Knowledge Acquisition and Transformation Expansion

Texas A&M Research Foundation – Literacy.IO Platform
Evaluation Team – WestEd & Analytica Insights Inc.
Technology Infrastructure – Problem Solutions

Contents

Knowledge Acquisition and Transformation Expansion

Knowledge acquisition and transformation expansion (KATE) is an Education Innovation Research (EIR) Expansion Phase grant focused on reading comprehension and submitted by the Literacy.IO Research group at the Texas A&M Foundation. We will partner with external evaluators WestEd and Analytica Insights Inc. Problem Solutions Inc. will support the comprehension-focused technology platform. The core elements and goals of KATE include:

- Improve reading comprehension for students in grades 4 and 5 using the web-based intelligent tutoring system for the structure strategy (ITSS), available in English and Spanish. ITSS research has shown positive effects for students' comprehension outcomes and been reviewed by the What Works Clearinghouse and met evidence standards without reservations (WWC, 2020). In the recent WWC review focused on distance learning, Sahni et al. (2021) recommended ITSS based on the positive and statistically significant effects for reading comprehension. In addition, Boegards-Hazinger et al. (2020) and others have conducted meta-analyses that also report ITSS's positive impact on reading.
- Support teacher fidelity of ITSS implementation with strong practice-based professional development (PD) using web-based massive open online virtual learning and a multi-dimensional, "360-degree" support system (MOOV 360), designed, developed, and tested through a Supporting Effective Educator Development (SEED) grant (U423A180074).
- Enhance best teaching practices for comprehension (e.g., recommendations from the WWC Practice Guide for Adolescent Literacy) through the MOOV 360 platform~~,~~ including (a) searchable teacher library of resources (e.g., lesson plans, model videos, Nearpod assessment activities, text structure games), (b) coaching and modeling, (c) a curricula integration

framework, (d) tools for professional learning communities (e.g., planning assistance, virtual office hours), (e) monthly check-in, and (f) fidelity of implementation checklists.

The goal of KATE is to examine the effectiveness of ITSS + MOOV 360 with an economically and culturally diverse population of students in grades 4 and 5 from multiple geographic regions of the U.S. A large-scale cluster randomized controlled trial (RCT) will be conducted with 64 elementary schools. KATE will be a replication of the WWC-reviewed findings in grades 4 and 5 (WWC, 2020), with a scale up of school and districts over a wider geographical region.

KATE addresses both absolute priorities and both invitational priorities of EIR.

Absolute Priority 1—***Strong Evidence—***ITSS has strong evidence with positive and statistically significant effects on reading comprehension for students at grades 4, 5, and 7. There is no contradictory evidence of negative effects. The strong evidence is based on review of three [redacted]-led RCTs reviewed in the WWC's ITSS Intervention Report (Wijekumar et al., 2012, 2014, 2017; WWC ITSS Intervention Report, 2021). These studies were conducted in two states with suburban and rural school districts.

Absolute Priority 2—*Field-Initiated Innovations*—KATE will take to scale ITSS + the MOOV 360 PD platform for disseminating and enhancing best teaching practices for comprehension. These entrepreneurial, evidence-based, field-initiated technology-based innovations can improve student achievement in reading, at scale.

Invitational Priority 1—*Innovative Approaches to Addressing COVID–19 Impact*—

KATE delivers evidence-based reading comprehension strategies using easily accessible and adaptive, web-based tools. ITSS and MOOV 360 already made an impact on underserved students during the COVID-19 pandemic. Given what we know about challenges that lie ahead

as students return to school and need to accelerate learning, the flexible, individually-adaptive intelligent tutoring, along with English and Spanish lessons, provide a strong platform that can be used in a variety of settings. During COVID-19, the literacy.IO team successfully implemented the ITSS + MOOV 360 in 28 economically disadvantaged schools in Texas and New Mexico. The team pivoted from classroom-based delivery to a hybrid model: at-home parent-supported delivery of intervention with ZOOM+Nearpod teacher delivery of instruction. Teachers and families downloaded over 250,000 resources from the MOOV and ITSS library. Findings demonstrated positive, meaningful, and statistically significant effects for teacher knowledge and student outcomes of underserved populations (Wijekumar et al., 2021a).

Invitational Priority 2—*Promoting Equity and Adequacy in Student Access*—KATE facilitates equitable access to educational resources and opportunities. The web-based tools promote high quality instruction among geographically, socioeconomically, and culturally diverse students and teachers. ITSS and MOOV 360 contains English and Spanish resources. With direct services to students, along with building pedagogical knowledge among teachers, KATE improves equity and adequacy of resources with particular emphasis on supporting underserved students.

A. *Significance*

A.1. National significance of the KATE project.

Literacy is the foundation of all human activity. Effective reading promotes academic achievement, professional prowess, health and wellness, and civic engagement (Marshall, 2013). Unfortunately, over 33% of elementary grade children fail to succeed in this fundamental skill, as observed in national and state assessments (NAEP, 2019; TX STAAR, 2018). Children from underserved groups and economically distressed communities are at a higher risk for failure. Literacy challenges in the elementary grades continue to impact children throughout their

schooling and beyond. Lyon has described illiteracy in the U.S. as a national health crisis, and numerous researchers have sounded the alarm bells about the long-term negative impacts that result from inequities in reading competencies in early grades (Lyon, 2000; Reardon, 2011, 2013).

Long before COVID-19 wreaked havoc on the world, children were experiencing the pandemics of poverty and prejudice (Ladson-Billings, 2021). Education, and particularly literacy instruction, is one of the best neutralizers of these inequities; schools have a unique opportunity to provide high-quality learning opportunities to all children. Addressing these challenges is the responsibility of teachers, school leaders, and researchers who can implement evidence-based practices that have the power to improve students' literacy skills. In particular, as students reach the upper elementary grades, reading demands increase as students read in order to learn in content areas, which present ideas and content in complex text structures (Kamil et al, 2008). Students must learn from different kinds of text and synthesize ideas and diverse perspectives presented across texts. Recently, Boegards-Hazenberg et al. (2020) synthesized influential theoretical models of reading (e.g. construction integration, landscape, reading systems framework) and practical instructional strategies. Among these, they identified (a) vocabulary knowledge, (b) selection and encoding of main ideas, (c) summarization, and (d) inferencing as key skills to improving reading comprehension. Main ideas and summarization are challenging to students at upper elementary grades (Wijekumar et al., 2020). Accumulating evidence by the What Works Clearinghouse showcases the impact of the reading comprehension strategies presented in ITSS and how they can improve reading outcomes in elementary grade children (Hudson et al., 2021; Sahni et al., 2021; WWC ITSS Intervention Report 2021). ITSS helps students by guiding the selection of important ideas based on the top-level structure sentence

stems. Students learn to classify the top-level text structure by identifying signal words in the passage (e.g., due to, because, consequently signal cause and effect). For example, when reading about the World War I, students understand the causes and effects of the war by using the cause and effect top-level structure and the associated sentence stem: the causes are ___, ____, and ___ and the effect is _____. The students then learn to encode the important ideas using the cause and effect logic and generate a main idea using the pattern. Students also learn how to extend the main idea with supporting details and create a summary of the text. Finally, students learn how to extrapolate inferences based on the causes and effects. The main idea, summary, and inferences appear regularly in all reading assessments. These constructs are also staples in all state standards and reading textbooks. Unfortunately, teachers struggle with teaching students to generate main ideas, summaries, and inferences (Wijekumar et al., 2020).

A.2. KATE is likely to increase knowledge and understanding of problems and issues related to reading comprehension and effective strategies for addressing them.

KATE will generate impact estimates of providing ITSS + MOOV 360 with grades 4 and 5 students and teachers. Strategic use of technology is a key strategy to scale up ITSS as a cost-effective literacy solution for upper-elementary reading comprehension. The ITSS platform with modeling, activities, assessment, feedback, scaffolding, and intelligent adaptive lesson pathways and English and Spanish lessons responds to the variety of student needs (e.g., reading ability, vocabulary knowledge, fluency). The technology tools present consistent evidence-based instruction to students regardless of the school circumstances (e.g., high teacher turnover). The project also uses technology to significantly reduce costs associated with teacher PD (by providing virtual coaching and support), to reduce curricular contradictions (e.g., when textbooks contradict ITSS instruction with non evidence-based comprehension strategies), and to administer pre- and post-tests. The tools have the potential to overcome any pandemic related

deficits in reading. Data collected from multiple geographic regions and underrepresented communities will inform educators, policy makers and other stakeholders about the:

1. Effectiveness of the ITSS + MOOV 360 to improve 4th and 5th grade students' reading comprehension on a general reading comprehension test and specific comprehension skills related to signaling words, main idea quality and top-level text structure.
2. Variations in effectiveness of ITSS + MOOV 360 by gender, socioeconomic status, English learner (EL) status, and prior reading comprehension skills.
3. Changes in teacher content and pedagogical knowledge about comprehension instruction – prior to and after one year of the intervention.
4. Correlations between the teacher content and pedagogical knowledge and student reading comprehension outcomes.
5. Implementation barriers and solutions (e.g., best practices in using ITSS student data; how to leverage MOOV 360 coaching supports at scale; solutions if textbooks contradict ITSS strategies – see Wijekumar et al., 2021a).

B. *Strategy to Scale*

B.1. KATE strategies that promote scaling up ITSS + MOOV 360

The KATE project will build on implementation approaches refined by the Literacy.IO group through multiple IES grant efficacy projects (R305A080704, R305A120593, R305A080133, R305A150057, R305A180060) and, most recently, during the SEED–MOOV grant, in which program implementation and the RCT were launched, just as COVID-19 began.

1. *ITSS is an easily scalable web-based solution to address reading comprehension needs of students regardless of language, race, or socioeconomic status.* ITSS was designed to improve reading comprehension by promoting the selection and encoding of important

ideas when reading. The instruction focuses on showing students how to identify signal words in text, classify text structures, utilize text structure main idea patterns to select important ideas (e.g., the causes are ___and the effect is ____), generate a main idea, extend the main idea to generate a summary by adding details, and finally extrapolating inferences about the reading. Appendix J: Figures J1-J2 present examples of the student interface and sample classroom posters. An animated pedagogical agent name I.T. presents modeling, practice tasks, assessment, immediate feedback, and scaffolding. Students practice the strategies in a variety of grade-appropriate passages in multiple domains (e.g., science, social studies), genres (e.g., expository, narrative, poetry, biography), and culturally diverse readings. The ITSS provides adaptive tutoring lessons based on student mastery of the lesson content; headphones allow students to hear the narrations and instructions. Students use ITSS for 60 minutes a week using an assigned username and password.

2. *Teacher PD is scalable with MOOV 360, a powerful web-based teacher professional system*. Teachers are an important part of the intervention and receive an initial 24 hours of training (practice-based PD) to learn the theoretical foundations of the text structure strategy and how they are applied in the classroom. The PD provides opportunities for teachers to plan, teach, reflect, and receive feedback from expert tutors. Teachers also learn how to integrate existing textbooks with the ITSS approach to generating main ideas, summaries, and inferences (see Appendix J Figures J4-J7, e.g., Hudson et al., 2021; Wijekumar et al., 2021a).

   Even before the pandemic, the literacy-IO team utilized the SEED-MOOV grant to produce a platform to deliver asynchronous, synchronous, and hybrid web-based PD that

promotes strong teacher knowledge about reading comprehension and cultivates applied reflective teacher practices. The system gained popularity during COVID-19 with the powerful suite of web-based tools that can easily adapt to the needs of the schools. MOOV 360 contains a library of thousands of resources. The system has 160 lesson guides to match textbook materials (e.g., HMH), 200 lesson guides for passages from expository, biography, narrative, and poetry from NEWSELA, 36 text structure games, 400+ Powerpoint with Nearpod lessons, and 280 videos with interactive activities.

3. *Teacher resources are scalable with an open-source, continuously-updating system that is crowd-sourced with teachers contributing lesson plans, videos, and games*. Utilizing this platform for KATE addresses the costs of delivering PD and the need for just-in-time, easily-accessible training resources for teachers. Once the teachers are trained, the system supports their on-going PD and engagement through weekly coaching and monthly catch-up sessions, along with a newsletter and updates to the resource library. Any teacher contributions go through expert review to maintain high quality resources.
4. *KATE uses an extant curricular integration framework that promotes alignment of school curricula and ITSS strategies.* Another barrier to scaling up were the contradictions presented in the curricula used by schools (e.g., Beerwinkle et al., 2018). In previous ITSS studies, we learned that classroom instruction frequently contradicts the ITSS instructions for main idea, summarization, and inferencing. We also learned that over 60% of teachers were unaware of evidence-based strategies for reading comprehension. The Literacy.IO team has now resolved this challenge with the Knowledge Acquisition and Transformation (KAT) guide (Appendix J: Figure J8), embedded in MOOV 360, that helps teachers integrate the text structure strategy into the schools' adopted curricula.

MOOV contains hundreds of resources in its library that direct align (or apply to) popular textbooks. During the teacher PD, we train teachers to use their own curricula materials to deliver instruction consistent with the ITSS delivered strategies. As an example, Keller Elementary School in Brownsville, TX started using ITSS with KAT in 2016 and has achieved 100% pass rates on the TX STAAR assessments for all grade levels in 2017, 2018, and 2019. These outstanding results were also achieved for special education students (Green, 2019). We will use the lessons learned from the KAT guide and related training and apply it to the EIR expansion.

5. *ITSS + MOOV 360 allow scaling of resources to the increasing number of Spanish speaking ELs in the US and their teachers.* The need for resources for Spanish speaking ELs has been a barrier to scaling up and our solution included developing ITSS and all teacher materials and PD in Spanish, in support of the large numbers of Spanish-speaking ELs in our schools. Linguistic supports are available at the word, sentence, and passage level (See Appendix J: Figures J3-J4). During the current SEED-MOOV project, 24 schools successfully used the resources to improve reading comprehension of ELs (Wijekumar et al. 2021a). The system has also produced significant and strong results for ELs as they learn English (Wijekumar et al., 2018).

B.2. The KATE dissemination approach promotes further development and replication.

The KATE dissemination plan is tailored to the practitioners, administrators, school leaders, policymakers, and the research community so that they may benefit from the findings from this project and spur further development and replication. Some dissemination tools have already been established through the SEED grant and will be further refined for KATE. The partnership with the WestEd allows the team to further extend dissemination. WestEd regularly uses its website and social media to share agency work and draws thousands of followers across

Twitter, Facebook, and LinkedIn. WestEd regularly disseminates research-to-practice information, as lead on the Regional Education Laboratory - West since its inception, and in partner roles o[redacted]. WestEd plays a key role on three regional Comprehensive Centers. To ensure research findings are relevant and actionable, WestEd develops tools to help educators use evidence-based practices. For example, REL West created Evidence-Based Improvement: A Guide for States to Strengthen Their Frameworks and Supports Aligned to the Evidence Requirements of ESSA.

Through these avenues, WestEd is accustomed to giving webinars and presentations that offer the opportunity to engage large audiences (and ensuring they are 508 compliant and readily accessible via WestEd's website for viewing after the event). In addition to webinars, we will seek out opportunities to present at local and national conferences and meetings of educators and policymakers. To maximize reach and attendance, we will propose symposia and panel presentations that include key stakeholders and practitioners from the KATE project.

Closed captioning and other ADA compliance is also available on the ITSS and MOOV 360, allowing broad access to materials. If, after meeting WWC standards, the results are meaningful and statistically significant, the research team plans to disseminate findings using 12 approaches customized to the needs of each group. Next we present the TAMUF approaches with specific customization for audiences that may benefit from this research.

Textbook for ELA produced by TAMUF. The text structure-based reading comprehension approach that is implemented in the ITSS and MOOV platform is being converted to textbooks for grades 3 to 5. The text will utilize links to the extensive resource library of the ITSS & MOOV.[redacted]

[redacted]. ***ITSS & MOOV Library.*** We will advertise the open-source resources library through literacy organizations such as the International Literacy Association, Texas Association of Literacy Educators, and Society for the Scientific Study of Reading to share the resources for practitioners and researchers. -***Webinars-First Thursday Catch-Up.*** Initiated in 2019, we will continue to hold monthly webinars-First Thursday Catch-Up sessions for all users (English and Spanish with closed captioning for hearing impaired). We will notify all participating *school teachers, administrators, and families* through school district newsletters, TAMU College of Education newsletters and listservs, and regional educational agency email listservs to advertise the webinars. During these sessions, we present descriptions of the intervention, research approach, and results. We also respond to questions from the participants and showcase exemplary teacher modeling and resources. These sessions are accompanied by newsletters for the attendees. The following highlight some of the recent events: **December Newsletter Link and January Newsletter Link**. ***Social Media and YouTube.*** To disseminate findings as we did with the previous results from the ITSS development and efficacy findings, social media and YouTube will be used. We will develop and post video clips that are suitable *for students, families, at-home and classroom use* like the videos: Coronavirus Poem Lesson (English link and Spanish link)**.** ***Publications, presentations at regional and national level by TAMUF.*** We plan to present findings at conferences where *school administrators and teachers* gather to hear about new developments. We will also present findings at AERA, SITE, and SSSR conferences. We will submit manuscripts for publication to high impact journals to share our findings with the research community *(See Hudson et al., 2021)*. ***Meetings with policymakers and legislative advisory groups.*** The Texas A&M University System of schools regularly meets

with Texas State Legislative groups and congressional representatives. We plan to present whitepapers to these groups noting the important contributions being made by this research.

C. *Quality of the KATE Project Design*

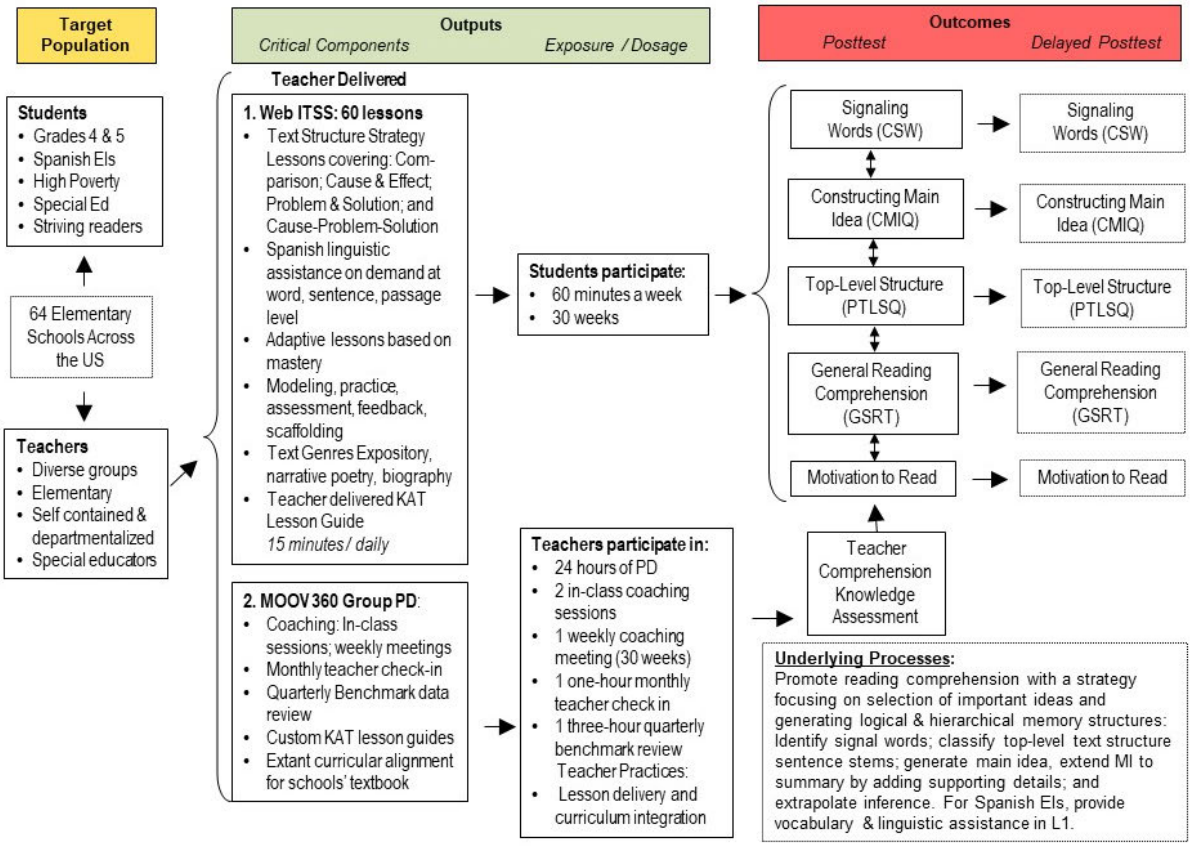
C.1. The KATE conceptual framework and supporting evidence

Skilled readers select important ideas while reading and generate a coherent mental representation of the text (Bogaerds-Hazenberg et al., 2020; Pearson & Cervetti, 2015 Shanahan, 2020). Multiple theories of reading explain how readers consume information from texts at a surface, text base, and situation-based level (i.e., Construction-Integration by Kintsch 2005; Kintsch & van Dijk, 1978) or through the activation and updating of ideas (i.e., Landscape Model by Van den Broek et al., 1999; 2005). The Text Structure Theory and its developments (Meyer 1975; Meyer & Poon, 2001; Wijekumar et al., 2012; Wijekumar et al., 2020b) made these connections explicit. First, the Text Structure theory suggests that the ideas that are most *relevant* and should be *intentionally activated* are those pertinent to the top-level structure (TLS) of a text (Meyer & Poon, 2001). The TLS promotes the generation of a logical main idea (Meyer, 1975; Wijekumar et al., 2021b).

Meyer and Poon (2001) identified problem and solution, cause and effect, and comparison as higher-order structures than description and sequence, as they prompt students to engage in constructing deeper level, more meaningful schemata than simply listing facts (description structure) or sequencing events (sequence structure). Wijekumar et al., (2020) suggested that a cause, problem, and solution structure is the fourth higher-order structure, as it differentiates problem and solution texts with and without a cause. The recent RCTs on ITSS by [redacted] and colleagues (e.g., Wijekumar et al., 2012; Wijekumar et al., 2014; Wijekumar et al., 2017) provide accumulating evidence that applying these higher-order structure strategies in

classrooms result in positive reading comprehension outcomes on standardized tests and researcher designed measures of main idea competence and TLS knowledge.

Figure 1. Logic Model



This conceptual framework is implemented through web-based ITSS lessons and a critical teacher delivered component of KATE as shown in the Figure 1 Logic Model. Students will receive 60 lessons throughout 30 weeks (for 60 minutes a week) of the school year with Spanish linguistic assistance on demand at word, sentence, and passage level. These adaptative lessons show students how to identify the top-level text structure, generate main ideas using specific sentence stems (e.g., comparison pattern ____ and ___ were compared on ____, ____ and ___), extend the main idea to form a summary, and extrapolate inferences. KATE teachers will participate in MOOV 360 PD, a second critical component of the intervention (Figure 1).

The same instructional features of ITSS lessons that student are expected to master are incorporated in to the MOOV 360 PD. By participating in the PD, teachers learn how to use the textbook passages to teach students the text structure strategy. This curricula integration further enhances student assimilation of the reading strategy to improve the reading comprehension outcomes depicted in the logic model. As mentioned earlier, and relevant to the logic model, the ITSS lessons have been shown in previous research to improve constructs of reading comprehension. The scale-up will also test whether these effects are sustained for fourth-grade students after six months through a delayed posttest (Figure 1).

C.2. The KATE goals, objectives, and outcomes are clearly specified and measurable.

The three major goals of the KATE project are to a) improve reading comprehension with 4th and 5th grade students across 64 participating schools, b) improve teacher practice about reading comprehension, and c) widely disseminate lessons learned about student comprehension, teaching, web-based instruction for students, and web-based PD for teachers,. Table 1 presents the highlights of the timeline and activities designed to achieve these goals.

C.3. KATE design is appropriate to, and will successfully address, the needs of the target population needs

The Literacy.IO team at TAMUF has devoted the past 20 years to developing and refining all the necessary resources to address the reading comprehension challenges facing a multitude of children across the world. RCTs conducted with fourth-grade students in rural and suburban schools showed statistically significant and positive effects favoring the ITSS classrooms over the control groups (Wijekumar et al., 2012). These results have met the WWC standards without reservations (WWC, 2020). The studies were also conducted with 5th grade and 7th grade students showing similar patterns of outcomes favoring the ITSS groups. Throughout these projects, the team has continued to develop resources and refine the PD to address the numerous

implementation science challenges to working in the school context (e.g., Beerwinkle et al., 2018; Hudson et al., 2021; Wijekumar et al., 2020; Wijekumar et al., 2021). The ITSS software was also adapted for Spanish speaking ELs and showed stronger effects favoring the intervention group (Wijekumar et al., 2018).

Through these refinements and deep understanding of the school context the project will address 4th and 5th grade student needs. The software has been proven usable and feasible. Teacher PD has been systematically developed and refined as a cost-effective and powerful platform for delivering the PD and then fully supporting the teachers as they implement the strategies in their classrooms. The ITSS teacher dashboard provides reports to inform the students on student progress. Teachers receive a Guidebook with all the resources and sample practice passages for use in the practice based PD sessions. The book also includes the classroom posters and samples of classroom resources.

By combining these student and teacher web-based resources that have been vetted in multiple states, we anticipate the KATE project will overcome many barriers to scaling up this evidence-based intervention.

D. ***Adequacy of Resources and Quality of the Management Plan***

D.1. TAMU Foundation, WestEd, Analytica Insights, and Problem Solutions capacity to bring KATE to national scale.

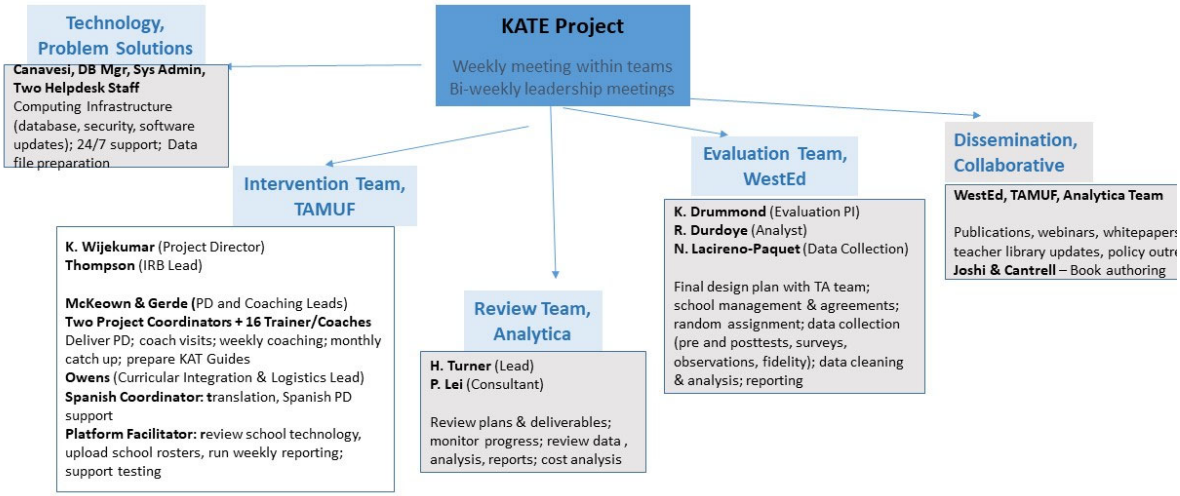
This proposal is being submitted by Texas A&M Research Foundation (TAMRF) with Evaluators WestEd and Analytica Insights Inc. supported by Problem Solutions Inc. for the intervention platform. The combined resources of all entities provide all necessary support to accomplish the goals of this project. TAMUF has extensive resources that currently hosts and supports the ITSS & MOOV 360 platforms. TAMUF has a state of the art library system,

multiple meeting rooms, an in house hotel with meeting spaces, numerous regional sites, technology infrastructure, and research support services. The endowed CUSP at TAMU provides a strong foundation of resources for this project with the necessary technology infrastructure, office space, collaboration tools, and established strong relationships with the participating schools. The College of Education and Human Development provides a strong support structure of post-award services including budget management, purchasing support, and outreach activities. The College has a dedicated staff to support the outreach activities for this project. Through a recent EDGES Fellowship to PI Wijekumar, TAMU has provided three post-doctoral researchers and a curriculum writer for the research activities supporting this project.

D.2. KATE management plan to achieve the objectives on time and within budget.

Responsibilities of the team are presented in Figure 2 and timelines and milestones are presented in Table 1.

Figure 2. **Organizational structure and responsibilities for accomplishing project tasks**



TAMUF Team – ITSS & MOOV 360

[redacted], PhD, **Project Director** is Houston Foundation Endowed Chair of Teaching, Learning, and Culture (TLAC), Chancellor's EDGES Fellow, and Director of the

endowed Center for Urban School Partnerships at TAMU. She has served as the PI for two IES Goal 2 Development Grants (SWELL & We-Write) that successfully transitioned to four Goal 3 Efficacy grants (ITSS Rural & Suburban, ITSS High Poverty, SWELL, and We-Write). The ITSS project was recently cited by US News and World Report and the WWC as an evidence-based practice. In all the projects, she has managed geographically dispersed teams in successfully completing the proposal, budget, research design, school recruitment, logistics for data collection, data analyses, and dissemination. She is the architect of the MOOV 360, ITSS, SWELL, and We-Write technologies and served as the schools' contact for all studies. CRTs lead by [redacted] have been reviewed by the What Works Clearinghouse and deemed to meet evidence standards without reservations. [redacted] serves on the editorial boards of *Journal of Educational Psychology, Contemporary Educational Psychology*, *Review of Educational Research*, *Reading and Writing: An Interdisciplinary Journal,* and *Educational Technology Research and Development*. [redacted] has developed extensive resources for dissemination through the CUSP including Teacher Summit conferences, webinars, MOOV, and social media outlet distribution of videos. She is also a member of the Texas Association of School Administrators and a regular presenter at their events. [redacted] will manage the full project in collaboration with a highly productive group of collaborators.

[redacted]**, PhD,** is an associate professor at TAMU TLAC. [redacted] is an expert in literacy and teacher professional development (PD). She has been designing both traditional and digital curriculum materials for over 20 years. She is a leader in creating effective and efficient professional learning experiences for practicing teachers and will lead this effort for the ACE project. She currently serves as the Co-PI on the SEED MOOV grant where she has developed and delivered PD for teachers on the reading strategies. She also has ten years of

experience teaching in the US and around the world, lending credibility to her presentations with teachers.

[redacted], **PhD,** is an associate professor at TAMU TLAC. [redacted] research focuses on children's literacy development, effective instruction to promote early literacy among young children living in poverty, and design of effective PD for early educators who are marginalized from university education. She uses a research-practice partnership model to collaborate with families and a range of early childhood programs (Head Start, Tribal Head Start, child care, state funded pre-K) serving diverse educators, and children. Her community-based work has been published in high impact early childhood and language and literacy journals and widely distributed practitioner journals (e.g., *The Reading Teacher*). She currently serves as PI on the IES-funded *iWRITE* project (partnership with Head Starts in Flint, MI and Atlanta, GA) and with state-funded preschool programs in Detroit, MI, and recently led partnership building and teacher coaching on an NSF-funded study with urban and rural Head Starts across Michigan. Her partnership with Tribal Head Start programs yielded a tech-based PD, overcoming geographic barriers to quality education and a model for designing culturally relevant curriculum freely available from *Head Start's National Center for Cultural and Linguistic Responsiveness*. [redacted] will provide the leadership to the coaching team and help in dissemination.

[redacted], **PhD**, is Professor of Literacy Education and Educational Psychology at Texas A&M University. He is writing multiple books with [redacted] and [redacted] in support of this reading comprehension project and will support the outreach efforts of the team.

[redacted], PhD, is Senior Research Scientist at TAMU CUSP. She is an award winning elementary school teacher and strong researcher. She is the author of the i-Lead intervention focusing on implementation of evidence-based interventions in schools. She served

as the project manager for the We Write and MOOV projects and has designed and delivered PD for thousands of teachers using multiple media platforms. She has conducted coaching and modeling, and assessments. She will lead the extant curricular review and KAT lesson guide preparation and support the MOOV teacher library and monthly teacher catch-up sessions

Independent Evaluation Team

WestEd has over 50 years of experience conducting research of programs that aim to increase student achievement. WestEd has extensive experience managing large, complex studies. WestEd's portfolio of current and recent projects demonstrates our ability to lead projects of similar size and scope, partner with stakeholders to collect and analyze data while conducting rigorous research, and ultimately to advance changes in educational practice to address racial and socioeconomic inequities in education. As examples, WestEd has led 41 grants funded through IES. For recent REL contracts, WestEd published 58 peer-reviewed research studies, including 9 RCTs. During the past three years, WestEd has been awarded 7 grants from the National Science Foundation and 2 grants from the EIR program. WestEd also runs the OSEP-funded National Center for Systemic Improvement and the National Research and Development Center to Improve Education for Secondary ELs.

[redacted], PhD, will serve as the PI and will have primary responsibility for ensuring that both impact and implementation study activities are carried out to the highest quality and in a timely and efficient manner. [redacted] will oversee the entire study, supported by Analytica and experienced quantitative analysts and the data collection team from WestEd. [redacted] brings more than 20 years of experience leading large-scale RCTs and other complex, mixed-methods third-party evaluations of literacy programs. She recently completed an RCT study for a study that is similar to the proposed one: the *Children's Literacy Initiative i3*

Scale-Up grant, which tested the efficacy of a teacher PD approach in 55 elementary schools. [redacted] serves as PI for an IES Systematic Replication RCT that evaluates the *Promoting Adolescents Comprehension of Text* intervention in 75 middle schools. Previously, [redacted] was co-PI for a multiyear IES Goal 3 RCT of the *HighScope Preschool Curriculum*. She played lead roles in two other large-scale federal RCT evaluations, both of which met WWC standards without reservations.

[redacted], PhD, will serve as the primary quantitative methodologist with responsibility for the final design plan and ensuring fidelity of the RCT methods. He brings experience with experimental and quasi-experimental evaluations of student learning initiatives. He regularly acquires, analyzes, and reports on individual student data, conducts educational finance analyses, and is accustomed to creating study design and analysis plans, data security protocols, and data sharing plans. Recently, as co-PI on the REMIQs Project – a multi-state investigation of school quality – he designed the study, oversaw data collection and analysis of state longitudinal data, and authored the technical report. He also completed the *Rhode Island Summer Readiness Evaluation* and the *Evaluation of the EL Parent/Teacher Training Certificate Project*.

[redacted], PhD, will serve as the WestEd data collection lead and manager. She has collected data —both remotely and in-person— from students and educators for national, state, and district projects. For example, as a task lead for the national NCEE-funded *Impact Study of Feedback for Teachers Based on Classroom Videos*, she executed weekly data collection from coaches and collected teacher surveys in 10 districts. For a study on the *Urban Advantage* middle school program, she collected pre- and post-intervention student and teacher science content knowledge assessments from schools across New York City.

[redacted], PhD, of Analytica Insights, Inc., will serve as a methodological review and monitoring consultant to the independent evaluation team and lead the cost study. He will apply two decades of WWC experience (since 2002), one decade of Investing in Innovation (i3) program experience (since 2010), as a Technical Assistance Team Leader, and one-week of training from the IES Cost Study Summer Institute to provide quality assurance review and monitoring of the impact, implementation, and cost studies as KATE is scaled up. He will also advise the evaluation team on how to be responsive and leverage the EIR technical assistance to ensure the impact study is designed, implemented, analyzed, and reported to meet EIR criteria and WWC Standards for Group Designs. [redacted] is a national expert with three decades of experience with rigorous impact evaluations to identify what interventions work in education and related areas. He contributed to innovative and cutting-edge initiatives focused on using evidence-based, scientifically valid methods when conducting federally-sponsored research. These initiatives include, in addition to the WWC and i3, the Campbell Collaboration (C2), the Regional Educational Laboratory (REL) Program, and the Social Innovation Fund (SIF). [redacted] is among a small and unique group of researchers currently certified in all three types of WWC Standards: Group Design 4.1; Regression Discontinuity Design 4.0; and Single Case Designs 4.1 with extensive experience conducting WWC reviews and WWC training on applying WWC Standards. [redacted] has applied his deep methodological expertise to apply mixed methods approaches to a wide range of research projects at several universities and research firms, including the Center for Responsive Schools, Buffett Early Childhood Institute at the University of Nebraska, Westat, WestEd, the HighScope Educational Research Foundation, and The Brady Foundation. He is an adjunct associate professor at the University of Pennsylvania's Graduate School of Education and The Lynch School of Education at Boston

College. He teaches courses on research methods and program evaluation. [redacted] continues to publish in peer-reviewed journals on standards for impact evaluations and related topics. He has written several methodological guidance documents to assist research teams in designing and implementing rigorous research.

[redacted]**, Consultant**, is Professor of Educational Psychology at Penn State and Advising Methodologist with Analytica Insights Inc. [redacted] was a co-PI and the methodologist for a U.S. Department of Health and Human Services Administration for Children and Families grant on developing early arithmetic, reading, and learning indicators for Head Start children. She also served as a co-PI for an IES efficacy grant to evaluate a Classwide Intervention Program in improving the social, behavioral, and academic outcomes in the primary grades and as a consultant for various funded projects on research design and measurement issues. [redacted] was co-PI on the recently completed ITSS efficacy study and on two ITSS extensions SWELL and WE-WRITE. She has been an Associate Editor of the *Journal of Educational Psychology* and is currently serving on the editorial board. [redacted] will serve as the counsel on the research design, conduct data quality reviews, data analysis reviews, and dissemination support.

Problem Solutions Team – Technology Delivery and User Support Team. Problem Solutions is a subsidiary of Allen Interactions Inc., with a long history of effective systems design and support. They have been integral to the ITSS, SWELL, We Write, and MOOV projects in designing the original systems to converting them to the current state of the art infrastructure. Their team will provide 24/7 support for all aspects of the technology infrastructure for the KATE project.

D.3. KATE costs are reasonable in relation to the objectives, design, and potential significance

Due to the advanced web-based intelligent tutoring platform hosting and delivering teacher PD and just-in-time resources the project costs are reasonable. Additionally, the outcomes of this project having the potential to make a difference in improving reading comprehension at upper elementary grade levels presents powerful support.

E. *Quality of the Project Evaluation*

E.1. KATE Evaluation is designed to meet WWC standards without reservations.

WestEd, working with consultation from [redacted] at Analytica Insights Inc., will be the independent evaluator for KATE and will conduct a multi-site cluster RCT. WestEd has conducted large-scale RCTs within the EIR program as well as for numerous Institute for Education Science (IES) grants. Analytica has been a subcontractor on the WWC since its inception in 2002, is a lead trainer for the WWC Statistics, Website, and Training team, and will provide consultation to ensure the study is conducted to meet EIR TA guidelines and WWC evidence standards. KATE is poised to meet WWC evidence standards without reservations: the outcome measures are content valid, psychometrically reliable, and aligned with the WWC's domains. We expect the RCT to have low school attrition because of strong incentives and low student attrition and non-response because of rigorous data collection monitoring. The evaluation team will evaluate fidelity of implementation consistent with the EIR TA guidelines and conduct cost analyses of KATE to establish a replicable and sustainable model beyond the EIR grant. The evaluation of KATE seeks to address the following confirmatory research questions about impacts on student outcomes and teacher practice:

Research Questions

(1) What is the impact of using ITSS with MOOV 360 teacher support for grade 4 and 5 students on measures of reading comprehension after one year of the intervention?

(2) *What is the impact of using ITSS with MOOV 360 for grade 4 and 5 students on reading motivation and self-efficacy outcomes after one year of the intervention?*

(3) *What is the impact of MOOV 360 on teachers' knowledge of reading comprehension and the text structure strategy among grade 4 and 5 teachers after one year of the intervention?*

(4) *What is the impact of MOOV 360 on teachers' instructional practices and use of the text structure strategy among grade 4 and 5 teachers after one year of the intervention?*

The evaluation will also answer the questions about fidelity of implementation and expansion:

(5) *To what extent is ITSS and MOOV 360 implemented with fidelity to the proposed model?*

(6) *What challenges are encountered during expansion, what conditions make implementation at scale more successful, and is the expansion cost effective?*

Finally, we will ask two exploratory research questions related to student outcomes:

(7) *What is the effect of ITSS with MOOV 360 on diverse student subgroups (e.g., grade 4 and 5, separately; gender, non-White students, Spanish speakers, students with disabilities)?*

(8) *To what degree do students who used ITSS with MOOV 360 in grade 4 continue to outperform students in control schools after summer break?*

The answers to the 8 research questions (RQ) will be helpful in support adoption or replication of KATE by other schools and districts.

Research Design

The Impact Study. The intent-to-treat (ITT) impacts will be estimated for two cohorts. Thirty-two Cohort 1 schools will participate in Year 2 of the grant and thirty-two cohort 2 schools will participate in Year 3 for a total of 64 schools in the study. For each year, the thirty-two schools in each cohort will be randomly assigned to the ITSS+MOOV 360 intervention or business-as-usual control within district blocks (clusters of similar schools that will be randomized to balance school locales, size, SES, prior ELA achievement). Random assignment will be conducted in

June for each cohort, in time for summer PD for the teachers (see Table 1 for study timeline overview). All general education classrooms in grades 4 and 5 within each school will be eligible to participate. The KATE project will carefully track student rosters, attrition, and non-response over the year-long intervention. Student joiners, or those who enroll six weeks after the school year starts, will be excluded from baseline and analytic samples.[1] As an incentive to participate and not to leave the study, schools randomly assigned to control condition will receive ITSS+MOOV 360 after the year-long study is complete. In the unlikely event of higher-than-expected student attrition or non-response, our collection of pretest data ensures that baseline equivalence between groups can be assessed, allowing the study to meet WWC evidence standards with reservations[2] (U.S. Department of Education, WWC 2021).

Sample and Statistical Power

To achieve statistical power, we estimate 32 schools per cohort, 4-10 classrooms per school, 20 students within each classroom (or approximately 4,800 students) to have an 80% chance of declaring a minimum detectable effect size (MDE) of .30 as statistically significant at the .05 alpha level.[3] This estimate assumes a 25% student non-response rate (i.e., 15 students per classroom in analysis sample). The inclusion of pretest and demographic data correlated with the outcome (explaining substantial residual variance) will included in the four-level analysis model to estimate effects. All assumptions that informed the power analysis are presented in Appendix J Exhibit J1. Thus far, eleven districts and one state have signed a letter stating interest in KATE and the supporting RCT. In addition, the Literacy.IO team has met with Digital Promise to

[1] The WWC Study Review Protocol 4.1 does not require this exclusion when the unit of random assignment is the school. However, for efficiency in consent and data collection, this exclusion rule is used.
[2] To be clear, meeting Standards with reservations is not the goal of the research team. However, collecting pretest data provides a second option by meeting Standards with reservations, consistent with EIR TA guidance.
[3] This estimate is conservative because we assume random rather than fixed effects.

become a member of the organization; they have a collection of districts who agree to participate in studies conducted by member institutions. The KATE leadership team will work with Digital Promise and continue to actively recruit the required number of schools.

E.2 Data Collection

All measures meet WWC standards for being reliable and valid. The pretest measures reduce residual variance in outcomes and increase precision of the impact estimate. The delayed posttests will provide valuable estimates of how well the intervention effects are sustained after summer break[4], a key area of interest in the reading field. Assessments will be administered on computers with the evaluation team proctoring administration of assessments (making data cleaning and analysis efficient and accurate). To answer questions about teacher knowledge and

Table 1. Evaluation Timeline

Phase	Cohort	Activities	Period
Year 1 Set-up		District/school final recruit/MOUs	Spring '22– Winter '23
		IRB approvals/consent	
2023-2024 School year	1 (n=32)	Random Assignment	June '23
		16 Treatment Schools, ITSS+MOOV 360	Aug '23–Jun '24
		Student Pretests	Fall '23
		Teacher Survey/Knowledge and Observations	Spring '24
		Cost Data Collection	Jan–Jun '24
		Student Posttest	Spring '24
2024-25 School year	1	Student Delayed Posttest	Fall '24
		16 Control Schools, ITSS+MOOV 360	Aug '24-Jun '25
	2 (n=32)	Random Assignment	June '24
		16 Treatment Schools, ITSS+MOOV 360	Aug '24-Jun '25
		Student Pretests	Fall '24
		Teacher Survey/Knowledge and Observations	Spring '25
		Cost Data Collection	Jan–Jun '25
		Student Posttest	Spring '25
2025-26 School year	2	Student Delayed Posttest	Fall '25
		16 Control Schools, ITSS+MOOV 360	Aug '25-Jun '26
		Analysis: Impact, Implementation, Cost	Winter–Spring '26
Year 5	-	Final Analysis, Reporting, and Dissemination	Summer–Fall '26

[4] For delayed effects, researchers will collect data from fourth graders who remain in the same schools (whereas fifth graders disperse to middle school and follow-up tests are time-intensive, costly and inclined to missing data).

instructional practices, the research team will use a knowledge assessment and classroom observations. Data to measure fidelity will be collected from the ITSS and MOOV 360 systems.

E.3.1. Student Measures. For RQs 1, 7, and 8, reading comprehension will be measured using a standardized test with multiple-choice questions mainly reflecting narrative text. Reading comprehension will also be measured using experimenter-designed tests about expository text.

E.3.1.1. Standardized test. The *Gray Silent Reading Test* (GSRT; Wiederholt & Blalock, 2000) will be administered at pretest, posttest 1 and delayed posttest 2. The GSRT was selected because of its characteristics enabling (a) group administration, (b) testing of deep comprehension processes with paraphrasing, application, and inference questions, and (c) at least two equivalent forms with strong psychometric properties (average alternate form reliability was .85 and delayed alternate-form reliability was reported at .83). Coefficient alpha reported for forms A and B were .95 and .94, respectively.

E.3.1.2. Researcher-developed measures. Three measures were created to capture specific aspects of comprehension: Comparison Text Structure Signaling Word Knowledge (CSW), Comparison Main Idea Quality (CMIQ), Problem and Solution Top-Level Structure Quality (PTLSQ). Each was validated under IES grants (Meyer et al., 2010; Meyer et al., 2012), used in several studies (e.g., Wijekumar et al., 2014; Wijekumar et al., 2017), and align with the reading comprehension outcome domain in the WWC Study Review Protocol 4.1 (SRP 4.1). The tests measure student competency in (1) understanding signaling words, (2) constructing main ideas, and (3) recall using top-level structure of the text.

CSW and CMIQ use three equivalent set of passages with 128 words, 15 sentences, 96 idea units, and equivalent scores on traditional measures of readability (e.g., Lexile®; Flesch-Kincaid, sentence length), text structure, and signaling (see Meyer, 2003). Each passage has four fill-in-

the-blank activities for CSW. Students then write a main idea for the passage to assess CMIQ. PTLSQ uses a set of three equivalent passages that have 98 words, 72 idea units, and equivalent scores on traditional measures of readability. Students read the passage and recall the text without the passage in view. The computer NLP scoring system for these constructs was trained with over 14,000 student samples and has a 99% accuracy to human scored essays from previous administrations (see Table 2).

Multiple forms of these measures were used as pre- and post-tests in the SEED grant (Wijekumar et al., 2021a) and evidence supported equivalence of forms and computer scoring. For KATE, we will administer three forms at pretest, posttest 1, and delayed posttest 2 for grade 4 students. We will administer two of the three forms at pretest and posttest 1 for grade 5 students. The CSW will be used as a covariate in the data analysis because it correlated with the GSRT scores in previous efficacy studies (.79 and .58, for grades 4 and 5, respectively Wijekumar et al., 2012; Wijekumar et al., 2014).

E.3.1.3. Motivation to Read. For RQ2, the reading self-concept scale from the Motivation to Read survey will be used at pre- and post-test, as a validated and reliable measure of student reading motivation and self-efficacy (Gambrell et al., 1995). This measure aligns with the WWC's Social-Emotional Outcomes, Behaviors, and Skills in the WWC SRP 4.1.

E.3.2. Teacher measures. To address RQs 3, 4, and 6, the research team will use a knowledge assessment survey and classroom observations.

E.3.2.1. Teacher Comprehension Knowledge Assessment. Teacher knowledge of comprehension strategies will be measured prior to PD and at the end of the academic year for each cohort. For RQ3, the survey contains (1) Teacher Comprehension Main Idea Assessment (TCMIA), a researcher-designed teacher knowledge measure of comprehension constructs such as signaling

words, recall, summarization and main idea as well as (2) Teacher Knowledge of Vocabulary and Comprehension (TKVC), which measures content knowledge, pedagogical knowledge, and application abilities related to vocabulary and reading comprehension. TKVC is based on previous studies of teachers' knowledge (Binks-Cantrell et al., 2012; Ely et al., 2014; Peltier et al., 2020; Spear-Swerling & Cheesman, 2012; Spear-Swerling & Zibulsky, 2014). TKVC involves 50 multiple-choice items. The vocabulary subscale has 24 scorable items (23 content knowledge, 9 application, and 5 pedagogical knowledge) and the comprehension subscale has 37 scorable items (10 content knowledge, 7 application, and 7 pedagogical knowledge). Finally, there are 13 items that assess teachers' perceived expertise for vocabulary and reading comprehension (rated on a Likert scale of minimal to expert). The TKVC was found to have acceptable reliability, with a Cronbach's $\alpha = 0.85$ (Hudson, 2021). The teacher knowledge measure is used in the SEED grant and has been validated by three experts and scored using the same accurate, consistent computer scoring system described in the student measures section (reliable at 99% accurate to human scores).

The teacher survey will also gather teacher demographic information and data on the hours and types of literacy-related PD teachers participated in during that school year, including questions relevant for time (cost) spent. Treatment teachers will also answer questions related to any barriers or facilitators to using ITSS and MOOV 360 (RQ6).

E.3.2.2. Observation Measure. To answer RQ4, the Classroom Reading Observation Tool measures specific elements of comprehension and vocabulary instruction, including main idea (with text structure or not), summarization, and inferencing. Each evidence-based practice tied to vocabulary and comprehension is indicated as *present* or *not present* and a short description is added by the observer. The measure has been validated by 3 experts and has inter-rater reliability of 98-99% (Hudson et al., 2021). This measure will be administered in a randomly-selected

subset of intervention and control classrooms. Appendix J: Figure J9 shows the constructs measured using the tool. Table E2 summarizes and provides reliability information for each measure described above.

Table E2. Measures and Reliability

Measure – Computer Scored			Properties and Use		
Level	Type	Name	Cronbach Alpha or IRR	Analysis	RQ
Student	General Reading Comprehension	Gray Silent Reading Test	.83 - .94	Impact	1,7,8
	Specific Comprehension Skills[a]	Comparison Text Structure and Signaling Word (CSW)	99%		
		Comparison Main Idea Quality (CMIQ)	99%		
		Problem and Solution Top-Level Structure Quality (PTLSQ)	99%		
	Reading Affect	Motivation to Read; reading self-concept[b]	.77 -.78		2
Teacher	Teacher knowledge[c], practices, and background	Teacher Comprehension Main Idea Assessment (TCMIA) Teacher Knowledge of Vocabulary and Comprehension (TKVC)	.85	Impact & Implem-entation	3, 6
	Observation	Classroom Reading Observation Tool	98% to 99%	Impact & Implem-entation	4

[a]Reliability from Wijekumar et al., 2012; 2014, 2017. [b]Reliability from items 7, 9, 13, & 15. [c]Hudson 2021

E.3. KATE evaluation plan clearly articulates the key project components, mediators, and outcomes, as well as a measurable threshold for acceptable implementation.

There is growing recognition within intervention research to clearly articulate and then evaluate the fidelity of implementation of critical components of an intervention (Wilson et al., 2020). To answer RQ5 related to implementation fidelity, WestEd will collect data on each component of KATE (depicted in the logic model): teacher PD including (1) *Initial 24 hours of training*; (2) *Use of MOOV 360 resources*; (3) *two sessions of in-class coaching; (4) Weekly coach meetings*; (5) *Monthly catch-up*; and (6) *Quarterly benchmark reviews*, as well as the student components including (7) *Weekly ITSS use of 60 minutes*; (8) *Coverage of text structures* across ITSS lessons (must complete 5 lessons of each structure); and (9) *Adaptive features* that respond to individual

students, such as Spanish supports and differentiated content. WestEd and Analytica will work with the Literacy.IO to determine thresholds for each component that indicates acceptable implementation. For example, the KATE team expects that a teacher would complete all 24 hours of initial training and both in-class coaching sessions, however, a teacher receiving >85% of these hours would be considered as having *high* fidelity. A teacher who receives 70-85% of these PD components will be deemed to have *medium* fidelity and <70% would have *low* fidelity. Individual-level ratings will be aggregated to the school and district levels and score for each indicator will be summed to create implementation scores at both the student and teacher level. Data sources related to implementation of these components are listed below.

KATE Staffing Records. Literacy.IO will provide records of resources delivered to teachers including training and meeting attendance records and coaching hours delivered to teachers.

ITSS and MOOV 360 Digital Records. Literacy.IO will provide access to its online system that tracks (a) the resources teachers utilize (number of logins, time spent with different teacher resources, text structures, and curriculum alignment materials). The system also contains (b) student tracking that will be used to determine total dosage (time in the ITSS program) as well as coverage for each text structure - comparison, problem and solution, cause and effect, cause problem and solution) and degree to which different adaptive features are utilized, including language supports.

Coaching Log. A MOOV web-based log will be used to gather data about the agendas coaches follow in their classroom visits and virtual weekly meetings with teachers. These forms document the content and practices covered during coaching sessions. The logs will also collect feedback from coaches about barriers and facilitators to implementation.

E.4**. The KATE evaluation will provide guidance about effective strategies suitable for replication or testing in other settings.**

To answer RQ6 related to expansion and how to address potential challenges to make replication and scale up more successful and cost-effective, the research team will triangulate a variety of data sources, including survey and observation data, as well as data on the usage of the ITSS and MOOV 360 systems, and cost data. The research team will document the process from training to use and report on the context of the study to inform future use of the software, teacher resources, and school implementation.

Based on prior project experiences, anticipated challenges could include potential local adaptations or weakness with fidelity. Researchers will begin with observing the summer PD for teachers to learn how they understand the intervention and the text structure instructional practices. For example, we will document teacher questions and discussion to understand teacher concerns that could be addressed in future training or used to refine support materials. Teacher surveys will also provide information on local context and conditions that can make implementation at scale more successful. The surveys will further provide data on things such as classroom practices and resources that can be used to understand implementation and provide lessons for scaling. For example, survey data could point to differences in implementation between large and small classes or differences in contexts with one-to-one devices compared to those that use computer labs. Finally, data from the online ITSS and MOOV 360 systems will also be analyzed to identify potential challenges to implementation that could be addressed to make replication and scale-up more successful. Usage data for teachers and students could inform improvements that would make future expansion smoother.

Cost Study. This portion of the study will systematically collect data on costs of implementing critical components of the ITSS + MOOV 360, and given these costs, whether it is a cost-

effective intervention. This analysis aims to provide education decision-makers with information on the resources needed for schools and districts to replicate and implement the ITSS + MOOV 360 in different settings and contexts. We will use the ingredients method and a "schools' perspective" to estimate the costs of implementing the critical components, as denoted in the logic model (e.g. Levin et al., 2018). This economic-based method accounts for both expenditures and opportunity costs incurred during implementation. Estimating the opportunity costs are essential, and the schools' perspective means the analysis will be limited to the costs incurred by schools to implement the intervention (IES, 2020). When educators invest time, space, and resources to implement ITSS + MOOV 360, there are alternative uses for these resources, such as other activities or programs, that may produce potential benefits that can no longer be realized due to their use to implement the ITSS + MOOV 360 (Hollands et al., 2016). These alternative uses will be identified in addition to identifying the accounting cost or direct expenditures incurred while implementing the intervention (Keeney & Raiffa, 2003). All analyses will be conducted consistent with the Cost Analysis in Practice Project funded by IES.

E.5 Data Analysis

The data to be analyzed will be free of confounds (because of random assignment), based on reliable and valid pre- and post-tests, free of risk of bias due to high attrition (cluster or student) and thereby qualifying for the highest WWC rating of "meets standards without reservations." Assessment of attrition will be conducted on the analytic sample prior to adjustments for missing data. We will calculate overall and differential attrition on the analytic sample and determine whether attrition is considered low or high using WWC's latest criteria, at the cluster level and subcluster levels. If the attrition rate exceeds the WWC thresholds in the WWC SRP 4.1, we will assess baseline equivalence of intervention and comparison groups on demographic characteristics and pretest scores with the analytic sample. All baseline characteristics that show

nonequivalence of larger than .05 effect size (i.e., Hedges' g for continuous variables or Cox index for dichotomous ones) will be included as covariates in all HLM analyses (i.e., statistical adjustment) to minimize bias in the impact estimates.

Missing data will be handled by listwise deletion if the percentage of missing data is 5% or less and the missing pattern is completely at random (MCAR) or by multiple imputation if missing is at random (MAR). Both missing data methods are acceptable under WWWC Standards 4.1. However, for multiple imputation, we will meet the additional WWC requirements of 1) imputation is conducted separately for intervention and comparison groups or include an indicator variable for intervention status, 2) the same covariates used in the analysis model are used in the imputation model, and 3) the outcome is included in the imputation model if baseline data are imputed (WWC Handbook 4.1, p.37).

HLM Model Specifications: Addressing Primary and Secondary Research Questions

For the HLM models, students are nested within classrooms, classrooms within schools, and schools are nested within sites. Because of the modeling complexity attributed to four-level structures, we will test initially the degree to which sites differ on each outcome variable using an unconditional four-level model. Should the test of the outcome variance at the site level demonstrate non-significance, we will simplify the models to three levels. Otherwise, we will analyze four-level models.

At the student level, predictor variables will include demographic variables such as gender, ethnicity, and special education status, reading comprehension pretest covariates, and affective pretest covariates (e.g., motivation to read). Demographic variables will be grand-mean centered, and pretest scores will be group-mean centered.

At the teacher/classroom level, classroom characteristics such as teacher experience, class-level pretest scores of comprehension and motivation. Teacher experience will be grand-mean centered, and class-level pretest scores will be group-mean centered.

Intervention condition and school characteristics such as percent of students receiving reduced-price lunch and school-level pretest scores will be examined at the school level. Potential cross-level interaction effects between treatment and student level variables (e.g., whether treatment effect varies depending on initial reading level; whether treatment reduces gender gap in reading) will also be explored by modeling level-1 coefficients as a function of treatment. Similarly, potential cross-level interaction effects between treatment and class level variables (e.g., whether treatment effect varies depending on teacher experience) will be explored by modeling level-2 coefficients as a function of treatment. Potential interaction effects between treatment and school characteristic variables will be examined by adding product terms at the school level.

Because our analyses include both cognitive and affective outcomes measured at different time points, potential mediating effects can be detected by examining the pattern of significant effects. For example, if affective variables mediate treatment effect on reading, treatment effect on the affective variables would be significant, and the effect of affective variables on reading would be significant. If any potential mediation effects are noted, we will use structural equation modeling to further examine the mediation effects.

In addition, we will estimate effect size of ITSS + *MOOV 360* as compared to the business-as-usual comparison group. Specifically, we will compute the effect size as a standardized mean difference by dividing the adjusted (for pretest scores and other covariates) group mean difference by the unadjusted *pooled* within-group student-level standard deviation of

the outcome measure (i.e., Hedges' g if standard deviations are similar for all treatment conditions) or by the *comparison* group student-level standard deviation (i.e., Glass' delta if standard deviations are very different, e.g., Lipsey & Wilson, 2001).

The Benjamini-Hochberg adjustment (B-H; Benjamini & Hochberg, 1995) will be used to control the false discovery rate in testing multiple hypotheses. The B-H procedure is more powerful than the Bonferroni technique (Williams, Jones, & Tukey, 1999) and has been adopted in many research studies.

References

Barshay, J. (2017, September 25). 3 lessons learned from education technology research. *U.S. News & World Report.* Retrieved from https://www.usnews.com/news/education-news/articles/2017-09-25/3-lessons-learned-from-education-technology-research

Beerwinkle, A., Wijekumar, K., Walpole, S. & Aguis, R. (2018). An Analysis of the Ecological Components within a Text Structure Intervention, *Reading and Writing: An Interdisciplinary Journal* , 31,9: 2041-2064.

Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bogaerds-Hazenberg, S.T.M., Evers-Vermeul, J., & van den Bergh, H. (2020). A Meta-Analysis on the effects of text structure instruction on reading comprehension in the upper elementary grades. *Reading Research Quarterly*. https://doi.org/10.1002/rrq.311

Boyd, D., Lankford, H., Loeb, S., Rockoff, J., & Wyckoff, J. (2008). The narrowing gap in New York City teacher qualifications and its implications for student achievement in high-poverty schools. *Journal of Policy Analysis and Management, 27*, 793-818.

Chall, J., Jacobs, V., & Baldwin, L. (1990). The reading crisis: Why poor children fail. First Harvard University Press.

Chin, T. & Phillips, M. (2004). Social Reproduction and Child-Rearing Practices: Social Class, Children's Agency, and the Summer Activity Gap. Sociology of Education, Vol. 77, No. 3 (Jul., 2004), pp. 185-210.

Common Core Standards Initiative (2012). Retrieved November 20, 2012, from http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf

Cook, L. K., & Mayer, R. E. (1988). Teaching readers about the structure of scientific text. *Journal of Educational Psychology, 80,* 448-456.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334. doi:10.1007/BF02310555

Gambrell, L. B., Morrow, L. M. E., & Pressley, M. E. (2007). Best practices in literacy instruction. Guilford Press.

Gambrell, L. B., Palmer, B. M., Codling, R. M., & Mazzoni, S. A. (1996). Assessing motivation to read. *Reading Teacher, 49,* 518-533.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American educational research journal*, *38*(4), 915-945.

Garza, J.V., (2019). Astronomical improvements in reading comprehension with text structure strategy. Presentation at the Texas Association of School Administrators Conference, April 2019, Austin, TX.

Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research, 71,* 279-320.

Green, A (2019). *Text structures for success: Improving test scores in Brownsville*. Retrieved from http://transform.tamu.edu/news/text-structures-success-improving-test-scores-brownsville

Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. American Educational Research Journal, 45(1), 184-205.Guthrie, J.T., & Davis, M.H. (2003). Motivating Struggling Readers in Middle School Through and Engagement Model of Classroom Practice. *Reading & Writing Quarterly*, *19*, 59-85

Hedges, L. V. & Hedberg, E. C. (2014). Intraclass correlations and covariate outcome correlations for planning 2 and 3 level cluster randomized experiments in education. Evaluation Review.

Hollands, F.M., Kieffer, M.J., Shand, R., Pan, Y., Cheng, H., & Leven, H.M.(2016). Cost-effectiveness analysis of early reading programs: A demonstration with recommendations for future research. Journal of Research on Educational Effectiveness, 9(1), 30-53.

Hudson, A.K., Owens, J.K., Moore, K.A., Lambright, K. & Wijekumar, K. (2021). "What is the Main Idea?": Using Text Structure as a Framework for Accelerating Strategic Comprehension of Text. Reading Teacher, doi:10.1002/trtr.2016

Hudson, A.K. (2021). Teachers' knowledge of reading comprehension, classroom practice, and student reading comprehension growth. Unpublished Dissertation, Texas A&M University.

Institute of Education Sciences. (2020). Cost Analysis: A Toolkit (IES 2020-001). U.S. Department of Education. Washington, DC: Institute of Education Sciences. Retrieved [month day, year] from https://ies.ed.gov/pubsearch/

Joshi, R.M., Binks, E., Graham, L., Ocker-Dean, E., Smith, D.L., & Boulware-Gooden, R. (2009). Do textbooks used in university reading education courses conform to the instructional recommendations of the National Reading Panel? *Journal of Learning Disabilities, 42,* 458-463.

Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., and Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A Practice Guide* (NCEE #2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc.

Keeney, R. L., Meyer, R. F., & Raiffa, H. (2003). *Decisions with multiple objectives: Preferences and value tradeoffs*. Cambridge: Cambridge, Univ. Press.

King-Dickman, K. (2013). How Elena Learned to Love Reading. Educational Leadership, 70(8),

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363–394. https://doi.org/10.1037/0033-295X.85.5.363

Kintsch, W. (2005) An Overview of Top-Down and Bottom-Up Effects in Comprehension: The CI Perspective, Discourse Processes, 39:2-3, 125-128, DOI: 10.1080/0163853X.2005.9651676

Ladson-Billings, G. (2021). The Four Pandemics of Minority Children. Presentation at Texas A&M University. February 2021.

Lemke, J.L. (1990). *Talking Science: Language, learning, and values.* Norwood, NJ: Ablex Publishing Corporation

Levin, H. M. (1975). Cost-effectiveness analysis in evaluation research. In E. Struening & M. Guttentag (Eds.), Handbook of Evaluation Research (Vol. 2). Beverly Hills, CA: Sage.

Levin, H. M., & McEwan, P. J. (2001). Cost-effectiveness analysis: Methods and applications (2nd ed.). Thousand Oaks, CA: Sage.

Levin, H. M., McEwan, P. J., Belfield, C. R., Bowden, A. B., & Shand, R. (2018). Economic evaluation in education: Cost-effectiveness and benefit-cost analysis. Los Angeles, CA: SAGE.

Lyon, G. R. (2001). *Measuring Success: Using assessments and accountability to raise student achievement.* Hearing before House Committee on Education and the workforce, subcommittee on Education Reform, 107th Congress. Available at http://www.nrrf.org/learning/statement-of-dr-g-reid-lyon-before-the-u-s-house-subcommittee-on-education-and-the-workforce-hearing/

Marshall, G. (2013). Educating literate career- ready citizens should be a national priority. *Reading Today, 30*(5), 27. Retrieved from http://search.proquest.com/docview/1349588909?accountid=13158

Meyer, B. J. F. (1975). *The organization of prose and its effects on memory.* Amsterdam: North-Holland.

Meyer, B. J. F., Middlemiss, W., Theodorou, E., Brezinski, K. L., McDougall, J., & Bartlett, B. J. (2002). Effects of structure strategy instruction delivered to fifth-grade children using the Internet with and without the aid of older adult tutors. *Journal of Educational Psychology, 94*, 486-519.

Meyer, B. J. F., & Poon, L. W. (2001).Effects of structure strategy training and signaling on recall of text. *Journal of Educational Psychology, 93*, 141-159.

Meyer, B.J.F., Wijekumar, K., Middlemiss, W., Higley, K., Lei, P., Meier, C., Spielvogel, J. (2010). Web-based tutoring of the structure strategy with or without elaborated feedback or choice for fifth- and seventh-grade readers. *Reading Research Quarterly, 41*, 62-92.

National Assessment of Educational Progress (NAEP) 2019. Retrieved Nov 1, 2019 from https://www.nationsreportcard.gov/highlights/reading/2019/

Pearson, P. & Cervetti, Gina. (2017). S.Israel (Ed). The roots of reading comprehension instruction. Handbook of research on reading comprehension. Guilford

Perfetti, C. & Stafura, J. (2014) Word Knowledge in a Theory of Reading Comprehension, Scientific Studies of Reading, 18:1, 22-37, DOI: 10.1080/10888438.2013.827687

Pressley, M., Wharton-McDonald, R., Mistretta-Hampston, J., & Echevarria, M. (1998). Literacy instruction in 10 fourth- and fifth-grade classrooms in upstate new york. *Scientific Studies of Reading, 2*(2), 159-194. Retrieved from

Raudenbush, S. W., & Bryk, A. S. (2002). *Hiearchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: SAGE publications.

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–116). New York, NY: Russell Sage Foundation and Spencer Foundation.

Reardon, S. F. (2013). "The Widening Income Achievement Gap." *Educational Leadership 70*,10-16.

Sahni, S.D., Polanin, J.R., Zhang, Q., Michaelson, L.E., Caverly, S., Polese, M.L., & Yang, J. (2021a). A What Works Clearinghouse rapid evidence review of distance learning programs (WWC 2021-005). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, What Works Clearinghouse. Retrieved from: https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/Distance_Learning_RER_508c.pdf

Shanahan, T. (2020). What Constitutes a Science of Reading Instruction? *Reading Research Quarterly*, *55*.

Spybrook, J., Bloom, H., Congdon, R., Hill, C., Martinez, A., & Raudenbush, S. (2011). Optimal design plus empirical evidence: Documentation for the "Optimal Design" software.

Retrieved from http://www.wtgrantfoundation.org/resources/consultation-service-and-optimal-design.

TX STAAR (2018). Data File of State of Texas Assessment of Academic Readiness 2010 to 2018.

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the on-line construction of a memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71-98). Mahwah, NJ: Erlbaum.

van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes,* 39, 299-316.

Wiederholt, J. L., & Blalock, G. (2000). *Gray Silent Reading Tests* (GSRT). Austin, TX: PRO-ED

Wijekumar, K., Meyer, B.J.F., Lei, P. (2012). Large-scale randomized controlled trial with 4th graders using intelligent tutoring of the structure strategy to improve nonfiction reading comprehension. *Journal of Educational Technology Research and Development.* 60, 987-1013

Wijekumar, K., Meyer, B.J.F., Lei, P.-W. (2013). High-fidelity implementation of web-based intelligent tutoring system improves fourth and fifth graders content area reading comprehension. *Computers & Education, 68,* 366-379.

Wijekumar, K., Meyer, B. J., Lei, P. W., Lin, Y. C., Johnson, L. A., Spielvogel, J. A., ... & Cook, M. (2014). Multisite randomized controlled trial examining intelligent tutoring of

structure strategy for fifth-grade readers. *Journal of Research on Educational Effectiveness*, 7(4), 331-357.Wijekumar 2017

Wijekumar, K., Meyer, B. J. F., & Lei, P. (2017). Web-based text structure strategy instruction improves seventh graders' content area reading comprehension. *Journal of Educational Psychology, 109*(6), 741-760. http://dx.doi.org/10.1037/edu0000168

Wijekumar, K., Meyer, B.J.F., Lei, P-W, Hernandez, A., August, D. (2018). Effects of web-based text structure instruction for 4-6th grade Spanish Els reading comprehension. *Reading and Writing: An Interdisciplinary Journal*, 31,9: 1969-1996.

Wijekumar, K., Beerwinkle, A., McKeown, D., Zhang, S., & Joshi, R. M. (2020). *The "gist" of the reading comprehension problem in grades 4 and 5*. *Dyslexia: An International Journal of Research and Practice* | ISSN: 1076-9242 PMID: NLM31903668 |

Wijekumar, K., McKeown, D., Thompson, J.L., (2021a). Annual report submitted to the Supporting Effective Educator Development (SEED) grant (U423A180074) on the Massively Open Online Virtual (MOOV) learning system.

Wijekumar, K., Zhang, S., Peti-Stantic, A., Joshi, R.M. (2021b). Why Textbooks Matter: A Global Review of Reading Textbooks. *Technology, Knowledge, and Learning*.

Williams, V.S.L., Jones, L.V., & Tukey, J.W. (1999). Controlling Error in Multiple Comparisons, with Examples from State-to-State Differences in Educational Achievement Journal of Educational and Behavioral Statistics 24(1) DOI: 10.2307/1165261

Wilson, S. J., Lipsey, M. W., Aloe, A., & Sahni, S. (2020). Developing evidence-based practice guidelines for youth programs: Technical report on the core components of interventions that address externalizing behavior problems. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services

Wilson, V. L., & Rupley, W. H. (1997). A structural equation model for reading based on background, phonemic, and strategy knowledge. *Scientific Studies of Reading, 1*, 45-63.

WWC ITSS Intervention Report (2021). WWC ITSS Intervention Report. Available at https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_ALitss_IR_apr2020.pdf