Transforming the Learning of Science for Second Grade Latinx Students Through Meaningful Interactions using Technology Outside of School (Project MITOS)

Table of Contents

A. Significance

The purpose of this Mid-phase project is to adapt and scale the use of a gamified app (MITOS) designed to be used both in school and outside of school to increase second grade student opportunities to develop science vocabulary and content knowledge. The app will provide culturally, and linguistically appropriate activities to expand student science vocabulary and content knowledge particularly for English learners (ELs), who are our targeted high needs population of students.

Science is one of the key skills required in the labor force, and improving and expanding science learning and engagement to increase student achievement in science is consistent with the efforts to ensure the Nation's economic competitiveness (Department of Education, 2023, EIR Mid-phase grant application, p. 9). Moreover, to obtain a deep understanding of science requires development of science vocabulary from the earliest stages of learning (Baker, et al., 2018; S. K. Baker et al., 2014).

MITOS involves the following: (1) teachers assign students lessons from the MITOS app to be completed outside of school based on the science content they will be teaching; (2) students work on the MITOS lessons for approximately 40 minutes per week (i.e., 10 minutes per day) guided by a friendly character or avatar who prompts them to answer questions out loud as part of the app activities; (3) these interactions are recorded by the system and the audio is transcribed and scored by an automated scoring system; (4) parents use a structured process to discuss the lessons students are learning at home for approximately 30 minutes per week (i.e., 10 minutes per day for three days); (5) teachers view student scores and student progress on the dashboard and make decisions on what science words or concepts should be revisited or discussed in small or large group discussions. If students are not able to engage adequately with the app activities,

teachers can reteach the concepts in small groups using the MITOS+ small group lessons.

MITOS proposes an innovative scalable solution that can overcome the persistent challenges to academic success for underserved and high needs students by (1) providing students increased opportunities to learn science outside of school through an interactive app that can be used flexibly on any device and setting; (2) developing a dashboard that will generate automated, easy-to-digest reporting for teachers using natural language processing (Nadkarmi et al., 2011), so they clearly understand which students are able to engage with the lessons on the app and which ones need supplementary support; (3) translating lessons into Spanish to address the needs of Latinx students with low English language proficiency; and (4) creating opportunities for families to engage in the science activities with their children. All materials used in MITOS to support student learning of science vocabulary and content are aligned with evidence-based practices proposed by the What Works Clearinghouse (WWC) Practice Guide for ELs (S. K. Baker et al., 2014). MITOS also aligns with the Framework for K-12 Science Education: Practices, Crosscutting Concepts and Core Ideas (National Academy of Sciences, Engineering, and Medicine, NASEM, 2021). MITOS addresses **three priorities** as defined in the EIR Mid-phase grant application: **Absolute Priority 1 and 3, and the Competitive Preference Priority.**

Absolute Priority 1- *Project supported by evidence that meets the conditions of moderate evidence of a key component in improving relevant outcomes for a similar population described in the cited studies.* MITOS is supported by **strong evidence based on the WWC review, Version 2.1** of two recommendations from the What Works Clearinghouse (WWC) Practice Guide: *Teaching academic content and literacy to English learners in elementary and middle school* (S. K. Baker et al., 2014). **Recommendation 1** indicates that ***teaching a set of***

academic vocabulary words intensively across several days *using a variety of instructional activities benefits EL learning* (S. K. Baker et al., 2014; p. 82). **Recommendation 2** indicates that *the* ***integration of oral and written English language instruction into content-area teaching*** *support EL vocabulary and content learning* (S. K. Baker, 2014; p. 88). The development of MITOS followed these two recommendations.

Regarding **Recommendation 1,** Project MITOS provides students opportunities to learn a set of academic words intensively through activities such as introducing these words in a brief and engaging informational text, practicing word learning with different modalities such as listening, speaking, reading, and writing, and teaching strategies to help students figure out the meaning of other words (S. K. Baker et al., 2014, pp. 82).

For **Recommendation 2,** MITOS will provide ELs the opportunity to learn and expand their knowledge of science. Activities in MITOS are anchored in learning words in a science content, using visuals and graphic organizers to illustrate the meaning and the use of the target word, teaching explicitly science words as well as general academic words that can be used in the future, providing ELs many opportunities to talk about the content with parents, a friendly character in the app, and their peers and teachers at school. MITOS also overlaps with the population of the recommendations by targeting ELs in second grade. The WWC Practice Guide indicates that all the activities in studies reviewed by the panel are recommended for ELs in grades K-8 (see the Evidence Form WS01153634-ED_Evidence-2_0-V2.0 provided with this proposal, and Appendix J10 for an example of activities based on Recommendations 1 and 2).

We have also added in the Evidence Form the reference of a study that has shown moderate effects of components of the MITOS intervention (see Baker et al., 2021). The study is publicly available, and included Latinx EL students, a key target population in MITOS.

Absolute Priority 3 – *Promoting equity in student access to educational resources and opportunities in science to improve student achievement and attainment in STEM.* Project MITOS is designed to improve and expand EL opportunities to develop their **science vocabulary and language** through engaging and motivating gamified activities. According to the WWC Practice Guide (S. K. Baker et al., 2014), NASEM (2021) and the National Research Council (NRC, 2013), students need opportunities to do what scientists do: **pose questions,** carry out investigations, analyze data, **draw conclusions based on evidence, communicate results, and debate with peers to develop a deeper understanding of science.** MITOS provides students with opportunities to converse with others and build science understanding through an app that exposes them to diverse types of words and discourse practices that have specific meaning in science concepts. For example, in a science lesson on *Rocks*, students will encounter three types of words: (a) common, everyday words, such as *soft*, *rock*, and *color*; (b) academic words, such as *texture*, *properties*, and *impact*; and (c) content knowledge words, such as *minerals*, *geologist*, and *luster*. For ELs to successfully navigate academic content, the introduction of academic words cannot wait until children have acquired basic word knowledge; vocabulary learning must occur on a regular basis **in both formal and informal contexts** (Baker, Basaraba et al., 2018; McKeown et al., 2017; S. K. Baker et al., 2014).

MITOS includes vocabulary and language in science that is suggested in the Next Generation Science Standards and the STEM Framework (NRC, 2013), the Texas Education Knowledge and Skills in science (Texas Education Agency, TEA, 2017). The MITOS app will activate student prior knowledge, review science concepts, and build connections between what students are learning at school with what they are experiencing at home. Teachers will be able to use results from a vocabulary formative assessment displayed in a dashboard to make data-

driven decisions about their science instruction. We envision MITOS to be vertically aligned with a focus of preparing ELs for college, and for STEM careers.

Competitive Preference Priority - *Promoting equity in student access to educational resources and opportunities*. Project MITOS will be implemented by UT Austin, a minority serving institution that is eligible to receive assistance under title V of the Higher Education Act (HEA, 2013). MITOS promotes educational equity and adequacy in resources and opportunities for underserved second grade ELs by providing them with the guidance to access high quality evidence-based science and academic language activities backed by strong evidence of effects as indicated in the WWC Practice Guide for ELs (S. K. Baker et al., 2014). We will also include in our team graduates and undergraduate students who are first generation college students, are currently ELs or former ELs, or who have lived in poverty to ensure they have opportunities to engage and participate in rigorous research.

In MITOS we are targeting ELs as defined by section 8101(2) of the Elementary and Secondary Education Act (ESEA, 2015). ELs are individuals who speak a language other than English at home and whose language difficulties deny them the ability to meet challenging state academic standards, succeed in classrooms where the language of instruction is English, or participate fully in society. According to the National Assessment of Educational Progress (NAEP), ELs scored 33 points lower than non-ELs in science and in reading (NAEP, 2019; 2022 respectively). Additionally, ELs are more likely to live in poverty and have parents who did not complete high school, characteristics that are strongly associated with low academic performance (Krogstad, 2014; Morgan et al., 2016). Latinx ELs, who represent 77% of the EL population (National Center for Education Statistics, 2023) are a particularly vulnerable population given their low socioeconomic status. Even though ELs are underrepresented and underperforming in

STEM K–12 subjects, their actual STEM interest is similar to their peers (NASEM, 2018), which supports the development of MITOS designed to address the language needs of all students, and particularly the language needs of ELs.

In this proposal, the terms *parents, family,* and *adults* in the household will be used interchangeably because Latinx children are more likely than White children to live in households with adults other than their parents (e.g., a grandparent, aunt or uncle, teenage siblings; Guzman et al., 2016). We focus on second grade because it represents students' transition from developing foundational reading skills to comprehending increasing amounts of informational text independently without extensive teacher support.

MITOS **is a promising approach to address the vocabulary and English Language Proficiency (ELP) needs of all students and particularly our target population, ELs, compared to standard instruction.** The app provides an alternative to traditional homework or independent practice by (1) allowing students to have engaging and substantive interactions that develop vocabulary and content knowledge, (2) leveraging family engagement, and (3) providing ongoing formative assessment data through natural language processing presented in a dashboard that teachers can use to inform their instruction. In addition, we make it scalable by allowing its use on an app that can be accessed at school or outside of school. This is a promising approach that improves standard practice, which typically does not support sufficient focus on vocabulary instruction in content areas (Baker et al.,2018, 2021), does not engage families in student's regular homework/independent practice (NASEM, 2016), and provides teachers with limited formative assessment information on language and vocabulary development (McKeown, 2017).

MITOS is anchored on two previous grants from the Institute of Education Science led by [redacted] The English Learner **Vocabulary Acquisition** project **(ELVA, Baker, 2014–2017)** and

the **Measuring of English Language Vocabulary Acquisition in Science (MELVA-S, Baker, 2020–2024).** Both include a speech recognition system (SRS) and an automated scoring system (AS system). In ELVA, we developed science and social studies lessons that improved Latinx students' vocabulary knowledge and comprehension with support from a virtual tutor in school settings. Results from two experimental studies indicated that students who used the ELVA app scored significantly higher on a researcher-developed vocabulary measure than students who did not use ELVA, including students who received science instruction in Spanish (*ES* = 0.38 in Study 1, Baker, Ma, et al, 2021; *ES* = 0.77 in Study 2, Baker & Le, under review). Moreover, in Study 2, the quality of the explanations in a multiple-choice assessment in science was higher in the group that used ELVA than in the comparison condition.

Other outcomes also indicated that students enjoyed the program and were able to use it easily. Teachers noticed that students who participated in the program related what they had learned with the ELVA app to what they were learning in class. Teachers valued that the program supported comprehension and language development, and they thought the words were aligned to the topics they were teaching (see Baker, Conry et al., 2018 for a video on ELVA). For Project MELVA-S, we are developing a formative science vocabulary assessment that will provide teachers with information about students' science word knowledge. The MELVA-S work on the SRS and AS system will inform the iterative development of the SRS and AS system in Project MITOS (see Sano et al., 2020; Wu, 2023). We **build on this prior work and expand it by capturing and analyzing student speech in the home in English, Spanish, and a mixture of these languages.**

B. Strategy to Scale Up MITOS

B1. Strategies That Overcome Barriers to Sustaining and Scaling

The MITOS scaling plan leverages strategies that take into account key barriers for teachers and students to use the evidence-based curriculum, including barriers particular to Latinx ELs. In a previous iteration of MITOS (Baker et al., 2021), we identified three primary barriers: (1) restrictions in where and how MITOS could be used, (2) teachers wanting guidance for how to effectively use their limited time to target extra support for students in need, and (3) limitations in how the program accommodates students with low English proficiency. We will overcome these barriers to scaling MITOS through the following strategies.

B1.1 Flexible Use of MITOS Across Devices and Settings. In the initial development of the MITOS app, we focused on the use of PCs, which tied implementation to core classroom time. We will take the existing evidence-based curriculum and develop a gamified app that can be used on any device and will be appropriate for use in multiple settings (i.e., in the regular school day, during after-school programs, and at home). We will create a website that includes the app for teachers and families to download onto any device available, including interoperability with Android and IOS (e.g., Apple) systems. Teachers also will be able to access the data dashboard and the support materials such as implementation guide and supplemental lessons taught in small groups by the teacher. Importantly, the MITOS app will be accessible to download via smartphone, the most common computing device owned in the country (US Census Bureau, 2021) This will allow students to engage in MITOS lessons as part of core instruction during the school day, but also allow flexibility for students to engage in MITOS in other settings, including at home as part of homework.

B1.2 Support for Teachers to Direct Limited Time to Students in Need. While in our prior implementation we found some students were able to use MITOS independently, others required more tailored support from their teacher, but teachers struggled with identifying the

students in need of further support and finding effective ways to follow-up with them given the limited time available for science instruction. We will support scaling by developing a teacher dashboard that will generate automated, easy-to-digest reporting for teachers, so they clearly understand which students are able to engage with the lessons on the app and which are not. In addition, the app will use our proven speech recognition system (SRS) and automated scoring (AS) system to score student responses as they engage in the app activities in English and in Spanish, reducing the time teachers would normally use to score student vocabulary formative assessment responses. As teachers use the dashboard to identify students who need extra support, they will have a set of supplementary lessons available to deliver to students in small groups twice a week. These lessons have already been created and are based on the Next Generation Science Standards, the STEM Framework (NRC, 2013), and the Texas Educational Knowledge and Skills standards (TEA, 2017). The lessons can be easily embedded in typical instructional routines, such as station rotation. Teachers will receive a one-time, start-up PD on the materials in the app, how to use the app, and how to use the supplemental materials for students in need of further support. To sustain and scale up the project, we will record the PD that can then be available on demand.

B1.3 Strategic Leveraging of Spanish Language Through Curriculum Translation, Formative Assessment, and Parental Involvement. In our previous study, some students with more nascent English proficiency struggled answering questions posed by an avatar even though they were equally engaged with the content as other ELs with higher language proficiency. Given that the majority of ELs in the country speak Spanish at home, we will translate the curriculum into Spanish and create the capability for the students to select to conduct the activities in Spanish when using the app. This will allow Latinx ELs to engage more fully in the

learning opportunity. In addition, the automated scoring to provide formative assessment information for teachers will be able to draw on students' language production in both English and Spanish. This will allow the teacher to have richer information about their Spanish-speaking students, and better distinguish between bilingual student understanding of science concepts and bilingual student vocabulary development in either of their languages.

In addition, we will draw on research underscoring the importance of family engagement for supporting ELs' academic learning (NASEM, 2016), and create opportunities for students to use the app at home with a family member. Specifically, students will engage in activities three times per week that ask them to have a 5–10-minute conversation with a family member about a science topic that they are learning about. Parents will follow specific question prompts to ask their child and cues to listen for or probe further about. Parents will be offered a one-time workshop on how to engage with MITOS with their children at home, which will be offered in both English and Spanish. We will also ask teachers to plan a monthly event for parents to maintain their engagement and motivation to talk science with their children at home. Having a structured opportunity for students and their family members to discuss science topics, and allow students to use their science vocabulary, will provide an important complement to the instruction they receive in the classroom. For students who can benefit from conversation and support in Spanish, engaging their families who can speak to them in Spanish will be of particular importance.

B2. Organizational Capacity to Bring the Project to Scale

Project MITOS is submitted through the University of Texas Foundation, and it entails the collaboration between Texas school districts, the University of Texas at Austin (UT Austin), Southern Methodist University (SMU), and American Institutes for Research (AIR). Within UT

Austin, the research team will be supported by the Meadows Center for Preventing Educational Risk (MCPER), and the Texas Center for Equity Promotion (TEXCEP) in the College of Education. Both Centers have a strong focus on reimagining education to ensure that high needs students who are culturally and linguistically diverse have all the opportunities and supports to raise the bar (EIR, Mid-phase grant application, 2023) and achieve academic success. MCPER is a world leader in identifying innovative school-related practices, initiating, and managing active programs of research to validate promising practices and disseminating user-specific information on evidence-based practices. With more than 150 employees and more than 15 currently funded projects, MCPER has ongoing partnerships with numerous school districts across Texas and the nation and has strong relationships with TEA.

B2.1 Team Approach. Our research team has worked together on two previous IES projects focused on recognizing and scoring student language samples, and team members have expertise in each of the major areas of the project, as indicated in the section below. We will also recruit family liaison specialists who will support the use of the app in the home. AIR who will conduct the external evaluation has many years of experience evaluating EIR grants. GRAs with an interest in technology, vocabulary and language development, and family engagement will constitute the next generation of scholars striving to overcome barriers to the development of home-school learning partnerships, technology accessibility, and EL opportunities to learn.

B2.2 Qualified Personnel. [redacted] **Ph.D., PI.** [redacted] will oversee all development conducted by the different project teams (i.e., the technology team, the family liaison team, the content development team, the data analysis team, and consultants). [redacted] is an associate professor at UT Austin, a board member at MCPER, and an affiliate of TEXCEP. Her research focuses on developing and testing interventions and formative assessments

designed to improve the academic outcomes of all students and particularly Latinx ELs. Her interventions include state of the art technology such as speech recognition systems and automated scoring as well as parental involvement. [redacted] has been the PI or co-PI on multiple Institute of Education Sciences grants among others to develop and test in large efficacy trials interventions addressing the vocabulary and language proficiency needs of ELs in their native language and in English. [redacted] is Latinx, and a native Spanish speaker. Three of her intervention studies that used an experimental design (2 as a PI, and one as a Co-Pi) have met the WWC evidence standards without reservations.

[redacted] **Ph.D., CO-PI, and supervisor of the SRS**. [redacted] is an Associate Professor at SMU. He will work closely with an engineering graduate student on the adaptation of the SRS and the user interface. [redacted] is an expert in machine learning, especially on applications involving audio and images. He has published works on analyzing speech and audio for applications in education, health, and security. He also has made image databases with subjective, psycho-visual scoring available to the research community, and the dataset has been downloaded more than 100,000 times and used in more than 1,500 research papers. [redacted] **Ph.D., CO-PI; Family Engagement Specialist**. [redacted] is a Research Associate at MCPER. She will help supervise family liaisons and facilitate feedback from the families during the adaptation of the app and during the pilot study and experimental studies. [redacted] Latinx heritage, native Spanish skills, and extensive work with family-school connections in large scale grants will support the use of MITOS at home and in the classroom. [redacted] **Ph.D., CO-PI, is a** professor of instruction at UT at Austin. **He will be the Supervisor of App Development**. [redacted] will oversee the adaptations of the MITOS app. He is the director of the Game Development and Design Program, and PI of the Simulation and

Game Applications (SAGA) Lab at UT Austin. [redacted] has produced more than 200 computer games and apps, ranging from those designed to advance academic research to those created for entertainment. [redacted] **Ph.D., evaluation PI.** [redacted] is a principal researcher at AIR with extensive expertise in evaluating teaching and learning programs for ELs. She serves as the evaluation PI for two current EIR grants and is the PI of a study on the efficacy of an EL-focused teacher PD program. She will be responsible for overseeing the evaluation and will direct all components including study design, random assignment, data collection, impact analysis, reporting, and dissemination. She has presented her work at national research conferences; convenings for practitioners and state and district administrators; and through webinars targeting diverse audiences. Her work has been published in peer-reviewed journals, including *Review of Research in Education* and *Educational Evaluation and Policy Analysis*. The leadership team will be supported by **a strong team of additional personnel and consultants** with well-defined roles, described in Appendix J9.

B3. Mechanisms for Dissemination to Support Further Development or Replication

B3.1 Dissemination. We will leverage the existing dissemination network and tools at MCPER, TEXCEP, UT Austin, SMU, and AIR. First, we will create a specific **project website or a landing page** on the MCPER website, with links to collaborating institutions' websites. The Project MITOS site will house papers and presentations that derive from Project MITOS. Second, we will utilize our **social media channels (Facebook, Twitter, Vimeo, YouTube) and email listserv** to share these resources with thousands of educators. Finally, we will utilize more **traditional methods of dissemination**, such as publishing papers in peer-reviewed journals and attending professional conferences such as the **American Education Research Association, the National Association for Research in Science Teaching, and the Society for Research on**

Educational Effectiveness, as well as practitioner conferences such as the **National Science Teaching Association**, the **Conference for the Advancement of Science Teaching**, and the **National Association of Bilingual Education.** The Project MITOS authors and graduate research assistants will be instrumental in developing, packaging, and distributing social media posts, information for educators, and other outreach materials.

For parents, particularly Latinx parents, we will offer workshops on how to enhance their academic supports for their children at home. We will offer these workshops through schools and national parent engagement support organizations such as **Parent Teacher Associations, the National Association for Family, School, and Community Engagement, Avance, Bachman Lake Together, and United Way.** In addition, we will produce a parent-friendly report of the findings that will be disseminated to parents through existing school communications such as newsletters. [redacted] and [redacted] are native Spanish speakers and will be able to share findings with parents in English and Spanish. They will also create a video to share through MCPER social media accounts.

B3.2 Supporting Further Development or Replication. If MITOS is successful, we will support replication by the project by (a) making databases available to the research community and (b) disseminating packages for processing data on mobile platforms. [redacted] will share the MITOS App process for aligning and integrating game-based learning. To further distribute the MITOS project we will create license agreements with districts, schools, and classrooms. A license would gain classroom access to all intervention content for six months. For example, a classroom would pay a small licensing fee to help support management of the Progressive Web App. All materials and content will be housed on the website. The classroom teacher would be able to use the data from the dashboard on the website to plan and deliver

instruction. The website would provide a step-by-step guide to intervention implementation. Thus, a key objective in our adaptation work will be to make sure the MITOS app and the website are user friendly.

B4. Utility of Products

The MITOS app uses NGSS and TEKS aligned content that is relevant for science learning in second grade in any setting nationally. MITOS also draws on strong evidence for how to effectively teach vocabulary and build science understanding for ELs. In addition, MITOS will be flexible in accessibility and use across devices and operating systems, ensuring its usability across settings. The MITOS app holds potential to be scaled further to other grades and content areas if the evaluation results indicate positive impacts for students. We envision using the same approach to develop an app across grades and in other content areas such as mathematics and social studies. Finally, national and state literacy and science standards make clear the expectation that all students beginning in the early primary grades should engage in expository text and in developing their argumentation skills in addition to developing their word reading skills. MITOS will provide all students the opportunities to engage in expository text and in argumentation.

C. Quality of the Project Design

C1. Conceptual Framework

Project MITOS is anchored on the recommendations provided in the WWC Practice Guide (S. K. Baker et al., 2014) on how to increase student vocabulary knowledge given the strong evidence from studies reviewed by experts. To carry out these recommendations, the research team in MITOS incorporated three key factors that support the learning of new

vocabulary in science for Els. First, project MITOS will engage parents in talking with their children about science. Second, in MITOS we use state-of-the art technology to support student learning and teacher enhancement of their instruction. Third, we acknowledge that student bilingualism and biculturalism are assets that support and build student science vocabulary and language within an inquiry-based approach.

C1.1 Parental Engagement. Research indicates that engagement of families is associated with positive student outcomes, such as higher grades and higher ELP (NASEM, 2017). [redacted] work with families also suggests that Latinx parental knowledge is key for parents to support their children academically, and that parents are eager to participate in academic activities with their children (Baker et al., 2016; 2019; York et al., 2019). However, Baker et al. (2019) also found that few schools can provide tools for parents to use at home or show parents how to support their children's academic achievement, especially in ways that take their sociocultural assets into account.

The MITOS app will provide opportunities for parents to engage in academic discourse together with their children, increasing their knowledge of **what** their children need to learn and **how** best to enhance their learning. These school-parent partnerships are foundational for **(a) increasing children's opportunities to learn, (b) establishing a proactive, rather than reactive, approach to parental engagement in their children's education, and (c) developing trusting relationships between educators and families** (NASEM, 2016; Olmstead, 2013).

C1.2 Learning Through Technology. Project MITOS incorporates Mayer's (2009) theories of learning using multimedia. The integration of auditory and visual modalities enriches student mental representation of the material and increases their active engagement in the critical thinking needed to integrate new information with prior knowledge (Kintsch, 1988). Compelling

evidence indicates that engaging students in meaningful interactions improves learning, and it increases their opportunities to reflect on what they see or hear (Chapin et al., 2003; Chiu, 2013; Murphy et al., 2009; Soter et al., 2008).

C1.3 Bilingualism and Learning the Language of Science. A premise of this project is the positive, causal relationship between students' first and second language as concepts developed through a student's primary language transfer to other languages (Baker, Cummings et al., 2022; Baker, McCoach et al., 2021; Baker, Park, et al., 2022; Cummins, 2005). **Project MITOS takes an additive approach to bilingualism, capitalizing on language-development processes that support parents and children in better understanding how home-language practices are inherently valuable, and provide a foundation for building content knowledge and academic language.**

C2. Clearly Specified and Measurable Goals, Objectives, and Outcomes

The purpose of Project MITOS is to adapt and scale the MITOS system for successful implementation at school and home that will lead to higher science achievement for all students, and particularly Latinx ELs. To accomplish this purpose, we propose the following objectives, strategies, and measurable outcomes.

Exhibit 1. Objectives, Strategies and Outcomes/Measures of MITOS

Objectives	Strategies	Outcomes/Measures
1. Adapt the MITOS app for game play in multiple devises for scaling	1.1 Gamify activities. 1.2 Translation of lessons into Spanish	1.1 Development team completes phases of DBR for gamifying activities. 1.2 Students and parents in pilot study can use Spanish to interact with MITOS 100% of the time.

Objectives	Strategies	Outcomes/Measures
	1.3. Recruit 10 students, parents, and teachers to provide feedback on the adaptations of MITOS. 1.4 Adapt and pilot the MITOS+ lessons for small group. 1.5 Enhance AS model to support multilingual scoring by collecting bilingual data from 100 students. 1.6 Enhance teacher dashboard for easy data visualization	1.3 All pilot participants provide feedback during the phases of DBR through surveys and interviews. 1.4 Ten teachers provide feedback on MITOS+. 1.5 Scoring by multilingual AS is comparable to multilingual human scoring. 1.6 Pilot teachers report high feasibility and ease of use.
2. Develop trainings for teachers and parents participating in the MITOS evaluation	2.1 Develop a one-time teacher professional development module. 2.2 Develop a parent workshop.	2.1 Pilot teachers report 100% satisfaction with training. 2.2 Pilot parents report 100% satisfaction with workshops.
3. Conduct a rigorous evaluation of MITOS to determine impact.	3.1 Recruit and randomly assign schools (Cohorts 1 and 2) to treatment and control. 3.2 AIR conducts implementation study to assess fidelity and quality.	3.1 60 schools are randomly assigned. 3.2 Fidelity and quality of implementation is above 90%

Objectives	Strategies	Outcomes/Measures
	3.3 Nonresponders receive MITOS+ lead by the teacher. 3.4 Complete data analyses	3.3 Increase in science content knowledge for all and ELP for ELs. 3.4 Research questions have been answered.
4. Dissemination	4.1 Creation of webpage to access all materials. 4.2 Written reports and manuscripts for researchers and practitioners 4.3 Use of social media to disseminate findings to parents	4.1 Website can be accessed by teachers. 4.2 Reports are completed. Publications in peer-reviewed journals and conferences are submitted. 4.3 Tools and strategies are shared through social media and workshops. Users indicate likes.

C3. Appropriate Project Design for Addressing the Needs of the Target Population

MITOS is designed to address the needs of second grade students to learn science but will include adaptations to meet the needs of Latinx ELs as a key target population to serve. For all students, MITOS includes NGSS and TEKS aligned content, engaging gamified activities, and accessibility by providing translations in Spanish and ease of use on any devise. For Latinx ELs, MITOS also draws on strong evidence for vocabulary development (S. K. Baker et al., 2014), using a variety of activities aligned with science content, and leveraging parent engagement (NASEM, 2019), and many opportunities to interact with others.

The project will draw from the existing evidence, including the evidence based on prior implementation of MITOS (S. K. Baker et al., 2021) but in addition will undertake piloting

during the initial year of the project to ensure suitability for students, teachers, and families on this project. After the piloting, the project evaluation (described in detail in section D) will rigorously test the efficacy of MITOS in authentic settings and explore for whom and in which settings it may be most effective.

C3.1 Design-Based Research to Pilot MITOS and Ensure its Use Addresses the Needs of Latinx ELs. In the first year of the project (2024-25) we will undertake pilot testing using a design-based research approach that enables input and equal participation of teachers, students, and families to adapt and refine MITOS. The front-end development team at UT Austin will follow a series of phases adhering to Design-Based Research (DBR, Barab & Squire, 2004) methodology, which includes an iterative, participatory process of developing and testing new activities in real settings.

DBR typically includes five phases—the design phase, the development phase, the research phase, the deployment phase, and the sustainability phase. We will recruit 10 teachers, students, and parents to provide input at each of the five phases across the year. Then, initial prototypes that use a gaming approach with the existing MITOS curriculum will be piloted by the collaborating students, parents, and teachers; teachers also will pilot the MITOS+ small group lessons. Participants will be asked questions based on the Structured Expert Evaluation Method (SEEM) for children's computer games (Baauw et al., 2005) and also will complete the Gameplay-Survey Scale (Zurita et al., 2020) to obtain quantitative data on their perceptions of the MITOS app. In addition, we will conduct interviews to gather information about the extent to which the translations are appropriate, the instructions are understandable in both languages, and activities are linguistically and culturally suitable for our target population. Data from surveys and interviews will be reviewed by the UT Austin development team to determine which features

or mechanics to modify, add, or delete, and what further linguistic or cultural tailoring is needed. We will revise as we move across the five design phases, attending to how well MITOS is serving students and Latinx ELs.

C.3.2 Piloting to Ensure the MITOS SRS and AS is Reliable in Spanish. MITOS will use the SRS and AS system from our previous work (Baker et al., 2014, 2018; Sano et al., 2020; Wu et al., 2023) with SMU leading the enhancement of model capability to support multi-lingual scoring. The game-based app will collect student speech as an input modality to the game. The backend of the MITOS app will save this audio into a cloud storage (e.g., Amazon S3 or equivalent) and then score students for their understanding of various scientific vocabulary words using machine learning methods from our previous work (Wu et al. 2023).

As we pilot over the 2024-25 school year with our 10 collaborating teachers, we will ask them to use the app with their bilingual students so that we can collect data from approximately 100 students using MITOS in English and Spanish. SMU will work inside the game environment and score student responses in the cloud. Responses in English and Spanish will be automatically transcribed, and we will update our existing algorithm for the AS system to score in both English and Spanish. To test reliability, SMU will employ human raters to listen to and provide scores for the responses as we have done in previous work (Wu et al. 2023) and compare the accuracy of the AS to the human raters in both English and Spanish. We will retrain the AS system until it provides scoring that is as reliable as human scoring. Our aim is to provide formative assessment information to teachers about student science understanding, regardless of the language that they are most comfortable using with the app.

C.3.3 Piloting to Ensure Parents and Families are Prepared to Support Student Engaging with MITOS. We will provide a one-time workshop for parents to support them in

their conversations about science at home. In the workshop, we will explain how the MITOS app works, what type of questions parents should ask, and how they can encourage their children to use the app four times per week for 10 minutes per day. Students who are not engaged will received MITOS+ at school. Latinx ELs will be able to speak to their parents in Spanish or in a combination of languages. [redacted] a Co-Pi on the project will work closely with families during the Pilot study to learn more about the best way to support parental engagement and motivation to support their learning of science with their children. Semi-structured interviews and observations during the Pilot study will be coded and used to refine the parent training and increase parental engagement.

D. Quality of the Project Evaluation

AIR will conduct an independent evaluation of MITOS that is designed to meet What Works Clearinghouse (WWC) standards without reservations. The evaluation will address the confirmatory, exploratory, and implementation research questions (RQs) shown in Exhibit 2. See a planned time for evaluation activities in Appendix J4.

Exhibit 2. Research Questions and Data Sources

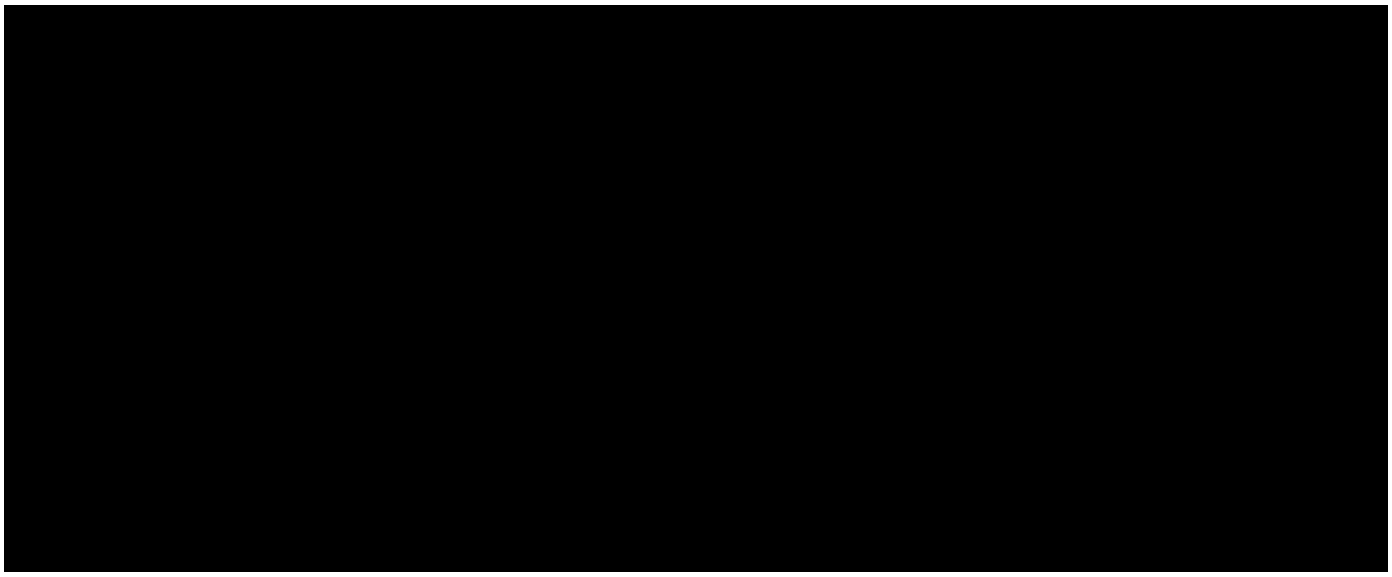
Research Question	Data Sources
Confirmatory Impact Analyses	
RQ 1: What is the impact of MITOS on student vocabulary development?	Teacher-administered student assessments
RQ 2: What is the impact of MITOS on student science knowledge?	Teacher-administered student assessments
RQ 3: What is the impact of MITOS on EL student ELP?	Student administrative data

Research Question	Data Sources
Exploratory Impact Analyses	
RQ 4: To what extent do student and school characteristics moderate the impact of MITOS on student outcomes?	Teacher-administered student assessments Student administrative data
RQ 5: To what extent does student vocabulary development mediate the impact of MITOS on student science knowledge and EL student ELP?	Teacher-administered student assessments Student administrative data
RQ 6: What is the impact of the MITOS small group adaptation (MITOS+) on student outcomes?	Teacher-administered student assessments Student administrative data
Implementation Analyses	
RQ 7: To what degree is MITOS implemented with fidelity across schools?	App usage data Teacher survey Teacher and family training records
RQ 8: What factors support or inhibit MITOS implementation?	Teacher and family interviews Teacher survey
RQ 9: What is the cost and cost-effectiveness of MITOS?	Teacher and family interviews Teacher survey

D1. Methods to Meet WWC Standards Without Reservations

The impact evaluation to answer RQs 1–6 will be conducted using an approach known as sequential, multiple assignment, randomized trial (SMART) design (Almirall et al., 2012; Oetting et al., 2007). Schools will be randomly assigned to treatment or control groups within district blocks. All grade 2 teachers and classes at treatment schools will be offered participation in MITOS. Grade 2 classes in the control schools will continue with "business-as-usual" classroom activities and opportunities for family engagement. The evaluation design includes an optimization sub study to test the efficacy of the scaling strategy, where teachers provide targeted small group support to students in need (MITOS+), as depicted in Exhibit 3 and detailed in Section D.2.

Exhibit 3. SMART Design for the Impact Evaluation and Program Optimization Sub study



Sample. The evaluation will take place across two cohorts (2025–26 and 2026–27) with 60 schools (30 treatment and 30 control) in the Austin Independent School District (AISD), Round Rock School District (RRISD), and the Cityscape School District (Cityscape ISD). Appendix J11 includes the description of the three sites. The sample will include approximately 240 teachers and 6,000 students, with an estimated 1,200 EL students. Each school will participate in the evaluation for 1 year, in either the first or second cohort. The confirmatory research analyses are powered to detect a minimum detectable effect size (MDES) of 0.15 for impacts on student vocabulary (RQ 1) and student science knowledge (RQ 2), and an MDES of 0.17 on EL student ELP (RQ 3).[1] These are reasonable and substantively meaningful, given a recent meta-analysis of the impacts of educational apps on the academic achievement of students

[1] The three-level power analysis with treatment at Level 2 (school) assumes the following: alpha = 0.05; a two-tailed test; power = 0.80; proportion of Level 2 units in the treatment group = 50%; $ICC_2 = 0.10$; R^2 Level 1 (student) = 0.75 (assumes pretest scores are available); R^2 block and Level 2 (district block and school) = 0.60. For RQs 1 and 2, the analysis assumes an average of 100 students per school. For RQ 3, the analysis assumes 20 students per school (across multiple sections).

in preschool to grade 3, which found average effects of 0.17 on standardized measures (Kim et al., 2021). Power analyses are provided in Appendix J.1.

D1.1 Confirmatory and Exploratory Impact Analyses (RQs 1–6). We will employ intent-to-treat (ITT) analyses to estimate the impact on focal student outcomes of being at a school randomly assigned to participate in MITOS. To answer RQ 1 and RQ 2, AIR will support teachers to administer two validated and reliable assessments to their students (in fall and spring for each cohort). For RQ 1, teachers will administer the easyCBM Vocabulary Subtest (Alonzo et al., 2006). For RQ 2, teachers will administer the Stanford Achievement Test Science Subtest, 10th Edition (Psychological Corporation, 2002). Both tests will be administered online in group settings and will take 10–15 minutes for students to complete. To examine EL students' ELP (RQ 3), AIR will collect student-level administrative data from participating districts on the statewide Texas English Language Proficiency Assessment System (TELPAS). To test MITOS+ (RQ 6), we will use models similar to those defined for RQs 1–3 but restricted to students in the treatment group with insufficient app usage (see Section D.2).

To assess the impact of MITOS on student outcomes, AIR will use a three-level model that accounts for student clustering within schools and districts, controlling for school and student characteristics, as well as baseline outcome measures. We will assess differential impacts of the MITOS program (RQ 4) on student outcomes by incorporating a treatment-by-moderator interaction term into the model, where the moderator is a characteristic of the student or school (see Section D.2). Finally, we will estimate the extent to which student vocabulary mediates the impact of MITOS on student science knowledge and EL student ELP outcomes (RQ 5) using a multilevel path analysis that estimates program effects on the mediator and the effect of the mediator on student outcomes. See Appendix J.2 for further details of all models.

AIR will use random school-level assignment, which means that minimal risk from contamination or other treatment crossover effects is expected. To minimize attrition—and importantly, school-level attrition—we will confirm participation and then carry out random assignment as late as possible at the beginning of each intervention year, thus positioning the evaluation to meet WWC standards *without reservations*. To mitigate the risk of joiners, we will follow the recommendations of the most recent WWC standards (version 5.0; 2022) by including only early student joiners, based on student lists collected during the first 6 weeks of the school year. To minimize missing student assessment data, we will offer $250 stipends to teachers for completion of the assessments and will use simple administration procedures to implement with students (e.g., 10–15 minutes per test, delivered via group administration) that we have successfully used in prior large-scale studies. In the case of high and/or differential attrition, we will use multiple imputation (Graham, 2009) to demonstrate that outcome imputation bias is within .05 standard deviations, or we will match non-attrited schools and students to ensure baseline differences on outcome measures are within .25 standard deviations and will then adjust for those differences in the analysis.

D2. Guidance About Effective Strategies Suitable for Replication or Testing

The proposed evaluation will provide robust information about the effectiveness of the MITOS scaling strategy and its suitability for use in other settings, through (a) an optimization study to help strengthen MITOS's adaptive approach, and (b) moderation and cost analyses.

D2.1 Optimization Study to Evaluate MITOS's Adaptive Approach. A key premise of the MITOS scaling approach is to create opportunities for student learning that can be leveraged both inside and outside of school settings, with students working independently or with an adult, including family members. For students who are not able to engage with the app as

intended, teachers will provide small group support to enhance their learning opportunities. To provide nuanced and rigorous evaluation of that scaling strategy, the evaluation's SMART design (Almirall et al., 2012; Oetting et al., 2007) will assess the impacts of randomly assigning students in 30 treatment schools across both cohorts who do not sufficiently engage with MITOS ("non-responders") to MITOS+ small group support with their teacher. Insufficient engagement is defined as using the MITOS app, on average, less than three times a week over the first 3 months of the program.

Based on UT Austin's previous experience with an earlier version of MITOS (Baker et al., 2021), we anticipate that approximately 35% of treatment group students will fall into this group of non-responders. The non-responding students will be entered into a second lottery, which will randomly assign half within school blocks to receive small group sessions twice a week with their teachers at school (see Section B1) and the other half to continue with the standard approach to MITOS, as represented in Exhibit 3. The SMART approach will allow for unbiased comparisons between non-responding students in the *small group* and *continue* groups and will estimate the impact of the small group sessions on student outcomes. This will provide rigorous evidence of the efficacy of the scaling strategy, which provides access through an app and then targets additional supports based on student needs.

D2.2 MITOS Suitability and Costs. The SMART design results will be combined with information from moderator and cost analyses that will elucidate for whom and where the program may be best suited, and the program's cost-effectiveness. Through RQ 4, we will test key moderators to understand for which students and in which types of school settings the program may be best suited (see Appendix J.2.1 for all moderators). For example, this will include examining differential impacts for current, former, and never EL students, and for

students from Spanish versus non-Spanish language backgrounds. Examples of school-level moderators include the percentage of EL students, average baseline achievement, and locale. In addition, for RQ 9, we will gather comprehensive information on resources used and time costs associated with MITOS implementation through multiple sources (e.g., app usage data, teacher and family interviews, teacher survey) and will apply an "ingredients" approach to build a resource cost model that captures the full underlying costs of the program, including distinguishing start-up costs from ongoing costs (Levin et al., 2018; see Appendix J.3). The cost analysis will calculate the average per-student cost and apply it to the impact findings for RQs 1–3 to estimate program cost-effectiveness.

D3. Clear Articulation of Components, Mediators, Outcomes, and Measurable Thresholds

The evaluation design is informed by clearly articulated key components and a key interim outcome (student vocabulary development) that mediates further student outcomes (science knowledge and ELP), as depicted in the logic model in Appendix G. MITOS key components include (a) a teacher training, (b) a parent workshop, (c) student use of the app, (d) teacher use of the data dashboard, and (e) MITOS+ small group sessions for students who need extra support. Through implementation, these components are expected to lead to improved student vocabulary development, which in turn is hypothesized to mediate the impact of MITOS on student science knowledge for all students and ELP among EL students.

The evaluation design will use multiple valid and reliable outcome measures that capture the outcomes in the logic model. Student vocabulary and science knowledge tests are well-established measures with demonstrated psychometric properties (α = .94 to .95; see Appendix J.5) that are widely used in research and by practitioners to inform regular instruction. AIR will measure EL student ELP by collecting student performance data on the TELPAS state

assessment for each evaluation year. This standardized measure is considered valid and reliable based on WWC standards and represents a policy-relevant assessment that is already embedded in the educational settings where the evaluation will take place.

D3.1 Measurable Implementation Thresholds. We will assess fidelity using measurable thresholds for acceptable implementation of all key MITOS components (see Appendix J.6 for details on data sources, collection methods, and timing). Fidelity indicators capture school-level implementation percentages for grade 2 classes and families based on the key components: (a) teachers complete the initial training, (b) families participate in a family workshop, (c) students use the app at least three times a week, (d) teachers review the dashboard at least once every 2 weeks, and (e) teachers provide small group support to students identified for MITOS+ two times per week. We will also create an overall, school-based fidelity measure that captures these indicators. Based on prior research on fidelity in randomized controlled trials (RCTs; Durlak & DuPre, 2008; Hill & Erickson, 2019), we will use the following school-level implementation fidelity thresholds for each key program component and for overall fidelity: low fidelity (i.e., inadequate; less than 60% complete the program component), moderate fidelity (60%–80%), and high fidelity (above 80%). We will consider moderate or high fidelity an acceptable level of implementation, based on prior research (Durlak & DuPre, 2008).

D4. Procedures for Ensuring Feedback and Periodic Assessment of Progress

D4.1 Performance Feedback. To provide UT Austin with meaningful performance feedback, AIR will evaluate implementation fidelity (RQ 7) of program use based on program training records and app usage data (see Section D.3 for specific thresholds to assess acceptable implementation). AIR will meet each month with UT Austin to discuss findings from the prior month's data. This will provide information not only on fidelity, but also on the feasibility of

implementing MITOS as designed.

To understand teacher and family perceptions of factors that support or inhibit implementation (RQ 8) and program costs (RQ 9), AIR will conduct teacher and family member interviews. Midway through each evaluation year, we will identify 20 MITOS classes with high and low levels of program fidelity (see Section D3). We will conduct semi structured interviews with the 20 teachers of those classes to capture teacher perceptions of the usability and usefulness of the MITOS app and its dashboard; the relevance of the content to their learning goals for students; their implementation of supplementary lessons and small group instruction; and their time spent on MITOS, to support the cost analysis. AIR will also conduct eight focus groups (with approximately six family members in each group) from the sampled classrooms in each cohort to gather family perspectives on MITOS and what supported or hindered their family's engagement. Teacher perceptions of MITOS and time spent on MITOS will be gathered from treatment group teachers in the spring teacher survey, allowing reporting from all treatment group teachers.

D4.2 Progress Toward Intended Outcomes We will assess progress toward intended outcomes in two ways. First, AIR's monthly monitoring of usage activity among classes participating in the program will support ongoing understanding of progress toward program participation as intended. Second, AIR will examine initial estimates of program impacts on student vocabulary development and science knowledge (measured via the assessments teachers administer in treatment and control groups) at the end of participation for the first cohort, providing an opportunity to gauge progress toward two key outcomes midway through the study.

References

Almirall, D., Lizotte, D. J., & Murphy, S. A. (2012). Smart design issues and the consideration of opposing outcomes: Discussion of "evaluation of viable dynamic treatment regimes in a sequentially randomized trial of advanced prostate cancer" by Wang, Rotnitzky, Lin, Millikan, and Thall. *Journal of the American Statistical Association*, *107*(498), 509–512. https://doi.org/10.1080/01621459.2012.665615

Alonzo, J., Tindal, G., Ulmer, K., & Glasgow, A. (2006). *EasyCBM® online progress monitoring assessment system*. Eugene, OR: Center for Educational Assessment Accountability. http://easycbm.com

Baauw, E., Bekker, M. M., & Barendregt, W. (2005). A structured expert evaluation method for the evaluation of children's computer games. In M. F. Costabile & F. Paternò (Eds.), *Human-Computer Interaction—INTERACT 2005* (Vol. 3585, pp. 457–469). Springer Berlin Heidelberg. https://doi.org/10.1007/11555261_38

Baker, D. L. (Principal Investigator). (2014–2017). *English learner vocabulary acquisition (ELVA): Promoting the vocabulary and language proficiency of Spanish speaking English learners in second grade* (Project No. R305A140471) [Grant]. National Center for Education Research. https://ies.ed.gov/funding/grantsearch/details.asp?ID=1467

Baker, D. L. (Principal Investigator). (2020–2024). *Measuring the English language vocabulary acquisition of Latinx bilingual students* (Project No. R305A200521) [Grant]. National Center for Education Research. https://ies.ed.gov/funding/grantsearch/details.asp?ID=3402

Baker, D. L., Basaraba, D. L., & Richards-Tutor, C. (Eds.). (2018). *Second language acquisition: Methods, perspectives, and challenges*. Nova Science Publishers.

Baker, D. L., & Le, N. (under review). *Developing the science vocabulary of second grade girls using a learning management system*.

Baker, D. L., Conry, J., & Polanco, P. (2018, May 14). *English language vocabulary acquisition project*. http://stemforall2018.videohall.com/presentations/1202

Baker, D. L., Cummings, K., & Smolkowski, K. (2022). Diagnostic accuracy of Spanish and English screeners with Spanish and English criterion measures for bilingual students in Grades 1 and 2. *Journal of School Psychology*, *92*, 299–323. https://doi.org/10.1016/j.jsp.2022.04.001

Baker, D. L., Ma, H., & Gallegos, E. (2019). Increasing parental knowledge through workshops on early childhood programs, home literacy, and technology use. *Bilingual Research Journal*, *42*(4), 491–509. https://doi.org/10.1080/15235882.2019.1666755

Baker, D. L., Ma, H., Polanco, P., Conry, J. M., Kamata, A., Al Otaiba, S., Ward, W., & Cole, R. (2021). Development and promise of a vocabulary intelligent tutoring system for Second Grade Latinx English learners. *Journal of Research on Technology in Education*, *53*(2), 223–247. https://doi.org/10.1080/15391523.2020.1762519

Baker, D. L., McCoach, B. D., Ware, S., Coyne, M. D., & Rattan, S. M. (2022). Effects of Spanish vocabulary knowledge on the English word knowledge and listening comprehension of bilingual students. *International Journal of Bilingual Education and Bilingualism*, *25*(6), 2269–2283. https://doi.org/10.1080/13670050.2021.1908219

Baker, D. L., Mogna, V., Rodriguez, S., Farmer, D., & Yovanoff, P. (2016). Building the oral language of young Hispanic children through interactive read alouds and vocabulary games at preschool and at home. *Journal of International Special Needs Education*, *19*(2), 81–94. https://doi.org/10.9782/2159-4341-19.2.81

Baker, S. K., Lesaux, N., Jayanthi, M., Dimino, J., Proctor, C. P., Morris, J., Gersten, R., Haymond, K., Kieffer, M. J., Linan-Thompson, S., & Newman-Gonchar, R. (2014). Teaching academic content and literacy to English learners in elementary and middle school. IES practice guide (NCEE 2014-4012). In *What Works Clearinghouse*. What Works Clearinghouse. https://eric.ed.gov/?id=ED544783

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences*, *13*(1), 1–14. https://doi.org/10.1207/s15327809jls1301_1

Chapin, S. H., O'Connor, C., & Anderson, N. C. (2003). Classroom discussions using math talk in elementary classrooms. *Math Solutions*. https://mathsolutions.com/uncategorized/classroom-discussions-using-math-talk-in-elementary-classrooms-pdf

Chiu, Y.-H. (2013). Computer-assisted second language vocabulary instruction: A meta-analysis. *British Journal of Educational Technology*, *44*(2), E52–E56. https://doi.org/10.1111/j.1467-8535.2012.01342.x

Cummins, J. (2005, September). *Teaching for cross-language transfer in dual language education: Possibilities and pitfalls. In Teaching and learning two languages in the EFL setting* [Symposium]. TESOL Symposium on Dual Language Education, Istanbul, Turkey.

Department of Education (2023). Applications for new awards: Education, Innovation and Research (EIR) Program-Mid-Phase Grants. *Federal Register* retrieved on 7/10/2023 from: https://www.federalregister.gov/documents/2023/05/23/2023-11001/applications-for-new-awards-education-innovation-and-research-eir-program-mid-phase-grants.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting

implementation. *American Journal of Community Psychology*, *41*(3–4), 327–350. https://doi.org/10.1007/s10464-008-9165-0

Guzman, L., Hickman, S., Turner, K., & Gennetian, L. (2016). *Hispanic children's participation in early care and education: Parents' perceptions of care arrangements, and relatives' availability to provide care*. https://www.hispanicresearchcenter.org/wp-content/uploads/2019/08/NSECE-Series-Parent-Perceptions-V21.pdf

Higher Education Act, 1067q 20 U.S. Code (2013).

Hill, H. C., & Erickson, A. (2019). Using implementation fidelity to aid in interpreting program impacts: A brief review. *Educational Researcher*, *48*(9), 590–598. https://doi.org/10.3102/0013189X19891436

Kim, J., Gilbert, J., Yu, Q., & Gale, C. (2021). Measures matter: A meta-analysis of the effects of educational apps on preschool to grade 3 children's literacy and math skills. *AERA Open*, *7*, 23328584211004184. https://doi.org/10.1177/23328584211004183

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Krogstad, J. M. (2014, September). Hispanics only group to see its poverty rate decline and incomes rise. *Pew Research Center*. https://www.pewresearch.org/short-reads/2014/09/19/hispanics-only-group-to-see-its-poverty-rate-decline-and-incomes-rise/

Levin, H. M., McEwan, P., Belfield, C., Bowden, A. B., & Shand, R. (2018). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis* (Third Edition). SAGE Publications. https://methods.sagepub.com/book/economic-evaluation-in-education-3e

Loewen, S., Crowther, D., Isbell, D. R., Kim, K. M., Maloney, J., Miller, Z. F., & Rawal, H. (2019). Mobile-assisted language learning: A Duolingo case study. *ReCALL*, *31*(3), 293–311. https://doi.org/10.1017/S0958344019000065

Mayer, R. E. (2009). *Multimedia learning* (2nd edition, pp. xiii, 304). Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

McKeown, M. G. (2017). *Vocabulary assessment to support instruction: Building rich word-learning experiences*. The Guilford Press.

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educational Researcher*, *45*(1), 18–35. https://doi.org/10.3102/0013189X16633182

Murphy, P. K., Wilkinson, I. A. G., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology*, *101*(3), 740–764. https://doi.org/10.1037/a0015576

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association*, *18*(5), 544–551. https://doi.org/10.1136/amiajnl-2011-000464

National Academies of Sciences, Engineering, and Medicine. (2016). *Parenting matters: Supporting parents of children ages 0-8*. National Academies Press. http://www.ncbi.nlm.nih.gov/books/NBK402024/

National Academies of Sciences, Engineering, and Medicine. (2018). *English learners in STEM subjects: Transforming classrooms, schools, and lives* (p. 25182). National Academies Press. https://doi.org/10.17226/25182

National Academies of Sciences, Engineering, and Medicine. (2021). *Teaching K-12 science and engineering during a crisis* (p. 25909). National Academies Press. https://doi.org/10.17226/25909

National Center for Education Statistics. (2023). English Learners in Public Schools. In *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/programs/coe/indicator/cgf

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. National Academies Press. https://doi.org/10.17226/11625

National Research Council. (2013). *Next Generation Science Standards: For States, By States*. National Academies Press. https://doi.org/10.17226/18290

Oetting, A. I., Levy, J. A., Weiss, R. D., & Murphy, S. A. (2011). Statistical methodology for a SMART design in the development of adaptive treatment strategies. In P. E. Shrout, K. M. Keyes, & K. Ornstein (Eds.), *Causality and Psychopathology: Finding the Determinants of Disorders and their Cure* (pp. 179–205). American Psychiatric Publishing.

Olmstead, C. (2013). Using technology to increase parent involvement in schools. *TechTrends*, *57*(6), 28–37. https://doi.org/10.1007/s11528-013-0699-0

Plonsky, L., & Sudina, E. (2023). *The effects of frequency, duration, and intensity on L2 learning through Duolingo: A 'natural' experiment* [Research Report]. https://duolingo-papers.s3.amazonaws.com/reports/Plonsky_etal_whitepaper_language_learning_grit_motivation_2023.pdf

Psychological Corporation. (2002). *Stanford achievement test* (10th ed.). Harcourt Brace & Company, Educational Measurement. https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-

Assessments/Academic-Learning/Comprehensive/Stanford-Achievement-Test-Series-%7C-Tenth-Edition/p/100000415.html

Sano, M., [redacted] [redacted] Collazo, M., Le, N., & Kamata, A. (2020). Measuring the expressive language and vocabulary of Latino English learners using hand transcribed speech data and automated scoring. *International Journal of Intelligent Technologies and Applied Statistics*, *13*(3), 229–256. https://doi.org/10.6148/IJITAS.202009_13(3).0003

Soter, A. O., Wilkinson, I. A., Murphy, P. K., Rudge, L., Reninger, K., & Edwards, M. (2008). What the discourse tells us: Talk and indicators of high-level comprehension. *International Journal of Educational Research*, *47*(6), 372–391. https://doi.org/10.1016/j.ijer.2009.01.001

Texas Education Agency. (2017). *19 TAC Chapter 112. Texas essential knowledge and skills for science*. Texas Education Agency. https://tea.texas.gov/about-tea/laws-and-rules/texas-administrative-code/19-tac-chapter-112

US Census Bureau (2021). Computer and Internet use in the United States: 2018. Retrieved on 7/10/2023 from https://www.census.gov/newsroom/press-releases/2021/computer-internet-use.html#:~:text=Smartphone%20ownership%20surpassed%20ownership%20of,ownership%20fell%20behind%20at%2063%25.

What Works Clearinghouse. (2011). *Procedures and standards handbook (version 2.1).* Retrieved from http://ies.ed.gov/ncee/wwc/ pdf/reference_resources/wwc_procedures_v2_1_standards_handbook.pdf

What Works Clearinghouse. (2022). *Procedures and standards handbook, (version 5.0).* U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). https://ies.ed.gov/ncee/wwc/Handbooks

Wu, Z., Larson, E. C., San, M., Baker, D., Gage, N., & Kamata, A. (2023, July 20). *Towards scalable vocabulary acquisition assessment with BERT*. ACM International Conference on Learning at Scale (L@S 2023), Copenhagen, Denmark.

York, B. N., Loeb, S., & Doss, C. (2019). One step at a time: The effects of an early literacy text-messaging program for parents of preschoolers. *Journal of Human Resources*, *54*(3), 537–566. https://doi.org/10.3368/jhr.54.3.0517-8756R

Zurita Ortega, F., Medina Medina, N., Gutiérrez Vela, F. L., & Chacón Cuberos, R. (2020). Validation and psychometric properties of the gameplay-scale for educative video games in Spanish children. *Sustainability*, *12*(6), 2283. https://doi.org/10.3390/su12062283