate the implementation of Legends of Learning?	Training sign-ins, Legends of Learning teacher and student data, teacher monthly log, including relevant background and demographic data
Main mediating and moderating analyses	
6. To what extent does the impact of Legends of Learning differ across school contexts, teacher and classroom characteristics, and student characteristics?	All data for RQs 1–3, including relevant background and demographic data
7. To what extent is the impact of Legends of Learning on student outcomes mediated by instructional activities and opportunities to learn?	All data for RQs 1–3
8. To what extent is the impact of Legends of Learning on student outcomes mediated by the dosage of Legends of Learning activities?	All treatment data for RQs 1–3, plus Legends of Learning teacher and student data

Exploratory analyses	
9. To what extent is the impact of Legends of Learning larger in teachers' second year of implementation?	All data for RQs 1–3
10. After the control teachers are trained, are their students' outcomes comparable to the outcomes of the original treatment group?	Same data as RQ 3

Impact Study Summary. Half of 80 recruited schools in Florida will be randomly assigned to implement Legends of Learning and the other half will continue with business-as-usual science instruction and materials (see D.1). Recruitment will target schools with a high percentage of high needs and underserved students, but will ensure an appropriate range of demographics for generalization. The sample will consist of all fifth grade teachers and students in participating schools. This design will allow participation from approximately 240 teachers (120 in each condition) and 6,000 students for the impact study. The 3,000 students in the control condition will experience Legends of Learning the following year. The proposed study is powered for a minimum detectable effect size of 0.11 to 0.12 for student outcomes, and 0.28 to 0.32 for teacher instructional activities. (See Appendix J.5 for details about the power analysis and Appendix J.6 for details about the hierarchical linear models for the impact analyses.)

WWC Acceptable Outcome: Student Science Achievement. To measure students' outcomes for RQ1, WestEd will use the Florida 5th Grade Statewide Science Assessment, which is considered valid and reliable by the WWC standards and measures student achievement of the Next Generation Sunshine State Standards (NGSSS) in science (CPALMS, 2019). The assessments were developed by a rigorous standards setting process to ensure high validity and reliability (Florida Department of Education, 2010).

Outcome: Student Opportunities to Learn. To measure students' opportunities to learn in fifth grade science (RQ2), WestEd will administer a student survey at the end of the intervention year. The survey will include four opportunities to learn measures adapted from

studies that provided evidence of sufficient validity and reliability for the measures (Rickles et al., 2019; Walters et al., 2018): opportunities to make real-world connections (Cronbach's α = 0.84), opportunities to justify reasoning (α = 0.82), opportunities to solve challenging problems (α = 0.78), and opportunities to demonstrate conceptual understanding (α = 0.80). (See Appendix J.7 for the complete set of items.)

Outcome: Student Engagement. All students will complete the Engagement items from the 5Essentials Survey as part of the end-of-year survey. These items measure student interest and engagement in the class as a whole, and have shown strong reliability ($\alpha > 0.77$) and validity (Hart et al., 2020; Hart et al., 2021).

Outcome: Student Attitudes towards STEM. All students will complete the science portion of the Student Attitudes Toward STEM (S-STEM) survey (Friday Institute for Educational Innovation, 2012) as part of the end-of-year survey. The survey is intended to measure changes in students' self-efficacy and outcome expectancy in STEM subjects and interest in STEM careers. Unfried et al. (2015) calculated reliability and found Cronbach's α values above 0.83 for all constructs in the survey.

Outcome: Teacher Instructional Activities. To measure the quality of instructional activities (RQ3), WestEd will administer a teacher log three times across the year during the impact study to provide an accurate picture of the instructional activities over the entire school year. Prior studies of teacher logs indicate that they can be a valid and reliable measure of instruction (Rowan & Correnti, 2009). The log will include the following measures adapted from a RAND study of inquiry-based instruction (Le et al., 2006): inquiry-based practices intended to actively engage students and promote problem solving skills (α = 0.83), inquiry-based activities intended to facilitate critical thinking (α = 0.77), discussion (α = 0.74), and developing

conceptual understanding ($\alpha = 0.58$). Together, these measures capture the types of instructional activities expected to be seen in conjunction with Legends of Learning in the classroom (see Appendix J.7 for the complete set of items). In addition, WestEd will perform at least 60 observations of treatment and control classrooms during the impact study, with the aim of our observation sample including at least half of the schools.

D.1 Meeting WWC Standards Without Reservations

The evaluation of the impact of Legends of Learning will be based on a school-level randomized controlled trial designed to meet What Works Clearinghouse standards without reservations (WWC, 2020). Participating schools will be assigned to the treatment or control condition using blocked random assignment. Blocks will consist of school-level demographic information and prior science achievement (see Appendix J.6).

After randomization, Legends of Learning will begin enrolling teachers in the treatment schools into its professional learning sequence. All teachers in the treatment schools who plan to teach fifth grade science in the coming year will be included. Treatment schools will receive Legends of Learning's professional learning program and those teachers and students will have full access to all of its features and resources. Treatment teachers will be asked to adhere to the implementation plan finalized during the implementation study (see D.4).

The control schools will administer business-as-usual science classes and will not have training with or access to Legends of Learning. These schools will use curriculum materials selected by their district based on a list of state-approved textbooks in conjunction with their own materials. They will have district-offered professional learning in accordance with Florida's Coordinated System of Professional Development (Florida Department of Education, n.d.).

School-level random assignment was selected since schools typically implement Legends of Learning as a school-wide program, in which all science teachers receive PL and access to resources. Second, a teacher-level assignment would raise the threat of contamination, as teachers in a school may discuss, view, and share instructional materials and strategies. Also, based on prior school-level randomized studies (Davenport et al., 2019), particularly with Impact Florida's support, we expect minimal school-level attrition during the impact study.

The analysis of the intervention's impacts will use an intent-to-treat (ITT) approach—schools and their teachers and students will be retained in their originally-assigned groups. Student rosters will be collected at the start of the 2026–2027 school year to identify students in the ITT student impact sample. Given that the proposed evaluation is based on a school-level RCT that is expected to have low cluster-level attrition and a student analytic sample where joiner bias is not a threat, the evaluation has the potential to produce strong evidence about the impact of Legends of Learning. Students who join the school after randomization will not be included in the analytic sample. Based on research (Kim et al., 2020; Taylor & West, 2020) and WestEd's prior RCT experience, we anticipate manageable levels of student attrition during the study (i.e., less than 20%) and minimal differential student attrition across conditions (i.e., less than 5 percentage points), so the student impact analyses will likely meet WWC standards without reservations (see WWC attrition White Paper; WWC, 2017).

D.1.a Generalizability and Scalability. In partnership with Impact Florida, WestEd will evaluate Legends of Learning across a large number of economically disadvantaged schools, particularly in high-minority and some rural settings to reach high needs students. Findings based on the diverse sample will provide valuable guidance for future replications of implementation as the program scales further across Florida and throughout other states. To inform generalizability,

the evaluation will include a set of moderator analyses (RQ6) to assess the extent to which the effects of Legends of Learning are moderated by the characteristics of students, teachers/classrooms, and schools. Results from these analyses will guide future scaling, as they may identify where the program is particularly effective or less suited and how it can be improved accordingly. Appendix J.6 lists the potential moderators and corresponding analysis plan. Furthermore, analyses will explore how implementation, opportunities to learn, and instructional practices mediate the direct effects on students' science achievement.

D.1.b Cost Effectiveness. The evaluation includes a cost analysis based on the Resource Cost Model (Levin & McEwan, 2002) to provide information about the cost of implementing Legends of Learning, including associated PL and support, and whether it is cost effective relative to the business-as-usual condition. Implementation and PL costs will be identified in both the Legends of Learning and business-as-usual conditions using the "ingredients method" (Levin et al., 2017). Analyses will identify the costs associated with each component of the program, distinguish start-up costs from ongoing costs, and convert total costs to per-student costs. We will then combine the cost information and effect size estimates to describe the impact of Legends of Learning on a per dollar basis following the most up-to-date recommendations for cost analyses (Cost Analysis Standards Project, 2021).

D.2 Strategies for Replication

Our approach aims to make findings generalizable so they may impact the most students as the program scales (see D.1.a). To this end, our research questions explore how, when, and for whom Legends of Learning is working. Legends of Learning has been successfully adopted in at least one district in every state except Alaska; however, evaluations of a systematic prescribed

administration model, including moderation and mediation analyses, have not yet been completed. This project will fill this important knowledge gap.

Additionally, with the measures and methods developed by this study, replication and ongoing research can occur to ensure that the program is working for all students, and deficiencies can be identified and corresponding improvements made. Legends of Learning's approach to entry using district-wide or larger partnerships also makes the program scale quickly, which can provide future opportunities for such studies. With the findings from this study and future studies, Legends of Learning will be able to identify areas for cost-effective improvements and will be better able to address any barriers to implementation found across various contexts. Legends of Learning will prioritize continued research and improvements using its own funds and potentially through other partners and grant monies.

D.3 Components, Mediators, Outcomes, and Acceptable Thresholds of Implementation

The design of the proposed evaluation is informed by clearly articulated key components, mediators, and outcomes of Legends of Learning as depicted in the conceptual framework presented in Appendix G. The impact analyses (RQ 1–3) will be based on valid and reliable measures: (1) 5th grade Florida Statewide Science Assessment, (2) Opportunities to Learn, and (3) Instructional Activities scales. The evaluation will include moderator analyses (RQ6) and mediator analyses (RQ 7–8) to explore the relationships among implementation context, intermediate outcomes, and student achievement outcomes, as discussed above. Implementation context data will be collected from multiple sources, including artifacts (e.g., sign-in sheets and agendas) from teacher trainings to determine participation and coverage, observations, interviews, and monthly teacher logs describing implementation of Legends of Learning and other curricular materials, as well as student data from Legends of Learning. Legends of

Learning tracks and will provide to WestEd data on the number and type of student activities completed (e.g., instructional game, quiz game, assessment), including time spent and the learning outcomes targeted by each activity. Acceptable thresholds of implementation will be accounted for within PL measures, teacher logs, and observations. Classroom administration thresholds will most likely be based on tracking the percentage of weeks in which teachers complete the number of lessons / games prescribed in the administration model (e.g., two Legends of Learning lessons per week). These thresholds will be defined during the formative evaluation stages (see D.4), and will be monitored to discover if the program is implemented with fidelity and what potential effect the level of fidelity may have on outcomes.

D4. Formative Evaluation, Performance Feedback, and Periodic Assessment of Progress

During the first two and a half years, performance feedback and periodic assessment of progress will be addressed through formative evaluation. Usability studies, a classroom feasibility study, and a classroom implementation study will build towards the impact study (see Figure 2 and Table 6), and will be guided by the research questions in Appendix J.8.

Table 6. Summary of the samples and timeline for each major evaluation component.

	Usability	Feasibility	Implementation	Impact
Timeline	Fall 2024	Spring 2025	2025–2026	2026–2027
Teachers	10	10	20	240
Students	50	250	500	6,000

Usability Study. In Fall 2024, WestEd will conduct multiple rounds of usability research, as needed, to iteratively test new program components and features. For each round, at least 5 teachers, each with 5 students, will participate. Research has shown that this is a sufficient number of testers to identify major issues, and that testing with additional users provides

diminishing returns (Nielsen, 2000). The student sample of these rounds will be intentionally balanced with regard to demographics. WestEd researchers will guide participants through the relevant tasks and ask them to "think aloud" and explain their thought processes as they go. At the end of each session, WestEd will interview participants about their overall experiences, the ease of use of the activities, and their understanding of the tools and content presented. Each round of usability will result in a qualitative summary of findings **with actionable recommendations to Legends of Learning.**

Feasibility Study. In Spring 2025, WestEd will conduct a feasibility study to evaluate whether teachers and students can use the lessons, components, and features as intended in an authentic education setting with only small-scale guidance and support. Another main goal will be exploring how teachers envision ideal usage of Legends of Learning to inform possible administration models for the implementation study. WestEd will complete the feasibility study with 10 teachers, each with one full class of students. Teachers will receive a one-hour webinar introduction to the Legends of Learning and two hours of PL on how to integrate it into their classrooms. Teachers and students will use Legends of Learning in their classrooms for six weeks. Through interviews and observations, WestEd will capture data related to various implementation models used by teachers and which may be the potentially most effective models. Feedback will be collected from teachers through a weekly log and a final interview. Qualitative and quantitative reports of findings will provide **actionable recommendations to Legends of Learning** throughout the year.

Implementation Study. In the 2025–2026 school year, WestEd will conduct a yearlong implementation study on Legends of Learning, to test the most promising 3–4 implementation models uncovered in the feasibility study. The study will involve 20 teachers, each with one full

class of students. Teachers will receive the same PL during this study as during the feasibility study and will implement Legends of Learning using one of the models. Administration strategies will be intentionally assigned so each includes classrooms with a variety of demographics. WestEd will perform three observations per classroom. Students will complete a post-survey, including the opportunities to learn measure and overall engagement and satisfaction items. Feedback from teachers will be collected through monthly logs and a final interview. The implementation study will result in mixed qualitative and quantitative findings, **with actionable recommendations to Legends of Learning.** WestEd will review the fidelity and results from each administration model to inform the model used for the impact study.

Performance Feedback and Periodic Assessment of Progress. During this formative period, WestEd will help develop the ideal administration model in partnership with Legends of Learning and Impact Florida, who has been heavily involved in scaling effective teaching practices through design and implementation of professional learning models and instructional materials. WestEd will also create measurable thresholds for acceptable implementation in consideration of teacher feedback and quantitative findings from the implementation study.

The formative evaluation will also provide structure for consistent, periodic feedback to Impact Florida on the progress of Legends of Learning towards its development goals. All parties will participate in virtual, bi-monthly calls to discuss upcoming goals and review progress to date on development, testing, and other project components and steps, as detailed in the Detailed Project Timeline and Management Plan (see J.4). WestEd will be responsible for monitoring and tracking all activities and reporting to Impact Florida, to make sure all goals stay within their scope and timeline. Any deviation or concern on progress or performance will be discussed among all stakeholders, including the Program Officer.

References

Allensworth, E.M. & Clark, K. (2020). High School GPAs and ACT Scores as Predictors of College Completion: Examining Assumptions About Consistency Across High Schools. *Educational Researcher*, 49 (3), 198-211.

Barab, S. A., Sadler, T., Heiselt, C., Hickey, D., & Zuiker, S. (2007). *Relating narrative, inquiry, and inscriptions: A framework for socio-scientific inquiry.* Journal of Science Education and Technology, 16(1), 59–82.

Barab, S. A., Scott, B., Siyahhan, S., Goldstone, R., Ingram-Goble, A., Zuiker, S., & Warrant, S. (2009). *Transformational play as a curricular scaffold: Using videogames to support science education*. Journal of Science Education and Technology, 18, 305–320.

Bathgate, M., & Schunn, C. (2016). Disentangling intensity from breadth of science interest: What predicts learning behaviors. *Instructional Science, 44*, 423-440.

Bishop, M. J. (2016). *The Impact of Ongoing Professional Development on Math Achievement.* (Doctoral dissertation, Carson-Newman University, Jefferson City, Tennessee). Retrieved on 08/17/2020 from https://www.cn.edu/libraries/tiny_mce/tiny_mce/plugins/filemanager/files/Dissertations/M_Bishop_Dissertation-1.pdf

Bryan, R. R., Glynn, S. M., & Kittleson, J. M. (2011). Motivation, achievement, and advanced placement intent of high school students learning science. *Science Education, 95*(6), 1049–1065. doi:10.1002/sce.20462.

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis. *Review of Educational Research*, 86(1), 79–122. https://doi.org/10.3102/0034654315582065

Clark, D. B., Tanner-Smith, E., Hostetler, A., Fradkin, A. & Polikov, V. (2018). Substantial Integration of Typical Educational Games Into Extended Curricula, *Journal of the Learning Sciences*, 27:2, 265-318.

Collaborate, Plan, Align, Learn, Motivate and Share. [CPALMS]. (n.d.). *Search Standards*. CPALMS. https://www.cpalms.org/public/search/Standard

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser*, 453–494. Lawrence Erlbaum Associates, Inc.

Cost Analysis Standards Project. (2021). Standards for the Economic Evaluation of Educational and Social Programs. Retrieved from: https://www.air.org/sites/default/files/Standards-for-the-Economic-Evaluation-of-Educational-and-Social-Programs-CASP-May-2021.pdf

Darling-Hammond, L., Hyler, M. E., Gardner, M. (2017). *Effective Teacher Professional Development.* Palo Alto, CA: Learning Policy Institute.

Davenport, J.L., Kao, Y. S., Matlen, B. J., & Schneider, S. A. (2019). Cognition research in practice: engineering and evaluating a middle school math curriculum. *The Journal of Experimental Education*, 88 (4), 1–20.

Dieterle, E. (2009). *Neomillennial learning styles and River City*. Children, Youth and Environments, 19(1), 245–278.

EdTech Top 40: Fall 2022 Report. (2022). Instructure. https://www.instructure.com/resources/research-reports/edtech-top-40-fall-2022-report?filled

Edwards, A. R. & Beattie, R. L. (2016). Promoting student learning and productive persistence in developmental mathematics: Research frameworks informing the Carnegie Pathways. *NADE Digest*, 9 (1), 30–40.

Florida Department of Education. (n.d.). *Professional Learning Catalogs (PLC)*. https://www.fldoe.org/teaching/professional-dev/master-inservice-plans-mip.stml

Florida Department of Education. (2010). *Standard Setting*. https://www.fldoe.org/accountability/assessments/k-12-student-assessment/stard-setting.stml

Florida Department of Education. (2018). *Florida standards assessments*. https://www.fldoe.org/accountability/assessments/k-12-student-assessment/results/2018.stml

Florida Department of Education. (2021). *Florida standards assessments*. https://www.fldoe.org/accountability/assessments/k-12-student-assessment/results/2021.stml

Florida Department of Education. (2022). *2022 Statewide Science Assessment Next Generation Sunshine State Standards State Report of District Results Grade 05*. https://www.fldoe.org/accountability/assessments/k-12-student-assessment/results/2022.stml

Florida Department of Education. (2023). PK-12 Enrollment Demographics [Dataset]. In *Know Your Data FL*. https://knowyourdatafl.org/views/PK12-Enrollment/ENROLLMENTDEMOGRAPHICS?%3AshowAppBanner=false&%3Adisplay_count=n&%3AshowVizHome=&%3AisGuestRedirectFromVizportal=y&%3Aembed=y

Friday Institute for Educational Innovation (2012). Upper Elementary School STEMStudent Survey. Raleigh, NC: Author.

Gee, J. P. (2007). *What video games have to teach us about learning and literacy* (2nd ed.). New York, NY: Palgrave Macmillan.

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood to late adolescence: A longitudinal study. *Journal of Educational Psychology*, *82*, 525–538.

Grabau L.J., & Ma, X. (2017). Science engagement and science achievement in the context of science instruction: a multilevel analysis of U.S. students and schools. *International Journal of Science Education, 39*(8), 1045-1068

Hart, H., Young, C., Chen, A., Zou, A., & Allensworth, E.M. (2020). *Supporting school improvement: Early findings from reexamination of the 5Essentials survey.* Chicago, IL: University of Chicago Consortium on School Research.

Hart, H., Young, C., Chen, A., Kheraj, N., & Allensworth, E.M. (2021). *5Essentials Survey in CPS: School improvement and school climate in high poverty schools.* Chicago, IL: University of Chicago Consortium on School Research.

Hawkey, R., & Clay, J. (1998). Expectations of secondary science: Realisation and retrospect. *School Science Review, 79*(289), 81–83.

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.

Kafai, Y. B., Quintero, M., & Feldon, D. (2010). *Investigating the "why" in Whypox: Casual and systematic explorations of a virtual epidemic.* Games and Culture, 5, 116–135.

Kaimara, P., Fokides, E., Oikonomou, A. et al. (2021). Potential Barriers to the Implementation of Digital Game-Based Learning in the Classroom: Pre-service Teachers' Views. *Tech Know Learn* 26, 825–844. https://doi.org/10.1007/s10758-021-09512-7

Ketelhut, D. J. (2007). *The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment.* Journal of Science Education and Technology, 16(1), 99–111.

Kim, D., Lee, Y., Leite, W. L., & Huggins-Manley, A. C. (2020). Exploring student and teacher usage patterns associated with student attrition in an open educational resource-supported online learning platform. *Computers & Education*, 156 , 103961.

Klopfer, E., Scheintaub, H., Huang, W., Wendal, D., & Roque, R. (2009). *The simulation cycle: Combining games, simulations, engineering and science using StarLogo TNG*. E-Learning, 6(1), 71–96.

Kozulin, A. (2000) Diversity of Instrumental Enrichment applications. In A. Kozulin and Y. Rand (Eds.), *Experience of mediated learning*. Oxford: Pergamon.

Le, V. N., Stecher, B. M., Lockwood, J. R.,Hamilton, L. S., & Robyn, A. (2006). *Improving mathematics and science education: A longitudinal investigation of the relationship between reform-oriented instruction and student achievement*. Santa Monica, CA: RAND Corporation. Retrieved from https://www.rand.org/pubs/monographs/MG480.html

Lee, W.S., Shih, M., Liang, Jl, & Tseng, Y. (2021). Investigating learners' engagement and science learning outcomes in different designs of participatory simulated games. *British Journal of Educational Technology, 52*(3), 1197-1214.

Levin, H. M., & McEwan, P. J. (2002). Cost-effectiveness and educational policy. Larchmont, NJ: Eye on Education.

Levin, H., McEwan, P., Belfield, C., Bowden, A.B., & Shand, R. (2017). *Economic Evaluation in Education: Cost-Effectiveness and Benefit-Cost Analysis*. Sage: Thousand Oaks, CA.

Martinez-Garza, M., Clark, D. B., & Nelson, B. (2013). *Digital games and the US National Research Council's science proficiency goals.* Studies in Science Education, 49, 170–208.

McKinney, D., Strother, S., & Schneider, S. (2023). *The Association of Legends of Learning Usage and Science Achievement in Four School Districts in the Southern United States.* WestEd, San Francisco CA.

National Association of Educational Procurement [NAEP]. (2022). Florida Demographics (2021-2022) [Dataset]. In *The Nation's Report Card*. https://www.nationsreportcard.gov/profiles/stateprofile/overview/FL?cti=PgTab_Demographics&chort=1&sub=MAT&sj=FL&fs=Grade&st=MN&year=2022R3&sg=Gender%3A%20Male%20vs.%20Female&sgv=Difference&ts=Single%20Year&tss=2022R3&sfj=NP

National Center for Education Statistics. (n.d.-a). *Current expenditures and current expenditures per pupil in public elementary and secondary schools: 1989-90 through 2030-31*. https://nces.ed.gov/programs/digest/d21/tables/dt21_236.15.asp

National Center for Education Statistics. (n.d.-b). *Fast Facts: Long-term trends in reading and mathematics achievement (38)*. https://nces.ed.gov/fastfacts/display.asp?id=38

National Center for Education Statistics. (n.d.-c). *Total and current expenditures per pupil in public elementary and secondary schools: Selected years, 1919-20 through 2006-07*. https://nces.ed.gov/programs/digest/d09/tables/dt09_182.asp

National Research Council. (2011). *Learning science through computer games and simulations.* Washington, DC: National Academies Press.

Nielsen, J. (2000). Jakob Nielsen's Alertbox, March 19, 2000: Why You Only Need to Test With 5 Users.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25*(9), 1049–1079.

Papadakis, S. (2018). The use of computer games in classroom environment. *International Journal of Teaching and Case Studies*, *9*(1), 1–25.

Polikoff, Morgan with Jennifer Dean. *The Supplemental-Curriculum Bazaar: Is What's Online Any Good?* Washington, DC: Thomas B. Fordham Institute (December 2019). https://fordhaminstitute.org/national/research/supplemental-curriculum-bazaar

Qian, M., & Clark, K. R. (2016). Game-based learning and 21st century skills: A review of recent research. *Computers in Human Behavior, 63*, 50-58.

Reeves, T., Romine, W., Laffey, J., Sadler, T., & Goggins, S. (2020). *Distance Learning through Game-Based 3D Virtual Learning Environments: Mission Hydro Science. Evaluation Report for Mission HydroSci.* https://eric.ed.gov/?id=ED605283

Rickles, J., Zeiser, K. L., Yang, R., O'Day, J., & Garet, M. S. (2019). Promoting Deeper Learning in High School: Evidence of Opportunities and Outcomes. *Educational Evaluation and Policy Analysis*, 41(2), 214–234. https://doi.org/10.3102/0162373719837949

Robb, M. (2022, March 23). Kids' Media Use Accelerated Rapidly During the Pandemic. *Common Sense Media*. https://www.commonsensemedia.org/kids-action/articles/kids-media-use-accelerated-rapidly-during-the-pandemic

Roblox. (2023). *Q1 2023 Supplemental Materials*. Report: San Mateo, CA.

Rogoff, B. (1990). Apprenticeship in thinking: Cognitive development in social context. Oxford University Press.

Rogoff, B. (1998). Cognition as a collaborative process. In W. Damon (Ed.), Handbook of child psychology: Vol. 2. Cognition, perception, and language, 679–744. John Wiley & Sons Inc.

Rowan, B., & Correnti, R. (2009). Studying reading instruction with teacher logs: Lessons from the study of instructional improvement. *Educational Researcher*, 38 (2), 120-131.

Simpson, R. D., & Oliver, J. S. (1990). A summary of major influences on attitude toward and achievement in science among adolescent students. *Science Education, 74*(1), 1–18.

Squire, K. D., & Jan, M. (2007). Mad City Mystery: Developing scientific argumentation skills with a place-based augmented reality game on handheld computers. *Journal of Science Education and Technology*, 16(1), 5–29.

Squire, K. D., & Klopfer, E. (2007). Augmented reality simulations on handheld computers. *Journal of the Learning Sciences*, 16(3), 371–413.

Squires, D. (2011). Curriculum Alignment Research Suggests That Alignment Can Improve Student Achievement. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 85*, 129-135.

Steinkuehler, D., & Duncan, S. (2008). Scientific habits of mind in virtual worlds. *Journal of Science Education and Technology*, 17, 530–543.

Taylor, J. A., & West, B. (2020). Estimating Teacher Attrition for Impact Study Design. *Educational Researcher*, 49 (1), 68–70. https://doi.org/10.3102/0013189X19880550

Tobias, S., & Fletcher, J. D. (Eds.). (2011). *Computer games and instruction.* Charlotte, NC: Information Age.

Turkay, S., Hoffman, D. J., Kinzer, C. K., Chantes, P., & Vicari, C. (2014). Toward Understanding the Potential of Games for Learning: Learning Theory, Game Design Characteristics, and Situating Video Games in Classrooms. *Computers in the Schools*, *31*(1–2), 2–22. https://doi.org/10.1080/07380569.2014.890879

Unfried, A., Faber, M., Stanhope, D. S., & Wiebe, E. (2015). The Development and Validation of a Measure of Student Attitudes Toward Science, Technology, Engineering, and Math (S-STEM). *Journal of Psychoeducational Assessment*, 33(7), 622–639. https://doi.org/10.1177/0734282915571160.

US Department of Education (2012). *Race to the Top District Competition Draft: Definitions*, retrieved from https://www.ed.gov/race-top/district-competition/definitions

Walters, K., Smith, T.S., Lennon, V., Ogut, B. & Griffin, M. (2018). The better math teaching network: Lessons learned from the second year. Quincy, MA: Nellie Mae Education Foundation.

Wang, M. T., & Fredricks, J. A. (2014). The reciprocal links between school engagement and youth problem behaviors. *Child Development Perspectives, 8*(4), 225-230.

What Works Clearinghouse. (2017). *Assessing attrition bias white paper*. Washington, DC: Author. https://ies.ed.gov/ncee/wwc/Document/243.

What Works Clearinghouse (2020). *WWC Procedures and Standards Handbook Version 4.1*. Washington, DC: What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education.

Whitworth, B.A., Chiu, J.L. (2015). Professional Development and Teacher Change: The Missing *Leadership Link. J Sci Teacher Educ 26*, 121–137. https://doi.org/10.1007/s10972-014-9411-2

Vu, P., & Feinstein, S. (2017). An exploratory multiple case study about using game-based learning in STEM classrooms. *International Journal of Research in Education and Science*, 3(2), 582-588.

Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B.,…Yukhymenko, M. (2012). Our princess is in another castle: A review of trends in serious gaming for education. *Review of Educational Research*, 82(1), 61–89.