AI Across the Curriculum for Virtual Schools

1 SIGNIFICANCE

The Concord Consortium, in collaboration with Texas Tech University (TTU), University of Florida (UF), and WestEd, proposes a five-year early-phase project to develop a year-long *AI in Math* supplemental certificate program for Algebra I or Integrated Math 1 class. The program will provide high-need students the opportunity to develop Artificial Intelligence (AI) literacy and self-efficacy in solving problems using AI and learning AI topics and simultaneously improve their math learning and attitudes toward math. This project defines **high-need students** as female or racial/ethnic minority students underrepresented in the computing field and underserved students who are economically disadvantaged, living in remote rural areas, or enrolled in under-resourced schools. To broaden its access to high-need students, the *AI in Math* program will be designed for self-directed virtual learners and suitable for virtual schools to utilize. This project will implement the *AI in Math* program in two statewide public virtual schools, Florida Virtual School (FLVS) and Texas Tech University K-12 (TTU K-12), and reach a total of 68 teachers and 6800 students. WestEd, the external evaluator, will provide continuous feedback and conduct an impact study to produce evidence of the program's effectiveness that meets What Works Clearinghouse (WWC) standards without reservations.

Absolute Priority 1—Applications that Demonstrate a Rationale. This project builds on the PIs' prior federal-funded works on integrating AI education into the core school curriculum. These prior works not only suggest that the disciplinary integration approach is promising but also provide a solid foundation for implementing and evaluating the approach as well as expanding it across disciplines, grade levels, and virtual school systems.

Absolute Priority 3—Field-Initiated Innovations—Promoting Equity in Student Access to Educational Resources and Opportunities: Science, Technology, Engineering, or

Mathematics (STEM). This project addresses the persistent and significant challenge for high-need students to access high-quality AI education. Distributing the *AI in Math* program via virtual schools will reach underserved students in rural areas and under-resourced schools. Integrating AI topics into the core curriculum prevents high-need students from opting out of elective opportunities. A suite of specialized learning technologies designed to visualize abstract math concepts will allow students to explore AI models in interactive ways.

Competitive Preference Priority 1—Promoting Equity in Student Access to Educational Resources and Opportunities: Implementers and Partners. TTU, a key partner in this project, is a Hispanic-serving institution (with 29% Hispanic and 49% female students, see Appendix J1 Eligibility Letter). Co-PI [redacted], a Latina professional and Chief Equity Officer at the Lastinger Center for Learning at UF, will oversee the equitable design of the *AI in Math* program. Both UF and TTU will recruit female and minority graduate students to join the project.

1.1 Background: AI Literacy in K-12

Artificial Intelligence (AI) is transforming numerous industries and generating enormous wealth. However, the advancement in AI is reshaping the workforce, impacting people whose jobs can be replaced or redefined by AI systems. The wealth generated by AI advancement is unevenly distributed across different demographic groups, exacerbating existing inequities in society. Inequalities arising from current AI development are partially rooted in the unequal access to AI educational opportunities.

AI literacy. K-12 is the critical stage for young people to develop AI literacy and interest in AI-related careers. AI literacy is "a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace.” (Long & Magerko, 2020, p.598). K-12 students should learn about the big ideas in AI, including (1) how computers perceive the world using sensors, (2) how AI agents maintain the representation of the world and reason with it, (3) how computers learn from data, (4) how AI agents interact naturally with humans, and (5) the social impacts of AI development (Touretzky et al., 2019). Ultimately, students should understand the roles and responsibilities of AI developers and pathways for their participation in AI development.

Existing strategies. In recent years, many organizations have been developing AI education resources for K-12 students and teachers. Code.org developed a 20-hour *AI and Machine Learning* curriculum unit for Grades 6-12 (Code.org, 2023). Exploring Computer Science (ECS), a year-long high school intro-level CS curriculum, also offered an alternative unit on AI (Exploring Computer Science, 2022). Computer Science Frontiers, a new open-source curriculum, created a 9-week AI and Machine Learning module. AI4GA, a team of university faculty and middle school teachers, has been developing a 9-week middle school AI elective course (AI4GA, 2023). MIT’s *Day of AI* curriculum covers many topics, such as natural language processing, generative AI, and AI ethics (MIT RAISE, 2023).

These existing curricula are commonly centered on AI and geared toward CS classes, standalone electives, or extracurricular programs. However, CS classes are only offered in 53% of U.S. high schools; disparities in access and diversity issues still persist (Code.org, CSTA, & ECEP Alliance, 2022). Many students, especially those from groups underrepresented in the computing field, may not perceive the relevance and value of learning about AI and hence opt out of such elective opportunities. Another common feature is that these existing curricula explicitly or implicitly assume brick-and-mortar schools as the users. However, many schools do not have the teacher capacity and resources to take advantage of these curricula.

1.2 New Strategy 1: Integrating AI Education into Core Disciplines

Instead of adding AI topics in CS courses or creating standalone AI courses, we propose integrating AI education into core disciplines such as mathematics, science, and English Language Arts. Within the scope of this early-phase project, we will focus on integrating AI education into the high school math curriculum. Specifically, we will develop a year-long AI in Math supplemental certificate program for Algebra I or Integrated Math 1 class. This new strategy is informed by the following perspectives, research or evaluation findings from the PIs' prior work and the knowledge base built by the broader STEM education research community.

Interdisciplinary nature of AI. AI is a highly interdisciplinary field. It builds on mathematical foundations and relies on disciplinary knowledge about the type of intelligence to simulate. Many AI innovations stem from the attempt to solve problems in subject domains outside CS. In the workplace, it is very common for non-CS professionals to learn and apply AI knowledge and skills and collaborate with computer scientists to solve problems in their own domains. In the same way, AI education can happen in a variety of settings where the problems of interest call for AI solutions.

Integrating foundational AI education into disciplinary studies promises to transform AI education and reach students most underrepresented and underserved in the field. The key to this approach is to situate student learning in scenarios where disciplinary insights are critical for AI development and AI applications give rise to new disciplinary practices. By leveraging the intrinsic connections between AI and disciplines already taught in schools, we envision a series of learning opportunities, presenting discipline-specific scenarios for students to dive deep into aspects of AI and to develop awareness and interest in various AI applications and careers.

Integrating AI education into English language arts classroom. In the NSF-funded StoryQ project (NSF DRL-1949110), PI [redacted]s team has developed an AI curriculum module for high school English Language Arts (ELA) class (Chao et al., 2023). In this module, acting as computational linguists-in-training, students work with linguistically and culturally rich text such as customer reviews and clickbait headlines. They interact with models that are simple in terms of the ML and feature extraction methods used but sophisticated with respect to how language works in context and the representation of unstructured text data. Students gain a nuanced understanding of language and how to wield it, not just as a data structure, but as a tool in our human-human encounters as well. Results from a pilot study in 2021 showed promising evidence that this approach supported students in participating in AI practices (Jiang, et al., 2023), developing AI literacy (Tatar et al., 2022), and increasing self-efficacy in explaining how AI works (Tatar et al., 2023).

Integrating AI education into science classrooms. Logic is a foundation for both AI and STEM. In an NSF-funded LPK12 project, Co-PI [redacted] has developed a logic-based framework and curriculum to integrate AI concepts and practices into science learning. Inspired by symbolic AI such as knowledge and expert systems (Jackson, 1998), in logic-based science learning, students use an intuitive logic programming language to represent objects, attributes, and relations, build models for the target domain, and iteratively refine the models so that they can automatically generate answers to science questions. This process of abstracting and representing not only allows students to develop computational thinking skills and fluency with a programming language but also helps them deepen their understanding of science topics. Three middle school science teachers have implemented the curriculum modules with 450 students. Results showed that the modules had positive impacts on student learning of both science and logic programming (Zhang et al., 2019a, 2019b, Archer et al., 2023) and on the professional learning of teachers with different STEM backgrounds (Wang et al., under review).

1.3 New Strategy 2: Virtual Learning

To broaden its access to high-need students, the *AI in Math* program will be designed for self-directed virtual learners and suitable for virtual schools to utilize. Virtual schools have emerged as a successful and innovative mechanism for addressing the educational needs of high-need students, especially in underserved regions (National Education Policy Center, 2015). Traditional brick-and-mortar schools often face limitations in accommodating students from remote regions or economically disadvantaged backgrounds due to geographical constraints and resource disparities (Lavalley, 2018). In contrast, virtual schools transcend these barriers by offering a flexible and accessible learning environment that can be accessed from anywhere with an internet connection. This inclusivity enables students who may have previously struggled to access quality education to now engage in rigorous academic programs and tap into a vast array of learning resources.

By embracing various learning styles and paces, virtual schools empower students to progress at their own rate, cultivating a deeper understanding of the subject matter and a higher level of engagement (Curtis & Werth, 2015). Moreover, the use of interactive multimedia, simulations, and gamification techniques in virtual learning environments enhances student motivation and promotes active learning (Baker & Gossman, 2013; Bouchrika et al., 2021; Vlachopoulos & Makri, 2017). These dynamic teaching methods can be particularly beneficial for high-need students who may thrive in a more interactive and immersive educational setting.

2 PROJECT DESIGN

2.1 Conceptual Framework

The design of this project is informed by the Social Cognitive Career Theory (SCCT, Lent et al., 1994; Lent & Brown, 2013). According to SCCT, students' interests, choices, and performance in a certain domain are largely influenced by their self-efficacy and outcome expectations. Self-efficacy is one's belief in his or her ability to perform a task. Outcome expectations are one's beliefs about the consequences of performing the task (Bandura, 1986). Both self-efficacy and outcome expectations are influenced by prior learning experiences that convey four types of information: personal performance accomplishments, observational learning, social persuasion, and physiological and affective states and reactions (Lent et al., 2017). Among these, personal performance experiences account for the most variance in self-efficacy beliefs than other information sources.

Based on SCCT, we identify personal performance accomplishments in the *AI in Math* program as the key project component. On the one hand, the integrated learning tasks provide a direct experience for students to develop an understanding of fundamental AI concepts, which will serve them well in future opportunities to learn advanced AI topics. Their presence in the core math curriculum also sets the expectation that all students should understand the fundamentals of AI. On the other hand, these integrated learning tasks offer contemporary, exciting contexts for students to apply and refine their math knowledge and skills. Students also gain an awareness of the important role of mathematics in cutting-edge AI technologies that are shaping their own lives and transforming society. Both improved math achievement and attitudes toward math will contribute to students' interest in further learning advanced AI topics. These mechanisms form the basis of our logic model (Appendix G).

2.2 Goals, Objectives, and Outcomes

The goal of this project is to promote equity in AI education resources for high-need students. The project team will accomplish this central goal by achieving the objectives in the table below. Each objective has multiple outcome measures from various data sources (see detailed measures in Grant Application Form for Project Objectives and Performance Measures Information) and aligned development and implementation activities (see details in Section 2.3).

Table 1. Project objectives, outcome measures, data sources, and proposed activities

Measures	Data Sources	Proposed Activities
Objective 1: Increase the number of high-need students having access to *AI in Math*		
Number of students having access to the program	LMS system logs Class rosters	Maintain CC's LMS; Create offline version; Embed lessons in virtual schools;
Objective 2: Increase the number of high-need students using *AI in Math*		
Percentage of students with satisfactory completion rate; Students' satisfaction;	LMS system logs; Students' demographics; End-of-lesson surveys;	Develop recruitment and retainment strategies; Design and award certificates;
Objective 3: Increase high-need students' AI literacy and AI self-efficacy		
Students' AI literacy; Students' AI self-efficacy;	Pre- and post- assessments, surveys, and interviews;	Develop and implement *AI in Math* program;
Objective 4: Improve high-need students' math achievement and attitudes toward math		
Students' math achievement; Students' attitudes toward math;	Prior year assessment scores; End-of-course assessment scores;	Develop and implement *AI in Math* program;
Objective 5: Develop teachers' competency for implementing *AI in Math* program		
Teachers' AI literacy; Teachers' AI self-efficacy	Pre- and post-PD surveys and interviews	Develop and implement a teacher PD program

2.3 Meeting the Needs of the Target Population

This project defines **high-need students** as female or racial/ethnic minority students underrepresented in the computing field and underserved students who are economically disadvantaged, living in remote rural areas, or enrolled in under-resourced schools. These high-need students typically have little access to high-quality AI education resources. Even when opportunities are available, they may perceive them as irrelevant due to low expectations of success and low subjective task value. Therefore, providing access and tailoring AI learning experiences to their needs is critical to our goal and objectives.

2.3.1 Design *AI in Math* program for high-need students

Aligning with math standards. Our curriculum design will be guided by the K-12 AI Guidelines (Touretzky et al., 2019) and the Common Core State Standards (CCSS), Florida and Texas State Standards for math. The *AI in Math* program will include ten lessons (1 to 3 hours per lesson), each introducing one or two AI concepts that apply or build on mathematical ideas. For instance, the *perceptron* is a classic linear classifier and the building block of artificial neural networks that underlie modern AI development. From the mathematics point of view, perceptron is an application of linear function in computation. Thus, perceptron can be used as an application context to reinforce students' learning of linear function or, alternatively, as a context to motivate the learning of linear function. Either approach would expose students to fundamental ideas in AI and highlight the relevance and importance of mathematics (see Appendix J3 Sample Lesson).

Culturally relevant pedagogy. Following the culturally relevant pedagogy (Byrd, 2016), the lessons will introduce AI concepts and practices in problem contexts that students care about and draw on their own funds of knowledge. For instance, the concept of perceptron will be introduced in the context of everyday decision-making, such as whether to watch a movie or cancel a sports event. Topics of interest are highly varied across different cultural groups and communities. We will work closely with teachers and students to select problem contexts that are relevant to students, especially those underrepresented in the computing field. The lesson will also expose students to real-world applications of AI concepts, the social impacts and ethical issues of the applications, and related careers and jobs performed by professionals with diverse demographic backgrounds.

Table 2. Math standards, AI concepts, and sample contexts in AI in Math lessons

Math Standards	AI Concepts	Sample Contexts
HSF-IF.B.4. Interpret functions that arise in applications in terms of the context	*Perceptron* is a linear classifier and the building block of artificial neural networks.	Decide to watch a movie or not.
HSF-BF.B. Build new functions from existing functions	*Multilayer Perceptron* stacks perceptrons together, or in math terms, builds new functions from existing functions.	Computer vision for handwritten digit recognition.
HSF-BF.A. Build a function that models a relationship between two quantities	*Logic* is a language for representing the world as objects and relations among them, which are core elements of AI. Functions are a special case of relations.	Find health solutions using rich medical knowledge represented by logic and functions.
HSF-LE.A. Construct and compare linear, quadratic, and exponential models and solve problems	*Simulated Annealing Local Search* is an important search method. A key step of this algorithm relies on exponential functions.	Schedule classes at the right time and venue with many constraints.
HSA-CED.A. Create equations that describe numbers or relationships	*Constraints satisfaction is* an AI approach to solving problems by representing *variables* and *constraints*, which include equations and inequalities.	Solving word problems or puzzles by representing them using constraints or equations
HSA-REI. Reasoning with Equations and Inequalities	*Decision Tree* algorithm uses a tree-like model to represent problems and make decisions or predictions by evaluating a series of equations and inequalities.	Decision trees for family weekend activity selection
HSA-REI.A. Understand solving equations as a process of reasoning and explain the reasoning	*Search algorithm*. Search is a fundamental idea in AI for solving general problems and logical reasoning.	Algorithms to find solutions to puzzles based on their representation.
HSS-ID.A Summarize, represent, and interpret data on a single count or measurement variable	*Zero Rule* is the benchmark procedure for classification algorithms.	Algorithm bias in training data, e.g., imbalance dataset favoring certain population.
HSS-ID.B. Summarize, represent, and interpret	*One Rule* algorithm is a simple yet powerful approach that searches for the	Predict conditions for canceling a sports event.

data on two categorical and quantitative variables	strongest relationships between input and output variables.	
HSS-ID.C. Interpret linear models	*Linear regression* algorithm computes the linear relationship between input and output variables and makes predictions based on the linear model.	Predicting influencer's reach on social media.

2.3.2 Support high-need students with specialized learning technologies

We will support high-need students with specialized learning technologies designed to visually represent abstract concepts and allow them to interact with multiple representations in a connected and dynamic fashion. CC has developed a suite of learning tools through funding from the National Science Foundation. This project will leverage these resources and expand their functionalities to achieve our objectives (see Appendix J4 Supporting Learning Technologies).

Maintaining mature technologies. We will maintain a set of mature technologies, including (1) Common Online Data Analysis Platform (CODAP) and many of its plugins. CODAP is a free educational software for data analysis. It is designed as a platform for developers and as an application for students in grades 6-14. (2) The Story Builder plugin allows students to capture moments as they work with CODAP. They can put these moments together to tell a story to others about an investigation or even to themselves as a reminder of what they were thinking and doing, and (3) Data portal plugins for CODAP.

Refining beta-version technologies. We will refine two beta-version technologies to support students in exploring machine learning concepts and practices: (1) *StoryQ plugin for CODAP*, a web-based machine learning and text mining tool that allows young learners (grades 6-12) to engage in machine learning practices and work with unstructured text data without coding; (2) *Decision Tree plugin for CODAP* allows students to develop and test decision tree models to make predictions or decisions for various situations.

Improving alpha-version technologies. We will also improve an alpha-version Markov Chain plugin for CODAP, a web-based Markov Chain modeling tool that allows students to build, test, analyze, and use sequence models for real-world phenomena.

Maintaining research-intensive LMS. In addition to these data science and machine learning tools, we will also use CC's Activity Player (AP) to deliver the program. AP is a full-featured and research-intensive Learning Management System (LMS). AP has several components: (1) a *Student Edition* with learning tasks, embedded interactives, and assessments; (2) a *Teacher Edition* that mirrors the Student Edition and provides additional background information, theories and pedagogies, sample student work, and implementation tips for teachers; (3) a *Class dashboard* that offers teachers a grid of students' work that updates in real-time; (4) *User logs* that capture altra fine-grained student and teacher usage data for researchers and evaluators to measure learning progresses and outcomes.

2.3.3 Provide access to high-need students through virtual schools

To provide access to high-need students, we will design the *AI in Math* program for virtual schools to offer to their students. Virtual schools provide flexible and convenient learning options to a wide range of students, especially those in remote or underserved areas, with disabilities or chronic illnesses, or who may benefit from self-paced and personalized learning. According to a recent report (Molnar et al., 2023), in 2022, there were a total of 726 full-time virtual schools enrolling 643,930 full-time students in the U.S. Curriculum designed for virtual learning settings has the potential to scale its reach and impact on high-need students through adoption by many virtual schools.

In this project, we will implement and evaluate the *AI in Math* program in two statewide virtual schools: Florida Virtual School (FLVS) and Texas Tech University K-12 (TTU K12).

Both virtual schools are our long-time collaborators and are committed to supporting the program implementation and research data collection (see Appendix C: Letters of Support).

FLVS is a fully accredited, statewide public school district offering more than 190 courses to Kindergarten-12 students. Its certified teachers use a variety of personalized instructional programs to create individualized educational plans for every student. FLVS currently enrolls 248,616 students in its full-time program and flex program. The demographics of Grades 9-12 flex program students are 58% female, 16.7% Black or African American, 2.1% American Indian or Alaska Native, 0.7% Native Hawaiian or Other Pacific Islander, and 6.4% Multi-Racial. 15% of the completed courses are by students in rural schools and 13% by students in high-minority schools. In addition, over the past three years, the flex program has consistently witnessed substantial engagement from Hispanic or Latino students, with approximately 25% actively enrolling in computing and technology-related courses. This trend highlights the program's success in attracting underserved students in these specialized areas of education.

TTU K-12, a unit of the Texas Tech University eLearning & Academic Partnerships division, is a state-approved online kindergarten through 12th grade school that has been meeting students' needs for more than 30 years. TTU K-12 serves a large number of Black or African American (12.8%) and Hispanic or Latino (52.8%) students in Texas.

We will establish connections between CC’s AP system and the LMSs used by the virtual schools. Students will be able to either sign in to AP using their virtual school accounts to complete the program or complete the program embedded in their virtual school’s LMS. Virtual school teachers will use *AI in Math* lessons in relevant places in their Algebra I or Integrated Math 1 course. For instance, the *Linear Function & Perceptron* lesson can be embedded in the linear function unit. Students who complete the *AI in Math* program as part of their math class will receive an AI Literacy certificate issued by UF or TTU.

2.3.4 Develop guidelines for recruiting and retaining high-need students

We will develop a set of high-need student recruitment and retainment guidelines for math teachers. The guidelines will be designed to maintain or exceed the gender and racial/ethnic ratio representative of the U.S. demographics. Recruitment materials will include (1) a letter to students explaining the benefits and requirements of the certificate program; (2) experience reports from students who participated in the pilot study and earned the certificate; (3) a letter to parents in English and Spanish explaining the benefits and requirements of the certificate program. Teachers will check in with high-need students on a regular basis to gauge their progress and provide formative feedback.

2.3.5 Train teachers to support high-need students

We will develop and implement a 30-hour online teacher professional development (PD) program to train virtual school teachers to use the *AI in Math* program. Both FLVS and TTU K-12 are committed to helping recruit all 60 Algebra I teachers (58 at FLVS and 2 at TTU K12) to participate (See letters of support from the virtual schools in Appendix C). The PD program will consist of 15 hours of synchronous sessions and 15 hours of offline assignments. Teachers will develop general AI literacy by reading selected book chapters and watching documentary videos (Lee et al., 2022), gain technological pedagogical content knowledge (TPACK, Koehler & Mishra, 2009) required to implement the program, reinforce culturally relevant teaching strategies (Byrd, 2016), develop the identity of AI educators and a sense of belonging to the AI education community, and build relationships with the program implementation team to uphold implementation fidelity and collaborate on data collection.

3 PROJECT PERSONNEL

CC team**:** [redacted]**, Ph.D., PI**, is a learning scientist at the Concord Consortium. She leads multiple federally funded projects that develop and research learning technologies and learning experiences that engage youth in STEM learning. CC team consists of a project manager, a curriculum developer and PD specialist, and a full-stack technology development group.

TTU team**:** [redacted]**, Ph.D., Co-PI**, is an associate professor of computer science at TTU. He has been working on the foundations of AI and its applications for two decades and novel curricula integrating computing and STEM using AI in the last decade. [redacted]**, Ph.D., Co-PI**, is an associate professor of computer science at TTU. He is a machine learning expert. He won the test-of-time research award from ACM SIGKDD, a premier AI conference. TTU team also includes two graduate assistants with CS and educational research backgrounds.

UF team: [redacted]**, Ph.D., Co-PI**, is the Chief Equity Officer and Mathematics Principal at the Lastinger Center for Learning at UF. Her research examines math teachers' instruction for underserved students and English learners. She is a member of the board of directors for the National Council of Teachers of Mathematics (NCTM).[redacted]**, Ph.D., Co-PI**, is the informatics for education associate professor of educational technology at UF with a research focus on AI and learning analytics for online and STEM education. He has led and co-led multiple federally funded projects and published 60 articles with over 2,800 citations. The UF team also includes a graduate assistant with expertise in math education and virtual learning.

FLVS team: [redacted]**, ED.D.**, is the administrator of research at FLVS and has extensive experience in research and innovation in the e-learning industry. She will be in charge of the interfacing between the research team and the FLVS team on all aspects of the project, including curriculum development, technology integration, implementation, data access and collection, etc.

TTU K12 team: [redacted]**, M.Ed.**, is the superintendent of TTU K-12 and has 28 years of experience in education. He will connect the research team to his team for the implementation of the proposed program in TTU K-12 and data access and collection etc.

WestEd team: [redacted] **Ph.D.**, is a senior research associate at WestEd. She has more than 15 years of experience in educational research, evaluation, and consulting. [redacted] brings highly sophisticated design and analysis approaches to research and evaluation and is a WWC-certified reviewer on group design and single-case design. She will lead major evaluation activities related to the evaluation plan, research methodology, implementation fidelity, and impact data analysis. [redacted] will serve as the senior methodologist and will be responsible for data management and analysis. With a Ph.D. in measurement and applied statistics, [redacted] has led, designed, and implemented rigorous experimental trials as well as measurement/assessment projects funded by the U.S. Department of Education and NSF.

4 MANAGEMENT PLAN

4.1 Responsibilities and Collaboration

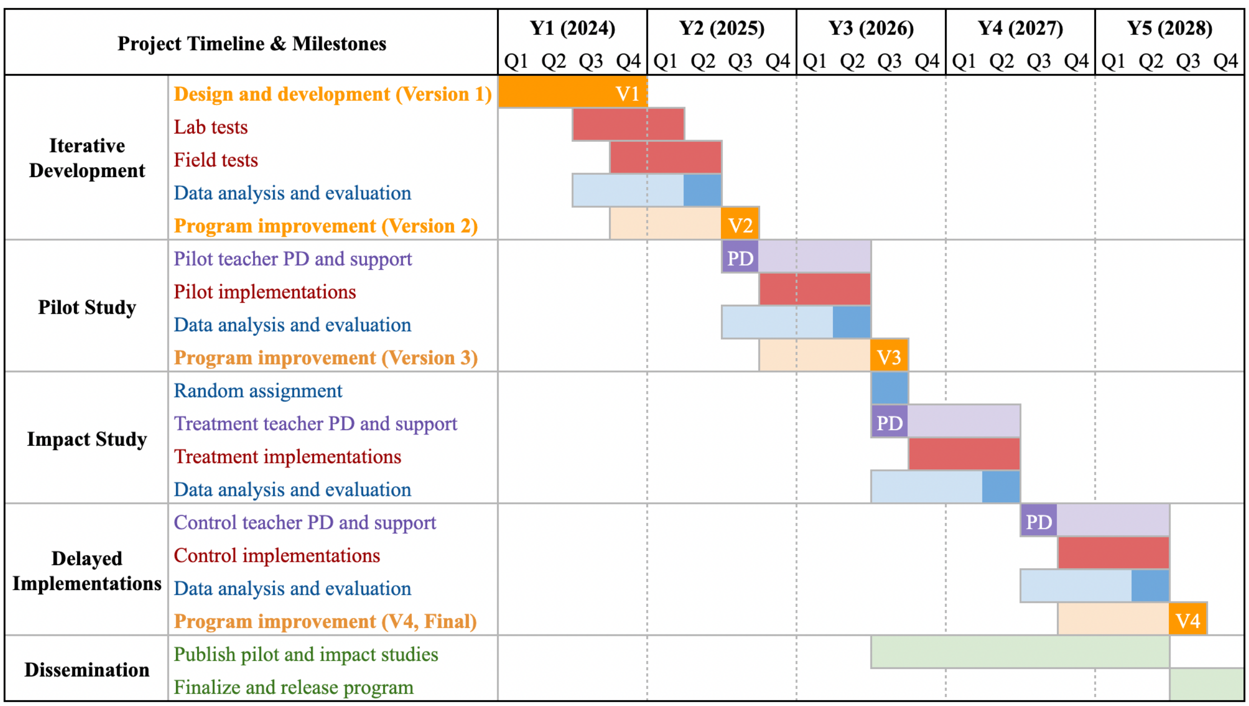
The CC team will be responsible for developing, testing, and improving the curriculum, assessment, technologies, and teacher PD program described in this proposal. The UF team will coordinate with FLVS to conduct the field tests, pilot study, and impact study described below. The TTU team will serve as the subject matter experts on this project and review all curriculum and assessment materials to ensure their content validity. The TTU team will also coordinate with TTU K-12 to conduct the field tests, pilot study, and impact study. CC, TTU, and UF teams will jointly implement the teacher PD program. The WestEd team, as the external evaluator, will periodically review project progress, provide formative feedback, and lead the analysis and publication of the impact study. The partners are long-time collaborators on multiple projects.

Over the years, we have established effective and efficient ways of communication and collaboration across multiple locations and time zones. In this project, we will use a combination of routine and need-based synchronous and asynchronous communications as well as standardized processes and documents to ensure seamless collaboration among the partners.

4.2 Project Timeline and Milestones

We propose to start the project on January 1st, 2024, and take five years to complete the proposed activities. Figure 1 below shows the project timeline and milestones.

Figure 1. Project timeline and milestones



Iterative Development. In Year 1, we will develop the curriculum materials and supporting technologies of the *AI in Math* program (version 1), conduct lab tests with student volunteers and field tests with teachers and their classes, collect and analyze observation, interview, and focus group data to inform program revision. In Year 2, we will develop and add teacher PD materials and research instruments to the program package (version 2) for use in the subsequent teacher PD and pilot study. **Pilot Study.** In the summer of 2025, we will train 8 teachers to pilot the program with 800 students in the 2025-26 academic year. Pilot data, including assessments, surveys, interviews, and focus groups, will be analyzed to inform program revision (version 3). **Impact Study**. In the 2026-27 academic year, we will conduct a randomized controlled trial (RCT) impact study with 60 teachers and 6000 students from the two virtual schools. The 30 teachers assigned to the treatment group will be trained in summer 2026 and receive continuous support from staff for their implementations (see details in section 5.1 Evaluation Methods). **Delayed Implementations**. In the 2027-28 academic year, we will train the 30 teachers previously assigned to the control group to implement the program. Implementation data from both 2026-27 and 2027-28 will be analyzed to inform the final revision (version 4, final).

4.3 Dissemination Plan

Curriculum and PD resources. We will host all curriculum materials and teacher PD materials on the Concord Consortium's STEM Resource Finder, a robust online learning platform visited by nearly half a million educators and learners, to ensure free and easy access and potential use beyond the funding period. We will optimize the project website for search engine queries and promote it through mathematics education and computer science education listservs and practitioner journals, including novel means such as joint promotion with these organizations. We will also reach out and offer demos to other virtual schools and encourage them to use our resources.

Newsletters and social media. Project news, research briefs, and infographics highlighting key findings will be broadcast to practitioners and policymakers via the CC newsletter, which currently reaches 63,000 subscribers, AI4K12's listserv, TTU K-12's newsletters, and UF's Kenneth Griffin CS Education for All Initiative reaching every CS teacher in Florida and beyond. We will send releases, updates, and invitations to CODAP users (20,000 active users per month at peak time), and CC's Data Science Education mailing list (1,100 members). We will actively and fully leverage social media channels (LinkedIn, Twitter, and Facebook), seeking out potential users and cultivating enthusiastic users to promote the site through their independent action, disseminate project information, and foster community.

Publications. We will publish in peer-reviewed journals and present at regional and national conferences geared at both practitioners and researchers. Potential journal targets include ACM Inroads, Computer Science Education, Educational Studies in Mathematics, Journal of Research on Educational Effectiveness, Educational Researcher, etc. Potential conferences include the annual meetings of the American Educational Research Association, Educational Advances in Artificial Intelligence, the Computer Science Teachers Association, the ACM Technical Symposium on Computer Science Education, and the National Council of Teachers of Mathematics. With a track record of invited talks and keynote speeches, the PI team will actively seek opportunities to disseminate the work of this project.

5 PROJECT EVALUATION

5.1 Evaluation Methods Designed to Meet WWC Standards Without Reservations

WestEd will conduct an independent evaluation to answer 7 research questions (RQs) that are aligned with the project logic model (see Appendix G), about the impact of *AI in Math* on students' AI self-efficacy, attitudes towards math, AI literacy, and math achievement (RQs 1–5); and the implementation of *AI in Math* (RQs 6–7). We propose to use a teacher-level randomized controlled trial (RCT) that will collect valid and reliable data on relevant outcomes so that our findings will meet the What Works Clearinghouse (WWC) standards without reservations.

Research questions (RQs)	Data sources and collection timeline
RQ1. What is the impact of *AI in Math* on students' AI self-efficacy?	Student survey on their AI self-efficacy Pretest: fall 2026; Outcome: spring 2027
RQ2. What is the impact of *AI in Math* on students' attitudes toward math?	Student survey on their attitudes toward math Pretest: fall 2026; Outcome: spring 2027
RQ3. What is the impact of *AI in Math* on students' AI literacy?	Student assessment of AI literacy Pretest: fall 2026; Outcome: spring 2027
RQ4. What is the impact of *AI in Math* on students' math achievement?	Algebra I End-of-Course assessments scores; Prior year's academic performance data; Student demographics; Teacher background survey
RQ5. To what extent is the impact of *AI in Math* on students' AI self-efficacy, attitudes towards math, AI literacy, and math achievement outcomes moderated by student and teacher characteristics and by implementation fidelity?	Same as RQs 1–5, RQ7
RQ6. To what extent is *AI in Math* implemented with fidelity?	Program records such as teacher attendance of training, teacher usage of *AI in Math* resources; student usage of *AI in Math* lessons; teacher implementation logs; teacher surveys (impact study year)
RQ7. What are the factors that hinder or facilitate the implementation of *AI in Math*?	Teacher surveys; teacher interviews; school leader interviews; student focus groups

Evaluation design. The evaluation will include a pilot study (the 2025-26 academic year) and a teacher-level RCT study (the 2026-27 academic year). The pilot study will include 8 Algebra I teachers and 800 students (100 students per teacher) and the evaluation will focus on collecting and analyzing initial implementation data to address RQs 6–7 and to inform refinement of *AI in Math* for the full implementation in 2026-27. The RCT study in 2026-27 will include 60 Algebra I teachers and 6000 students (100 students per teacher). Teachers will be randomly assigned to treatment or control within blocks (i.e., groups of teachers within a virtual school), with students in the same teachers' classrooms receiving the same assignment.

Teachers in both treatment and control conditions will participate in their schools' normal professional development requirements and opportunities, but teachers in the treatment condition also will participate in *AI in Math* training. Similarly, students in treatment classrooms will use *AI in Math* supplemental program, whereas students in control classrooms will continue their current Algebra practices (business-as-usual), including any use of additional resources, but not *AI in Math*. After the RCT study is completed, control teachers will also receive training and implement *AI in Math* with their students in the 2027-28 academic year. Student rosters will be collected three weeks prior to the start of the school year. Teachers will be randomly assigned into treatment and control groups two weeks prior to the start of the school year. Treatment teachers will participate in the training during the week before the start of the school year. Students who join the study after random assignment will be excluded from the evaluation sample. The team will track both overall and differential teacher- and student-level attrition from both conditions.

Outcome measures. WestEd will use student surveys and student assessments to measure outcomes. The evaluation will use measures that are directly related to the intended program outcomes as indicated in the logic model (Appendix G) but are not over-aligned with the intervention, and that meet WWC's validity and reliability requirements (WWC, 2022).

Student AI Self-Efficacy. We will measure student AI self-efficacy using the AI Self-Efficacy subscale of the AI literacy survey developed by Carolus et al. (2023). This 11-point Likert subscale (0-10) includes 6 items on students' self-efficacy in AI problem-solving and learning. Confirmatory factor analysis indicates that the 6 items on the AI Self-Efficacy subscale demonstrated a good internal consistency coefficient (.93-.97).

Student AI Literacy. We will measure student AI literacy using an objective assessment developed by Weber, et al. (2023). The assessment includes 16 items that measure socio user AI literacy, socio creator/evaluator AI literacy, technical user AI literacy, and technical creator/evaluator AI literacy. The instrument demonstrated acceptable agreement between allocated (participants) and theoretical (experts) groups (77.5%).

Student Math Attitude will be measured using the math subscale of High School Student Attitudes Toward STEM developed by North Carolina State University (Unfried, et al., 2015). This survey includes 8 five-point Likert scale items and is designed to measure high school student attitudes toward math (e.g., "Math has been my worst subject" and "I would consider choosing a career that uses math"). The internal consistency coefficient was .90.

Student Math Achievement. The **primary student math outcome measure** will include State Algebra I End-of-Course assessments scores from spring 2027. State Algebra I assessment also includes statistics standards due to course requirements in the state. Because the study involves assessments from two states, we will convert scaled scores to *z* scores separately for each state using the statewide means and standard deviations. We will also collect data on prior-year student achievement, attendance, and discipline, as well as student demographics (e.g., gender and ethnicity), which will be included as covariates in the impact analysis.

Power Analysis. We will recruit 60 Algebra I teachers from the two virtual schools. A power analysis was conducted using PowerUp! (Dong & Maynard, 2013) for teacher-level random assignment designs. Accounting for attrition, we conservatively estimated 54 teachers and 5400 students (100 students per teacher: 25 students per session, 4 sessions per teacher) from two virtual schools in the analytic sample. The proposed study has sufficient power to detect an effect size of 0.196 *SD* for student achievement (see Appendix J2 for power analyses).

Impact Analysis. We will use two-level hierarchical linear modeling (HLM) to estimate the intent-to-treat impact of *AI in Math* on student outcomes to accommodate the nested nature of the design (students nested in teachers' classrooms). Because the study involves assessments from two states, we will convert state assessment scaled scores to *z* scores separately for each state using the statewide means and standard deviations. Analyses will test the overall impact of *AI in Math* for Algebra I students, adjusting for a baseline measure, as well as student characteristics. HLM models also will be used for moderation analyses (RQ5) (see Appendix J2 for the data analysis plan).

Minimizing Attrition to Meet WWC Standards Without Reservations. We expect the study to meet WWC standards without reservations. The project team has established strong partnerships with participating virtual schools and will provide resources for virtual school staff to support teacher engagement to minimize teacher-level attrition (See letters of support from the virtual schools in Appendix C). Given that *AI in Math* is a 1-year program, teacher turnover and student mobility within the same school year are expected to be relatively low. The program implementation team will provide *AI in Math* to the control teachers after the implementation year to reduce teacher-level differential attrition. Since the analysis of student outcomes will rely on state End-of-Course assessment data, we will collect data for all students in the study sample. This ensures that the analysis of student math achievement outcomes will have low teacher- and student-level attrition and will be able to meet WWC standards without reservations.

5.2 Performance Feedback

WestEd will examine project implementation, assess project progress in achieving its goals, and provide iterative feedback for program improvement through frequent collection and analysis of both implementation and outcome data. We will regularly collect implementation fidelity data to document implementation quality for both pilot and impact phases (See Appendix J2 for implementation measures). Qualitative and quantitative data collected during both the pilot and impact phases of the evaluation will allow us to closely track the progress of program implementation and shed light on factors that may hinder or facilitate the implementation of *AI in Math* (RQs 6 and 7). These data will flow into feedback to the development team for program refinement and continuous improvement.

In addition to regular monitoring of implementation fidelity, we will conduct analyses at multiple time points, provide interim briefs with key findings (including available impact findings), and jointly interpret the evidence with the development team to support the continuous improvement processes. The findings will help the development team learn from successful approaches and identify common or localized problems of implementation that may need intervention. We will meet with the development team regularly to share progress and discuss challenges, define model implementation, and engage stakeholders in understanding and interpreting the findings.

5.3 Components, Mediators, and Outcomes & Measurable Threshold

The design of the evaluation is informed by the key components and student learning outcomes as illustrated in the logic model in Appendix G. The key component for students includes *AI in Math* supplemental certificate program. The program is designed to provide high-need students the opportunity to develop Artificial Intelligence (AI) literacy and self-efficacy in solving problems using AI and learning AI topics, and simultaneously improve their math learning and attitudes toward math (see section 2.3.1 Design AI in Math program for high-need students). The key component for teachers includes the use of online training and virtual support to help teachers learn about *AI in Math* and address their questions and concerns in a collaborative manner. We hypothesize if *AI in Math* is implemented with fidelity, students will increase their AI literacy, AI self-efficacy, attitudes toward math, and math achievement.

The fidelity thresholds for teachers include (1) attending 10 or more hours of synchronous PD sessions (a total of 15 hours); (2) reading the briefings or watching the video recordings of the missed synchronous PD sessions; (3) completing 90% of the professional learning assignments (a total of 15 hours); (4) meeting with staff members at least once to discuss implementation plan and research participation; and (5) implementing 8 or more *AI in Math* lessons (a total of 10 lessons). The threshold for students is completing at least 6 lessons in *AI in Math* program. The acceptable implementation by a class requires that 70% of students meet the acceptable implementation thresholds by the end of the implementation year. For the program to be considered to meet fidelity implementation expectations, we expect (1) at least 80% of the treatment teachers need to meet thresholds for teachers; and (2) at least 70% of the classes need to meet thresholds for students. WestEd will test and finalize the implementation matrix and the proposed thresholds.

6 REFERENCES

AI4GA.org (2023). *Living and working with Artificial Intelligence*. https://ai4ga.org/curriculum/

Archer, J., Eckel, R., Hawkins, J., Wang, J., Musslewhite, D., & Zhang, Y. (2023). *A study of students' learning of computing through a logic programming based integrated curriculum for middle schools.* Paper presented at the 13th AAAI Symposium on Educational Advances in Artificial Intelligence. Washington, D.C.

Barker, J., & Gossman, P. (2013). The learning impact of a virtual learning environment: students' views. *Teacher Education Advancement Network Journal (TEAN), 5*(2), 19-38.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bouchrika, I., Harrati, N., Wanick, V., & Wills, G. (2021). Exploring the impact of gamification on student engagement and involvement with e-learning systems. *Interactive Learning Environments, 29*(8), 1244-1257.

Byrd, C. M. (2016). Does culturally relevant teaching work? An examination from student perspectives. *Sage Open, 6*(3), 2158244016660744.

Carolus, A., Koch, M., Straka, S., Latoschik, M. E., & Wienrich, C. (2023). MAILS--Meta AI Literacy Scale: Development and Testing of an AI Literacy Questionnaire Based on Well-Founded Competency Models and Psychological Change-and Meta-Competencies. *arXiv preprint arXiv:2302.09319*.

Chao, J., Ellis, R., Jiang, S., Rosé, C., Finzer, W., Tatar, C., Fiacco, J., & Wiedemann, K., (2023, Feb). *Exploring Artificial Intelligence in English Language Arts with StoryQ* [Paper Presentation]. EAAI-23: The 13th Symposium on Educational Advances in Artificial Intelligence. Washington, D.C., USA.

Code.org. (2023). *AI and machine learning (‘23-’24)*. https://studio.code.org/s/aiml-2023

Code.org, CSTA, & ECEP Alliance (2022). *2022 State of Computer Science Education: Understanding Our National Imperative.* Retrieved from https://advocacy.code.org/stateofcs

Curtis, H., & Werth, L. (2015). Fostering student success and engagement in a K-12 online school. *Journal of Online Learning Research, 1*(2), 163-190.

Dong, N., & Maynard, R. (2013). PowerUp!: A tool for calculating minimum detectable effect sizes and minimum required sample sizes for experimental and quasi-experimental design studies. *Journal of Research on Educational Effectiveness, 6*(1), 24-67.

Durik, A. M., Vida, M., & Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: A developmental analysis. *Journal of Educational Psychology*, 98, 382– 393. https://doi.org/10.1037/0022-0663.98.2.382

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*, 103– 127. https://doi.org/10.1037/a0018053

Exploring Computer Science. (2022). *Artificial Intelligence alternative curriculum unit*. https://www.exploringcs.org/for-teachers-districts/artificial-intelligence

Huerta, L., Shafer, S. R., Barbour, M. K., Miron, G., & Gulosino, C. (2015). Virtual schools in the US 2015: Politics, performance, policy, and research evidence. *National Education Policy Center*.

Jackson, P. (1998). *Introduction to Expert Systems* (3rd ed.). Harlow, England: Addison-Wesley.

Jiang, S., Tang, H., Tatar, C., Rosé, C. P., & Chao, J. (2023). High school students’ data modeling practices and processes: From modeling unstructured data to evaluating automated decisions. *Learning, Media and Technology*, 1-19. https://doi.org/10.1080/17439884.2023.2189735

Long, D., & Magerko, B. (2020, April). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-16).

Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)?. *Contemporary Issues in Technology and Teacher Education, 9*(1), 60-70.

Lavalley, M. (2018). Out of the loop: rural schools are largely left out of research and policy discussions, exacerbating poverty, inequity, and isolation. *Center for Public Education*.

Lee, I., Zhang, H., Moore, K., Zhou, X., Perret, B., Cheng, Y., ... & Pu, G. (2022, February). AI Book Club: An Innovative Professional Development Model for AI Education. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education-Volume 1* (pp. 202-208).

Lent, R. W., & Brown, S. D. (2013). Social cognitive model of career self-management: toward a unifying view of adaptive career behavior across the life span. *Journal of counseling psychology, 60*(4), 557.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance [Monograph]. *Journal of Vocational Behavior, 45*, 79 –122. doi: 10.1006/jvbe.1994.1027

Lent, R. W., Ireland, G. W., Penn, L. T., Morris, T. R., & Sappington, R. (2017). Sources of self-efficacy and outcome expectations for career exploration and decision-making: A test of the social cognitive model of career self-management. *Journal of vocational behavior, 99*, 107-117.

Logan, S., & Johnston, R. (2009). Gender differences in reading ability and attitudes: Examining where these differences lie. *Journal of Research in Reading, 32*, 199– 214. https://doi.org/10.1111/j.1467-9817.2008.01389.x

Marinak, B. A., & Gambrell, L. B. (2010). Reading motivation: Exploring the elementary gender gap. *Literacy Research and Instruction, 49*, 129– 141. https://doi.org/10.1080/193880709028037

MIT RAISE (2023). *Day of AI: Empower your students with the knowledge and skills to flourish in a world with AI*. https://www.dayofai.org/

Molnar, A. (Ed.), Miron, G., Hagle, S., Gulosino, C., Mann, B., Huerta, L.A., Rice, J.K., Glover, A., & Bill, K. (2023). *Virtual schools in the U.S. 2023.* Boulder, CO: National Education Policy Center. Retrieved July 27th, 2023 from http://nepc.colorado.edu/publication/virtual-schools-annual-2023

OECD. (2013). Mathematics self-beliefs and participation in mathematics-related activities. In *OECD, PISA 2012 results: Ready to learn* (Vol. III, pp. 87– 112). OECD. https://doi.org/10.1787/9789264201170-8-en

OECD. (2019). *PISA 2018 results (Volume II): Where all students can succeed.* OECD. https://doi.org/10.1787/b5fd1b8f-en

Pajares, F. (2005). Gender differences in mathematics self-efficacy beliefs. In A. M. Gallagher, & J. C. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 294– 315). Cambridge University Press. https://doi.org/10.18848/2327-7920/CGP/v23i03/1-9

Schleicher, A. (2019). *PISA 2018: Insights and interpretations.* Organisation for Economic Co-operation and Development.

Tatar, C., Culbret, D., Jiang, S., Rosé, C., Chao, J., Ellis, R., Wiedemann, K., and Jiang, S. (2022). *High school students' sense-making of Artificial Intelligence and Machine Learning.* In Koli Calling '22: 22nd Koli Calling International Conference on Computing Education Research (Koli 2022), November 17–20, 2022, Koli, Finland. ACM, New York, NY, USA. https://doi.org/10.1145/3564721.3565958

Tatar, C. , McClure, J., Bickel, F., Ellis, R., Wiedemann, K., Chao, J., Jiang, S., & Rosé, C. (2023). Examining high school students' self-efficacy in machine learning practices. *Proceedings of the Annual Meeting of the International Society of the Learning Sciences-ICLS 2023.* Montréal, Canada.

Touretzky, D. S., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019, June). K-12 guidelines for artificial intelligence: what students should know. In *Proc. of the ISTE Conference* (Vol. 53).

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.

Vlachopoulos, D., & Makri, A. (2017). The effect of games and simulations on higher education: a systematic literature review. *International Journal of Educational Technology in Higher Education, 14*(1), 1-33.

Wang, J., Zhang, Y., Guo, Y., Hawkins, J., & Romero M. (2023). *Examine a workshop for STEM teachers about the integration of computational thinking and science instruction*. Manuscript submitted for publication.

Weber, P., Pinski, M., & Baum, L. (2023). Toward an objective measurement of AI literacy. *PACIS 2023 Proceedings*. 60. https://aisel.aisnet.org/pacis2023/60

What Works Clearinghouse. (2022). *The WWC procedures and standards handbook (Version 5.0).* U.S. Department of Education. https://ies.ed.gov/ncee/wwc/handbooks

Wigfield, A., Tonks, S., & Klauda, S. L. (2009). Expectancy-value theory. *Handbook of motivation at school, 2,* 55-74.

Zhang, Y., Wang, J., Bolduc, F., & Murray, W. G. (2019a, May). LP based integration of computing and science education in middle schools. In *Proceedings of the ACM Conference on Global Computing Education* (pp. 44-50).

Zhang, Y., Wang, J., Bolduc, F., Murray, W. G., & Staffen, W. (2019b, July). A preliminary report of integrating science and computing teaching using logic programming. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, No. 01, pp. 9737-9744).