"HIGH TOUCH HIGH TECH" (HTHT) MATHEMATICS IN URUGUAY – IMPACT EVALUATION REPORT

Report Prepared by
Yoon Soo Park and Janice Heejin Kim

Education Commission Asia (ECA)

December 6, 2023

Evaluation of the HTHT project was led by the Education Commission Asia (ECA), in partnership with Plan Ceibal and the Inter-American Development Bank (IDB). We thank the project team for the collaborative partnership and feedback in the development of this report.

Questions may be directed to Yoon Soo Park (email: yspark@mgh.harvard.edu) or Janice Heejin Kim (email: janice.hjkim@gmail.com).

EXECUTIVE SUMMARY

Delivering effective learning in traditional classroom settigns can be significantly enhanced using the "High Touch High Tech" (HTHT) learning paradigm, incorporating a synergistic instructional approach that (1) integrates computer-assisted asynchronous personal learning ("High Tech") and (2) promotes synchronous active learning facilitated by teachers in the classroom setting ("High Touch").

This report presents the *impact* and *effectiveness* of the HTHT educational intervention implemented in Uruguay targeting 5th grade mathematics teachers and students. Evaluation of the HTHT project was led by the Education Commission Asia (ECA), in partnership with Plan Ceibal and the Inter-American Development Bank (IDB).

This project incorporates an experimental design (two-group pretest-posttest design) – targeting teacher interventions across two groups ("Soft" and "Hard" interventions) for comparison with a "Control" group following standard education curriculum. Teachers in the intervention groups ("Soft" and "Hard" interventions) receive training and mentoring; in addition, teachers in the "Hard" intervention group receive supplemental individualized teaching feedback. Data were collected from teachers and students, targeting educational constructs on learning and teaching behavior, in addition to students' mathematics performance, measured using a standardized national assessment (SEA+). Data were collected in two phases: April 2022 (baseline) and October 2022 (post intervention).

Data were gathered from 108 schools, 155 teachers, and 2,709 students: "Control" Group (n = 28 schools, 39 teachers, 711 students); "Soft"Intervention Group (n = 51 schools; 74 teachers; 1,233 students); and "Hard" Intervention Group (n = 29 schools; 42 teachers; 765 students). Response rate for participation across both baseline and post-intervention study phases was 74.5% (Baseline $[T_0]$ = 2,709; Endline $[T_1]$ = 2,018). We note that sample sizes vary depending on unit of analysis (school, teacher, and students). To allow comprehensive analyses with HTHT data gathered, we created two panels.

- **HTHT increased the individualized use of computer-assisted learning by 49%**
    - Students in HTHT Intervention groups ("Soft" and "Hard" Interventions) had significantly greater completion rates for Matific episodes.
    - Between students in the "Soft" and "Hard" Intervention groups, the latter completed 20% more Matific episodes. Teachers who received classroom observation and individualized feedback sessions had greater use of Matific.
    - There were no significant differences at the aggregate classroom level, signaling variation in Matific usage was individualized at the student level between the control and intervention groups.

- **HTHT with higher use of Matific improved student performance by 0.33 SD**
    - The intervention had a null effect (of about 0.02 SDs) on the math achievement of the average students after accounting for students' initial achievement and characteristics.

- However, the null average effects masked important heterogeneous impacts. The intervention had a medium-to-large positive effect of 0.33 SD for students with higher use of Matific, who completed more Matific episodes at school.
- Notably, treatment students in the highest Matific use group outperformed their control counterparts by 0.76 SD. HTHT intervention had much smaller effects of 0.12 SD for students with lower Matific use and null effects for students who did not complete any Matific episode.
- Between students in the "Soft" and "Hard" Intervention groups, the difference in mathematics performance was modest.

- **A combined HTHT intervention integrating synergistic components of "High Touch" and "High Tech" leads to improved student performance.**
  - Computer-assisted personalized learning ("High Tech") or teacher interventions ("High Touch") alone does not independently lead to improved student performance.
  - A combined strategy utilizing both HTHT components demonstrated improvements in student performance.

This is the first large-scale experimental study demonstrating the effectiveness of the HTHT paradigm. Interventions that leverage both HTHT components in large-scale implementation through a controlled experimental design are uncommon. Educational policy may be designed to maximize interventions that combine HTHT components synergistically, while allowing flexibility in the large-scale implementation of teacher interventions.

"High Touch High Tech" Mathematics in Uruguay –
Impact Evaluation Report

I. Project Background and Goals

This report presents the *impact* and *effectiveness* of educational components associated with the "High Touch High Tech (HTHT)" project, led by the Education Commission Asia (ECA), in partnership with Plan Ceibal and the Inter-American Development Bank (IDB).

1.1. Project Background: "High Touch HighTech" (HTHT) Learning

Creating effective learning environments for students and teachers have been an ongoing challenge across the continuum in education (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). Educators and policymakers have wrestled to identify effective teacher interventions (henceforth "professional development" [PD] programs) that translate their impact on teaching effectiveness and also on student achievement outcomes. Traditionally, PD programs have been used as a mechanism to improve the quality of classroom instruction and student achievement; however, there has been a lack of well-documented empirical studies that provide evidence on the impact of PD programs on teachers' knowledge and behaviors, in addition to student achievement. Prior PD interventions have relied on single-shot, one-day workshops that have generated superficial and incoherent effects on teaching effectiveness (Ball & Cohen, 1999; Wilson & Berne, 1999; Pianta, 2011).

This project aims to transform the landscape of teacher PD interventions with enhanced learning environment for students – by incorporating a "High Touch High Tech" (HTHT) learning environment. The HTHT framework incorporates two educational components:

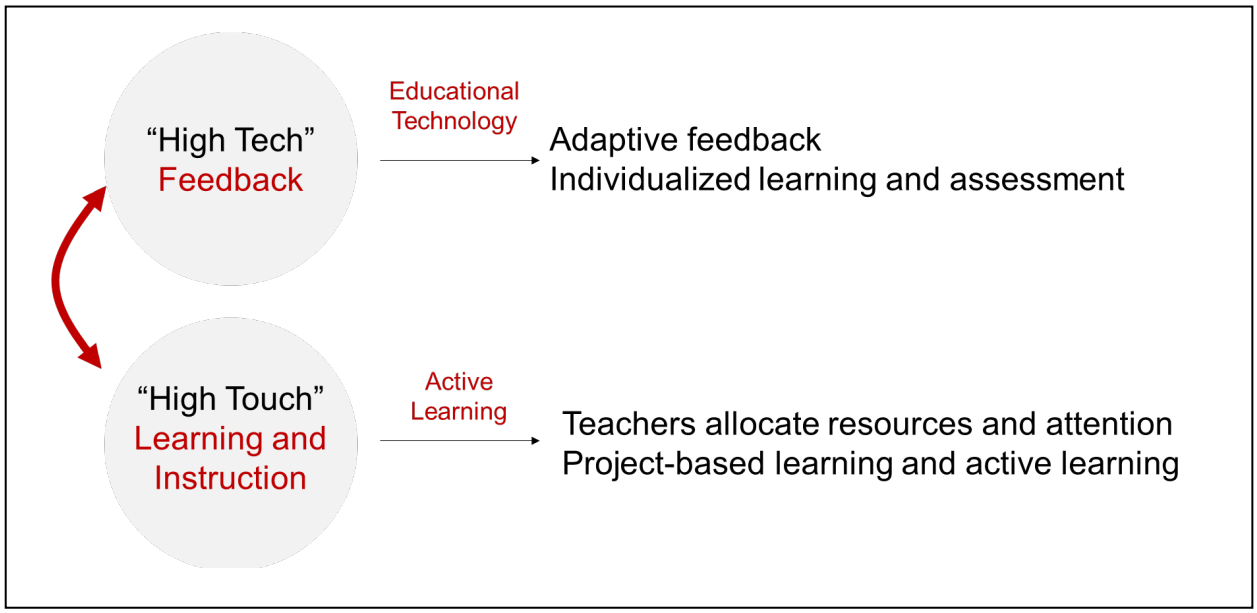
(1) **"High Tech" Learning:** Computer-assisted technology (including adaptive intelligent tutoring systems [ITS]) to provide *asynchronous* personalized learning and assessment to students

(2) **"High Touch" Learning:** *Synchronous* learning in the classroom facilitated by teachers to focus their teaching and resources on promoting project-based learning and active learning in small groups

The "High Touch High Tech" learning model aims to create a synergistic impact that increases teaching effectiveness and overall student achievement in cognitive and non-cognitive skills.

While studies have examined the unique contribution of computer-assisted learning or active and personalized learning in classrooms, there are no large-scale experimentally controlled studies that investigate the joint synergistic effect of these interventions on teacher and student outcomes. In this study, we conceptualize the "High Touch High Tech (HTHT)" learning model from a translational evaluation model, hypothesizing that the HTHT model that aims to transform teaching and learning environment that yield benefits to improve teaching effectiveness and student outcomes. We report on the implementation of the HTHT learning framework in Uruguay targeting mathematics.

Figure 1. The "High Touch High Tech" Learning Paradigm



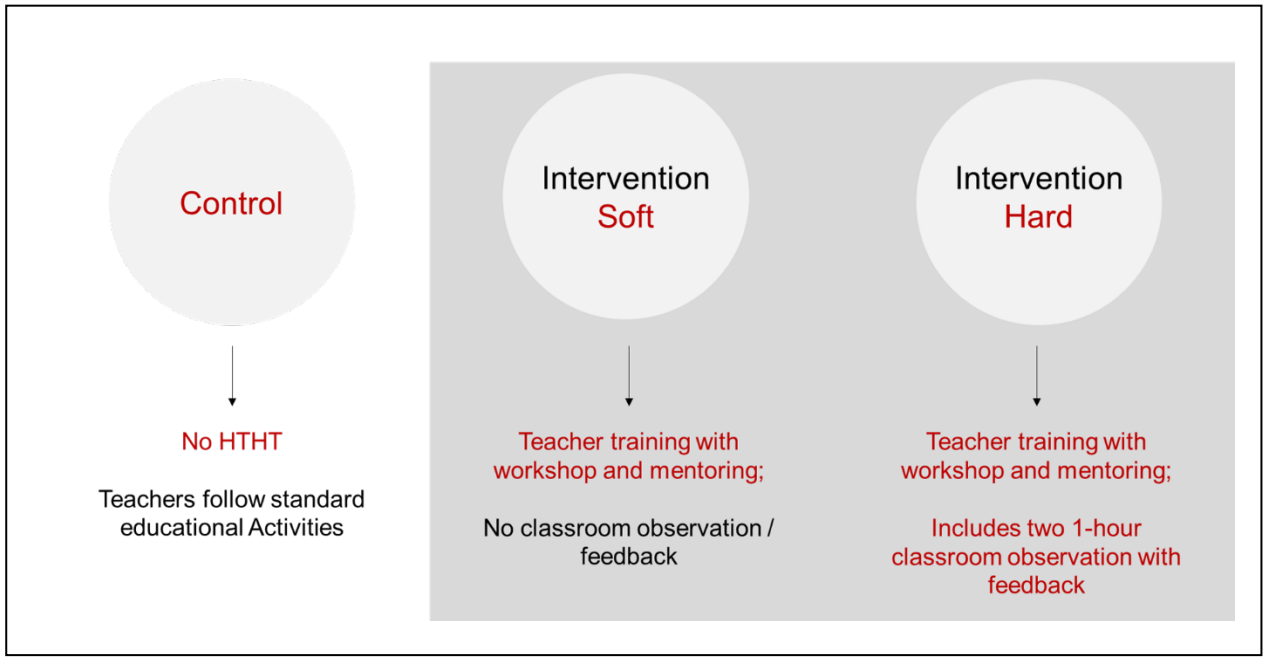
Note: The "High Touch High Tech" learning paradigm leverages the synergistic impact of (1) "High Touch" learning (facilitated by teachers incorporating active learning through synchronous classroom setting) and "High Tech" learning (using computer-assisted technology for asynchronous personalized learning).

1.2. Project Goals

We evaluate the impact of the "High Touch High Tech" learning project in Uruguay on teaching effectiveness and student outcomes in cognitive mathematics performance and non-cognitive skills (21st century skills targeted in the HTHT interventions, collaboration, and communication).

We examine teaching effectiveness and student outcomes using an experimental design with schools (teachers) randomized to the treatment group (receiving the HTHT intervention) and the control group (standard education). The treatment group is further stratified into schools receiving "hard" intervention and "soft" intervention, differentiated by the intensity and type of teacher training received to facilitate the HTHT application in practice. Figure 2 illustrates the key study design and treatment conditions; additional details describing the differences in "hard" and "soft" interventions and randomization process are described further in the methodology section.

Figure 2. HTHT Intervention: Study Design and Treatment Conditions



Note: The study incorporates three groups: (1) Control, (2) "Soft" Intervention, and (3) "Hard" Intervention. Treatment conditions are only applied to teachers. For teachers in the Control group, there is no HTHT intervention (i.e., teachers follow standard educational activities). For teachers in the "Soft" and "Hard" Intervention groups, teachers receive training with workshop and mentoring. For teachers in the "Hard" Intervention group, they also receive classroom observation with feedback; teachers in the "Soft" Intervention group do not receive individualized feedback on their classroom teaching.

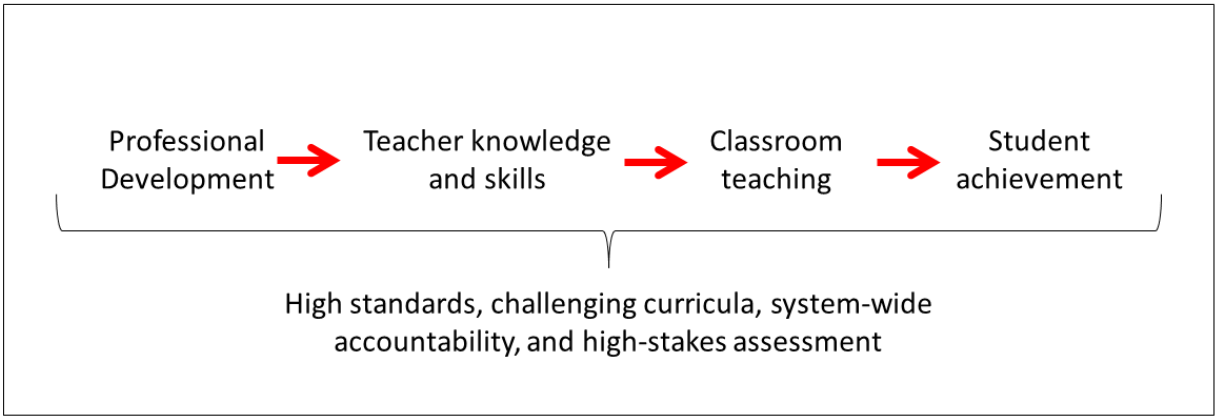
II. Conceptual Model and Study Design

Most models of effective PD programs that affect student achievement follow the path diagram illustrated in Figure 1 (Yoon et al., 2007; Blank, 2009). The effect of PD intervention on student achievement is mediated by teacher knowledge and skills and practices of classroom teaching; if one link is weak or missing, then increased student learning may not occur. PD programs must be of high quality in their theory of action, planning, design, and implementation. Teachers must also be motivated and be equipped with skills to apply the PD training to classroom teaching. Other frameworks such as Pianta (2011) and Allen, Pianta, Gregory, Mikami, and Lun (2011) characterize the mediators in Figure 1 as "teacher-student interactions."

Using the traditional PD evaluation framework illustrated in Figure 3, the fundamental task of the study design is to take into account the mediator and estimate both direct and indirect effects of PD intervention on student achievement.

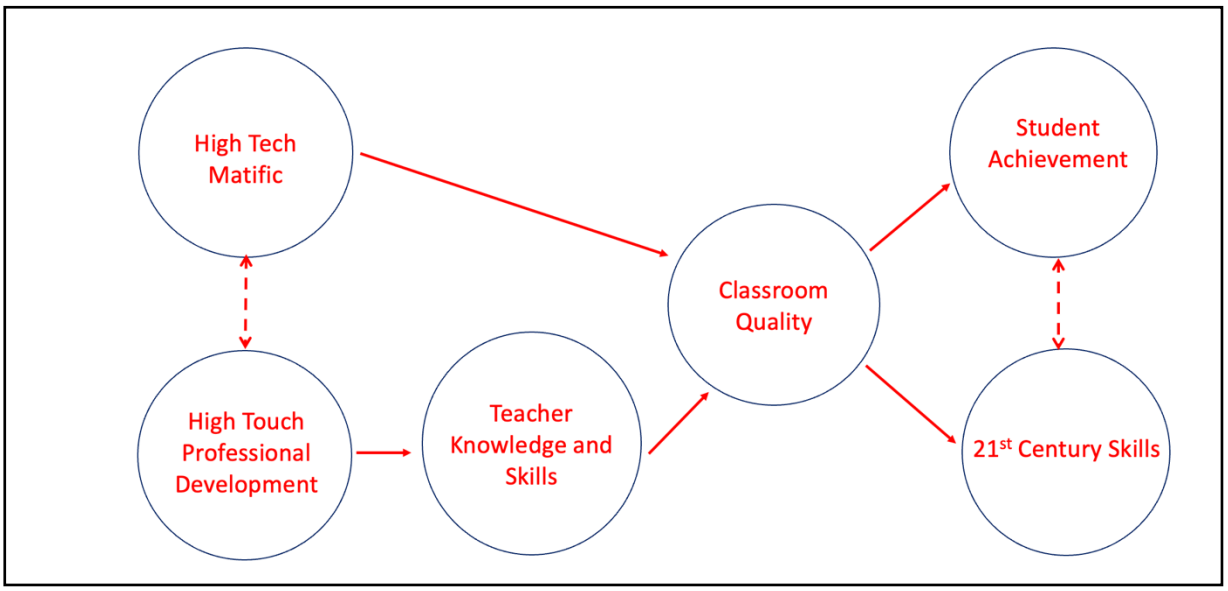
We use this evaluation design to examine the impact of the HTHT learning model on teaching effectiveness and students learning outcomes. See Figure 4.

Figure 3. Effect of professional development on student achievement



Note: Figure modified from Yoon et al. (2007)

Figure 4. Hypothesized Impact of the HTHT Learning Paradigm: Path Diagram



Note: Path diagram in Figure 2 reflects the conceptualized interaction between the "High Touch" and "High Tech" components that influence improvement in teacher knowledge and skills that thereby improves classroom quality. These frameworks imbed the PD intervention model in Figure 1, which subsequently lead to improvements in student achievement, moderated by changes in students' 21st century skills. This project aims to capture the direct and indirect effects modeled in this path diagram.

2.2. Study Design

Study Conditions: "Hard" Intervention, "Soft" Intervention, and "Control" Groups. An experimental study design will be used to stratify the HTHT learning model of schools/teachers into three study conditions varying in the level of teacher training (see Appendix Part 5 for specific differences of study conditions for "hard" and "soft" treatment groups):

(1) **"Hard" intervention group:** Teachers receive training through workshops, instructional materials, and mentoring, including classroom observations and feedback (two 1-hour classroom observations that include written feedback report and meeting with teachers for comments).
(2) **"Soft" intervention group:** Teachers receive training through workshops, instructional materials, and mentoring, but not the classroom observations and feedback components.
(3) **Control group:** Teachers without HTHT learning interventions; educational activities following standard and normally accepted practices.

As such, schools in the control group receive standard and accepted instructional approaches. Schools in the intervention groups receive the HTHT learning model, including workshops, instructional materials, and mentoring. Between the "hard" and "soft" intervention groups, schools receiving the "hard" intervention will also receive individualized classroom observations and feedback which are specifically targeted for teacher-specific professional development following the HTHT approach. All schools, including schools in the "control" group will have access to Matific, a computer-assisted learning platform already accessible to schools in Uruguay.

2.3. Sampling and Group Assignment

Random assignments to "hard", "soft" and "control" groups were determined using a sampling model based on the following criteria:

(1) **School Type:** (1) Regular ("Urban"); (2) Appender; and (3) Full-time schools
(2) **Sociocultural Context:** School's sociocultural context, stratified based on quntile distribution

Schools in the "soft" intervention group were oversampled at a 2:1 ratio, relative to the "hard" intervention and "control" group schools. The oversampling of "soft" intervention schools was intended to facilitate voluntary recruitment of schools and to allow comparisons with control and hard intervention groups. A total of 116 schools (including 176 teachers, and 3,880 students) were initially targeted for recruitment. The target recruitment by school classification are shown in the Appendix (see Supplemental Appendix Table A1).

2.4. Power and Sample Size Calculation

Power calculation was based on standard crtiteria (alpha = .05) and moderate effect sizes (standardized effect of .40 for students and teachers). These results yield power estimates of .851 for teachers and .958 for students, respectively, using conservative comparison settings between

the control and "hard" intervention groups. For student-level power calculation, cluster randomized design (CRD) was used, specifying clustering effects with intraclass correlation of .20 to allow within school variation. Power calculation was estimated using Stata using the power suite of commands.

Standard bias analyses (i.e., balance or equivalence test) were used to check for comparability of teacher characteristics and student socio-demographic characteristics between groups. Assuming a randomization process to assign teachers (schools) to intervention and control groups, we aim to measure outcomes associated with teachers using the conceptual model in Figure 4. The "High Tech" component of the HTHT will be delivered through the Matific within Plan Ceibal's digital learning platform. Recruitment of schools and randomization will be undertaken by Plan Ceibal, in partnership with ECA and IDB. Schools are not randomly selected to take part; but they are willing to participate in HTHT intervention voluntarily.

III. METHODS

Data were collected from two-time intervals (pre-test and post-test) to evaluate changes in teacher behavior and knowledge and their influence on teaching effectiveness. The evaluation strategy, analytic plan, and manuscript development (technical report and peer-reviewed article for publication) were be led by ECA, in partnership with Plan Ceibal and the Inter-American Development Bank (IDB).

3.1. Study Measures

Study measures are summarized in Table 1, stratified by constructs targeting teachers and students.

Table 1. Study Measures and Outcomes: Surveys and Assessments

Target	Type	Domain	Subdomain Measures
Student	Survey	Perceptions of Learning Environment and Education	Study Habits
			School Environment
			Classroom Environment
			Teaching Quality
		Mathematics learning and performance	Mathematics learning
			Preference toward mathematics
			Independent study of mathematics
		Access to technology and adaptive learning	Access to technology
			Anticipated effectiveness of HTHT
		21st century skills	GRIT (i.e., perseverance of effort and consistency of interest)
			Collaboration
		Basic Information	Demographic characteristics
	Assessment	Mathematics Performance	SEA+ mathematics assessment
Teacher	Survey	Teacher Self-Efficacy	
		Perceptions on the HTHT	
		Infrastructure and curriculum	School data systems
			Individualized learning
			Project-based learning
			Personalized technology use
		Student awareness of goals and progress	
		Competency-based learning	
		student choice and engagement	
		School professional environment	
		Student respect and motivation	
		Access to technology	

3.1.1. Study Measures – Teachers

Outcomes corresponding to teacher and students will be based on validated scales with sound psychometric property to ensure robust measurement. Scales will be examined in collaboration with the project team and partner collaborators. We plan to administer two subset measures from teachers (perceptions of HTHT and access to technology) to school principals, intended for school-level measures.

The following outcome measures will be gathered as part of an online survey completed by participating teachers. These surveys were selected from validates scales and measures used in prior online adaptive learning platforms. Replicating the use of these scales allow comparison of results obtained in the Uruguayan setting to other measures collected in the United States, Vietnam, and in other international contexts.

1. Teacher Self Efficacy
2. Perceptions on the "High Touch High Tech" intervention
3. Perceptions on infrastructure and curriculum
    a. School data systems
    b. Individualized learning
    c. Project-based learning
    d. Personalized technology use
4. Perceptions of student awareness of goals and progress
5. Competency-based learning
6. Perceptions of student choice and engagement
7. School professional environment
8. Student respect and motivation
9. Access to technology

Details of these existing measures are available in Adelfinger et al. (2017) and in Pane et al. (2015). The appendix includes the complete set of items administered to participating teachers.

The 24-item "Teacher's Sense of Efficacy" Scale (TSES; Tschannen-Moran & Woolfolk, 2001) was used to measure teacher efficacy and its subdomains using a 9-point rating scale: (1) student engagement (8 items), (2) instructional strategies (8 items), and (3) classroom management (8 items). "Anticipated effectiveness and perceptions of the adaptive learning platform" (10 items) were taken from Adelfinger et al (2017) and measures teacher views on the HTHT using a 5-point rating scale, focusing on student learning, preparation, ability to solve critical problems, and school community.

The ensuing measures relating to learning, access to data, and the school environment were taken and adopted from Pane et al. (2015). "Perceptions of quality and utility of data and school data systems" is a 6-item scale measuring based on a 5-point rating scale measuring teacher's access, use, and understanding of detailed and actionable student assessment data. "Views on technology to support individualized learning" is a 3-item scale based on a 5-point rating scale, measuring teachers' views on student's use and access to individual learning progress using technology. "Access to high-quality technology and non-technology-based curriculum" is a 14-

item scale (with 7 items for technology-based curriculum and 7 separate items for non-technology-based curriculum) based on a 5-point rating scale measuring access to technology or non-technology curriculum that address and support readily accessible learning needs of students.

In "curriculum and instruction" there are three subdomains based on a 5-point rating scale: (1) student awareness of goals and progress (3 items; measuring student awareness and teacher strategies on goals and progress of assignments and activities), (2) competency-based learning (4 items; measuring students' understanding of topics and competency-based progress), and (3) student choice and engagement (5 items; measuring student's individualized choice and engagement for instructional materials and topics of learning).

For "professional environment, respect, and motivation", there are three subdomains based on a 5-point rating scale: (1) student respect (2 items; measuring respect for peers and school staff), (2) student motivation (2 items; measuring student motivation for achievement and family support), (3) teacher collaboration (2 items; measuring peer collaboration toward improving student learning), and (4) administrator support for students and teachers (3 items; measuring administrator support for teachers, student learning, and trust of teachers).
Finally, items also measured teachers' use and access to different technology devices (smart phone, computer, tablet, video game system, and television) and Internet access. We also collected information on teacher's grade level, subjects taught, educational background, and years of teaching.

3.1.2. Study Measures – Students

Assessment and Survey Development. The baseline and endline surveys consist of three main sections, including an assessment component:

1. **Student background:** The first section includes demographic information, parents' educational background, and socioeconomic status (SES). Plan CEIBAL will provide students' background information collected by schools (i.e., school-level aggregated indicators) and from parents (i.e., a separate survey for parents). These variables would be used as control covariates.

2. **Mathematics learning attitude, perceptions of teachers and teaching quality, and 21st century skills:** The second section asks students' mathematics learning experiences and perceptions of their teachers. The ensuing measures relating to 21st century skills in the HTHT project-based learning (i.e., critical thinking/problem solving, collaboration, and communication) were taken from the 8-item "GRIT" Scale (Alan et al., 2019) and the 8-item on collaboration skills with others (Musa et al., 2012). Grit has been shown to be highly predictive of educational outcomes at the primary level (Alan et al., 2019) and the GRIT scale was adapted and piloted in Uruguay previously.
    - Study habits
    - Perceptions of school environment
    - Perceptions of classroom environment
    - Perceptions of teaching quality
    - Mathematics learning

- Mathematics preference
- Independent study of mathematics
- Access to technology and adaptive learning
- Anticipated effectiveness of HTHT
- Grit (i.e., perseverance of effort and consistency of interest)
- Collaboration

3. **Mathematics assessment:** The final section consists of mathematics assessment, administered using an adaptive assessment, aligned with the Uruguayan national curriculum and a reflective journal:

    - **Uruguayan mathematics assessment (SEA+ assessment).** Adaptive assessment from SEA+ will be used to estimate mathematics performance, aligned with the Uruguayan national curriculum

Cognitive interviews with subset of targeted students and teachers were conducted in March 2021 to refine the surveys, examining item clarity and response process. During Year 1, baseline teacher survey and student assessments were administered in April 2022. Teacher surveys and student measures (survey and assessment) were simultaneously administered to both "intervention" and "control" schools, to account for potential bias due to seasonality effects.

3.3. Empirical strategy

We examine descriptive statistics of responses obtained from the participating teachers. First, teachers' responses were stratified and compared between schools receiving the HTHT intervention and schools without the intervention. Statistical analyses included bivariate comparisons based on *t*-tests and χ^2 tests examining differences in means and proportions, respectively. Second, to examine the effect of the intervention between comparison schools (treatment and control), we used standard growth curve model based on mixed-effects (or multi-level) regression. Analyses controlled for teacher and student background characteristics.

Given the evaluation adopts clustered Randomised Controlled Trial (RCT) design, the impact of the HTHT intervention on student's academic outcomes were determined by fitting the following model:

$$Y_{ij}^{Post} = \alpha + \beta.Treat_j + \gamma.Y_{ij}^{Pre} + \delta.C_{ij} + \varepsilon_{ij}$$

Y^{post} = Student's post-test scores
Y^{pre} = Student's baseline test scores
Treat = A binary variable indicating whether the student was enrolled in a treatment or control school (0= control; 1= treatment)
C = Baseline (pre-treatment) controls for other student characteristics
ε = Error term
i = student i
j = school j

Controlling for prior achievement improves statistical power and account for differences in prior achievement between treatment and control groups. To allow for the fact that the program is a school level intervention and there is clustering of pupils within schools, all standard errors will be clustered at the school level. The coefficient of interest is β, which how whether or not there is a positive effect of the HTHT intervention. Additionally, mixed-effects (or multi-level) model including school fixed or random effects were used to estimate school-level variance (e.g., school resources). In principle, any analysis including student achievement needs to reflect the hierarchical nature of the data, where students are nested within schools and classrooms (See Goldstein, 1995; Raudenbush & Bryk, 2002). School fixed effects that account for variation in unobserved, time-invariant school-level characteristics, were estimated using the dummy variable-based approach added to the OLS model.

We fit variations of this model that interacted the treatment dummy with students' Matific use (continuous or by categorical groups & individual or classroom levels) to understand whether the intervention was more effective for some sub-groups of students. We also conducted sub-group analysis with three student characteristics that we observe at baseline (i.e., gender, father's education level, and students' baseline test score quartile) to test for heterogeneous effects.

In addition, outcomes were linked between teacher and student outcomes to examine the translational impact (i.e., HTHT intervention on teachers → improvement in interactive classroom learning → improvement in student outcomes). To incorporate direct and indirect effects in the analysis, structural equation models were used. To help inform how teacher perceptions yield direct or indirect effects on student learning and their assessment performance, we merged data with the student scores, following the model in Figure 1 (Yoon et al., 2007). In addition, to evaluate the role of teacher effects on the overall student performance, structural equation modeling analyses will be used to examine factors that predict performance and outcomes. Details of the statistical approaches in these models are in Park et al. (2018) and in Park and Lee (2014).

3.4. Data Overview

3.4.1. Analytical Sample: Panel A and Panel B

Data were gathered from a total of 108 schools, 155 teachers, and 2,709 students: "Control" Group (n = 28 schools, 39 teachers, 711 students); "Soft"Intervention Group (n = 51 schools; 74 teachers; 1,233 students); and "Hard" Intervention Group (n = 29 schools; 42 teachers; 765 students). Response rate for participation across both baseline and post-intervention study phases was 74.5% (Baseline [T_0] = 2,709; Endline [T_1] = 2,018).

We note that sample sizes vary depending on unit of analysis (school, teacher, and students). To allow comprehensive analyses with HTHT data gathered, we created two panels.

- Panel A: n = 2,002 (without student survey), representing 72% of students with student characteristics (n = 2,776).
- Panel B: n = 1,550 (with student surveys), representing 56% of students with student characteristics (n = 2,776); Panel B also represents 77% of data from Panel A.

Panel A: Student Asessment and School Data (*without* Student Survey). In Panel A, we use data with combined student and school data sample of **2,002 students**. In this panel, we exclude data from student surveys, as there was respondent attrition in the survey sample. The corresponding data with teacher survey linked has sample size of 1,628 (teacher surveys in both survey waves: 114). See Appendix Table A2.

Panel B: Student Asessment and School Data (*with* Student Survey). In Panel B, we use data with combined student and school data sample of **1,550 students**. In this panel, we include data merged from student surveys, resulting in retention of 77% of data from Panel A (Panel B/Panel A = 1,550/2,002). The corresponding data with teacher survey linked has sample size of 1,228 (teacher surveys in both survey waves: 114).

Participants corresponding to Panels A and B use data from the following sources (Table 2). The corresponding sample sizes for each panel are provided in Appendix Table A2.

Table 2. Data for Panels A and B

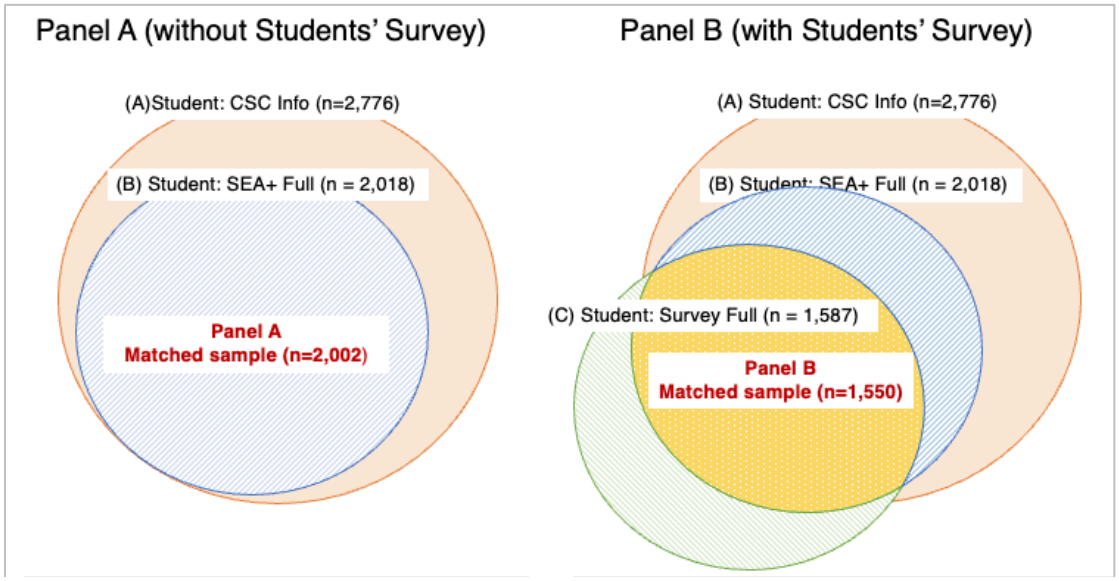
Data Source	Collection Methods	Panel A (n = 2,002)	Panel B (n = 1,550)
SEA+ Math Assessment	Computer-assisted test	X	X
School CSC (Living in capital, school quintile, school type)	Administrative data	X	X
Student CSC (Afampe, TUS, Matific use)	Administrative data	X	X
Student survey: Demographics (Age, gender)	Self-reported		X
Student survey: Family characteristics (Living with mother/father, parental education level, wealth)	Self-reported		X
Student survey: Non-cognitive outcomes (GRIT, collaboration, etc)	Self-reported		X

Attrition and Changes to Sample Sizes between Panel A (without student survey) and Panel B (with student survey). The rationale for creating two panels is the retention of records across linking different data gathered in this study. In the data with SEA+ mathematics assessment, we have 74% retention between basline (n = 2,709) and endline (n = 2,018). Linking this data with the student data results in sample size of 2,002 which was used in Panel A. Student survey had sample sizes of 2,238 and 2,414 between baseline and endline surveys, respectively. The combined data (to allow impact evaluation incorporating student characteristics), however, yielded sample size of 1,550, following combination of data with the assessment data. This consititutes the sample size in Panel B.

Overall, Panel A (without student surveys) represents 72% (n = 2,002) of students with CSC information (n = 2,776). Panel B (with student surveys) represents 56% (n = 1,550) of students with CSC information (n = 2,776) and 77% of data from Panel A.

Figure 4.1 graphically illustrates the sampling differences between Panel A and Panel B.

Figure 5. Analytical Sample – Panel A and Panel B



3.4.2. Randomization and Balance

Balance and Comparability of Groups – Students. Overall, there was comparable balance between control and treatment groups, across student characteristics: age, gender, parental education status (secondary education completion), living with mother, living with father, household size, household socioeconomic status, living in capital, school sociocultural status, and whether a student is a recipient of conditional case transfers (Afampe, TUS). Balance comparison between control and treatment groups are presented in Tables 3 and 4 for students in Table 5 and 6 for teachers.

We note modest differences in student characteristics between treatment and control groups for student living in capital (7% greater students living in capital in treatment) in Panel A and student gender (5% greater female students in control) in Panel B. In addition, mathematics performance as measured using SEA+ was significantly greater for students in the control group (7.64 SEA+ units for Panel A; 11.92 SEA+ units for Panel B). In general, these results demonstrate comparability across groups. Substantive factors with meaningful group differences (e.g., SEA+ mathematics performance, living in capital city, gender) were statistically adjusted in subsequent analyses.

Table 3. Baseline Statistics and Group Balance: Students – Panel A (*n* = 2,002)

Variable	(1) Total		(2) Control		(3) Treatment		(4) *t*-test
	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	Mean difference
Living in capital	2,002	-.20 (.01)	591	0.15 (.01)	1,411	0.22 (.01)	-0.07 (.01)***
School quintile	1,960	3.13 (.03)	574	3.06 (.06)	1,386	3.16 (.04)	-0.11 (.07)
Afampe	1,954	.52 (.01)	572	0.55 (.02)	1,382	0.51 (.01)	0.04 (.02)
TUS	1,954	.27 (.01)	572	0.28 (.02)	1,382	0.28 (.01)	-0.01 (.02)
Baseline math score (T0)	2,002	935.22 (1.58)	591	940.60 (2.90)	1,411	932.96 (1.88)	7.64 (3.46)**

Note: Difference in Control and Treatment groups examined using *t*-test. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4. Baseline Statistics and Group Balance: Students - Panel B (*n* = 1,550)

Variable	(1) Total		(2) Control		(3) Treatment		(4) *t*-test
	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	Mean difference
Age	1,550	10.23 (.01)	434	10.22 (.02)	1,116	10.23 (.01)	-0.005 (.03)
Female	1,503	0.49 (.01)	425	0.53 (.02)	1,078	0.47 (.02)	0.05 (.03)*
Mother secondary edu.	1,226	0.30 (.01)	338	0.31 (.03)	888	0.29 (.02)	0.02 (.03)
Father secondary edu.	1,129	0.18 (.01)	303	0.17 (.02)	826	0.18 (.01)	-0.009 (.03)
Living with mother	1,076	0.54 (.02)	292	0.90 (.02)	784	0.89 (.01)	0.009 (.02)
Living with father	1,080	0.89 (.01)	292	0.56 (.03)	788	0.54 (.01)	0.02 (.03)
Household size (< 4)	1,061	.11 (.01)	286	0.10 (.02)	775	0.11 (.01)	-0.006 (.02)
Household wealth ter.	1,045	2.00 (.03)	286	2.02 (.05)	759	2.00 (.03)	0.03 (.06)
Living in capital	1,550	.21 (.01)	434	0.15 (.02)	1,116	0.23 (.01)	-0.07 (.02)**
School quintile	1,516	3.20 (.04)	420	3.20 (.07)	1,096	3.19 (.04)	0.001 (.08)
Afampe	1,510	0.50 (.01)	418	0.50 (.02)	1,092	0.50 (.02)	0.001 (.03)
TUS	1,510	0.26 (.01)	418	0.24 (.02)	1,092	0.27 (.01)	-0.03 (.03)
Baseline math score (T0)	1,550	942.15 (1.81)	434	950.73 (3.43)	1,116	938.81 (2.12)	11.92 (4.02)**

Note: Difference in Control and Treatment groups examined using *t*-test. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Balance and Comparability of Groups – Teachers. Table 5 shows the baseline descriptive statistics for participating teachers (n = 114) in the control and in treatment groups, including a comparison by school quintile, school type, teacher age, educational level, experiences, and instructional preparation time. There were modest differences in teacher education level (0.47 years greater in control group).

In addition, Table 6 shows the baseline and endline differences in teacher outcomes, including pre-post differences (comparison of mean difference), demonstrating significant changes across self-efficacy and other instructional domains. In the treatment group, the differences were significant improvements across all (excluding access to technology), whereas for the control group, there was significant decrease for technology curriculum access, non-technology curriculum access, and professional environment. Table 7 shows the intervention effects using mixed-effects regression, taking into account the baseline and post-intervention changes.

Table 5. Baseline Statistics and Group Balance: Teachers (*n* = 114)

Variable	(1) Total		(2) Control		(3) Treatment		(4) *t*-test
	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	Mean difference
School quintile	114	3.04 (.13)	26	3.12 (.26)	88	3.01 (.15)	0.10 (.31)
School type	114	2,85 (.13)	26	2.65 (.27)	88	2.91 (.15)	-0.26 (.31)
Teacher's age	114	43.15 (.81)	26	44.15 (1.94)	88	42.85 (.88)	1.30 (1.92)
Teacher's highest education level	114	2.21 (.07)	26	2.58 (.26)	88	2.10 (.05)	.47 (.17)**
Teacher's experience (total)	114	16.11 (.81)	26	17.42 (1.93)	88	15.73 (.89)	1.70 (1.94)
Teacher's experience (current)	114	5.32 (.45)	26	6.23 (1.03)	88	5.06 (.50)	1.17 (1.07)
Teacher's instruction prep. time	114	9.32 (.64)	26	10.19 (2.21)	88	9.06 (.52)	1.14 (1.53)

Note: Difference in Control and Treatment groups examined using *t*-test. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6. Descriptive Statisics of Outcome Measures: Teachers

Outcome measure	Baseline Mean (SD)		Endline Mean (SD)		Pre-Post Difference	
	Control	Treatment	Control	Treatment	Control	Treatment
Teacher Self-Efficacy (standardized)	4.01 (.49)	4.02 (.50)	4.07 (.67)	4.08 (.42)	.06**	.06***
Teacher SE: Engagement	4.10 (.61)	4.05 (.53)	4.17 (.62)	4.11 (.47)	.07**	.06***
Teacher SE: Instruction	3.89 (.51)	3.92 (.53)	3.98 (.73)	4.06 (.43)	.09**	.13***
Teacher SE: Management	4.05 (.57)	4.09 (.55)	4.08 (.75)	4.08 (.55)	.03	.006
HTHT Perception	3.99 (.41)	3.99 (.59)	4.06 (.73)	4.25 (.58)	.10**	.25***
School Data System	3.75 (.61)	3.68 (.86)	3.75 (1.1)	4.14 (.58)	.08	.45***
Personalised Learning	3.88 (.79)	3.62 (.90)	3.90 (.52)	3.87 (.69)	.05*	.25***
Technology Curriculum Access	3.59 (.71)	3.56 (.95)	3.17 (1.41)	3.90 (.92)	–.42***	.34***
Non-Tech Curriculum Access	3.49 (1.10)	3.51 (.96)	3.24 (1.33)	3.55 (.89)	–.26**	.04*
Curriculum Instruction	3.82 (.49)	3.74 (.67)	3.94 (.38)	3.96 (.49)	.11***	.22***
Professional Environment	4.25 (.57)	4.20 (.54)	3.94 (.74)	4.22 (.48)	–.31***	.03**
Access to Technology	2.41 (.35)	2.44 (.38)	2.40 (.38)	2.42 (.38)	–.01	–0.02**

Note: Pre-post difference based on t-test. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7. Intervention Effect on Teachers: Standardized Coefficients from Mixed-Effects Regression

Outcome measure	Treatment Effect (SE)	*p*-value
Teacher Self-Efficacy (standardized)	0.34 (0.21)	$p > 0.1$
Teacher SE: Engagement	–0.16 (0.09)	$p > 0.05$
Teacher SE: Instruction	–2.06*** (0.18)	$p < 0.001$
Teacher SE: Management	–1.57*** (0.18)	$p < 0.001$
HTHT Perception	2.36*** (0.29)	$p < 0.001$
School Data System	0.73** (0.23)	$p < 0.01$
Personalised Learning	–0.94*** (0.15)	$p < 0.001$
Technology Curriculum Access	0.12 (0.22)	
Non-Tech Curriculum Access	1.15** (0.18)	$p < 0.01$
Curriculum Instruction	–0.18 (0.10)	
Professional Environment	0.96** (0.19)	$p < 0.01$
Access to Technology	–0.44** (0.12)	$p < 0.01$

Note: Treatment Effect is the mixed-effect regression coefficient estimating difference between treatment and control groups, taking into account baseline and post-intervention changes.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

3.4.3. Attrition Aanalysis

Response rate for participation across both baseline and post-intervention study phases was 69.3% in Panel B. The results of attrition analysis are presented in Table 8 (students) and Table 9 (teachers).[1] We note that certain analyses use varying sample sizes due to participant (student or teacher) non-reponse between baseline and post-intervention data collection phases.

In the student analysis for attrition, student characteristics no longer significantly influenced attrition, except student's treatment status (see result [5] in Table 8). There was higher attrition among students who participated in HTHT intervention than control group. In the teacher analysis for attrition, teacher characteristics did not significantly influence attrition; however, some school characteristics (school quintile and school type) remained to influence attrition in the survey participation (see result [3] in Table 9).

[1] Attrition analysis for Panel A (n=2,002; without students' survey) is not available because the initial sample size was 2,018 for those who participated in the baseline SEA+ Math assessment and the attrition rate is less than 1%.

Table 8. Relationship between Survey Attrition and Baseline Characteristics: Student – PANEL B (*n* = 1,550)

114/Variables	= 1 if missed in the endline math tests and survey				
	(1)	(2)	(3)	(4)	(5)
	Treatment	Adjusted	Adjusted	Adjusted	Adjusted
	Odds ratio	Odds ratio	Odds ratio	Odds ratio	Odds ratio
	(SE)	(SE)	(SE)	(SE)	(SE)
HTHT treatment	1.61***	1.85***	1.85***	2.06***	2.28***
	(0.18)	(0.22)	(0.23)	(0.33)	(0.48)
SEA+ Baseline math		1.00**	1.00	1.00	1.00
score		(0.00)	(0.00)	(0.00)	(0.00)
Female			0.90	0.85	1.04
			(0.09)	(0.09)	(0.14)
Age			1.39***	1.21*	1.22
			(0.13)	(0.13)	(0.18)
Afampe			1.16	1.07	0.96
			(0.14)	(0.14)	(0.18)
TUS			1.00	0.97	0.91
			(0.13)	(0.15)	(0.19)
School quintile			1.06	1.06	1.10*
			(0.04)	(0.05)	(0.06)
Mother completed					0.92
secondary					(0.16)
Father completed					1.05
secondary					(0.21)
Living with father					0.97
					(0.14)
Living with mother					1.10
					(0.27)
Household size					1.00
					(0.23)
Household wealth					0.95
index					(0.06)
Constant	0.31***	1.25	0.02***	0.05**	0.03*
	(0.03)	(0.83)	(0.03)	(0.07)	(0.05)
Observations	2,238	2,177	2,058	1,733	1,143

Note: Robust standard errors in parentheses are clustered at the level of class-by-school.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9. Relationship between Survey Attrition and Baseline Characteristics: Teacher

Variables	= 1 if missed in the endline math tests and survey		
	(1)	(2)	(3)
	Treatment	Adjusted	Adjusted
	Odds ratio	Odds ratio	Odds ratio
	(SE)	(SE)	(SE)
HTHT treatment	0.71	0.71	0.66
	(0.27)	(0.30)	(0.28)
School quintile 1 (*ref*)	-	-	-
	-	-	-
School quintile 2		0.45	0.52
		(0.33)	(0.38)
School quintile 3		0.18**	0.22*
		(0.16)	(0.20)
School quintile 4		0.44	0.52
		(0.36)	(0.45)
School quintile 5		0.21*	0.25
		(0.19)	(0.24)
School type: urban (*ref*)		-	-
		-	-
School type: rural		1.03	1.02
		(0.58)	(0.57)
School type: apprender		0.10**	0.12**
		(0.09)	(0.12)
School type: full-time		0.39	0.41
		(0.27)	(0.29)
School type: practice		0.73	0.76
		(0.41)	(0.43)
Teacher's edu: magisterio (*ref*)			-
			-
Teacher's edu: otra carrera docente			0.60
			(1.23)
Teacher's edu: otros estudios terciarios			1.11
			(1.16)
Teacher's edu: Universidad (grado)			1.34
			(1.51)
Teacher's edu: Universidad (postgrado)			-
			-
Teacher's experience in current school			1.00
			(0.04)
Teacher's preparation time for instruction			1.01
			(0.03)
Constant	0.58*	2.95	2.38
	(0.19)	(2.72)	(2.57)
Observations	165	165	162

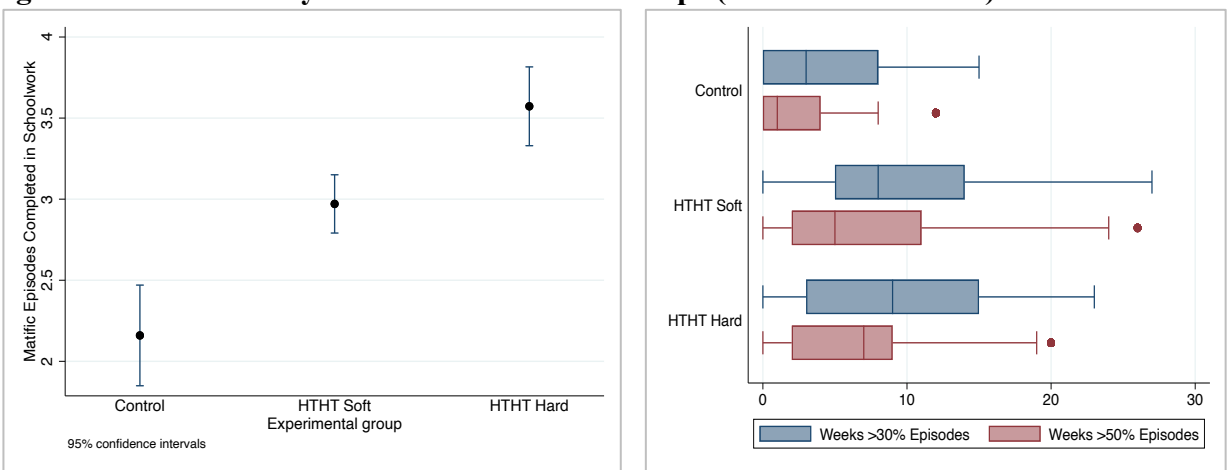
Note: Robust standard errors in parentheses are clustered at the level of class-by-school.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

IV. Results

4.1. HTHT Intervention and Students' Matific Use

HTHT increased the individualized use of computer-assisted learning by 49%. Students in HTHT intervention group ("Soft" and "Hard" treatments) had significantly greater completion rates for Matific episodes, as visually shown in Figure 7. Students in treatment group completed 1.49 times more Matific episodes than those in control group (control=15.81/ treatment=31.93). Between students in the "Soft" and "Hard" Intervention groups, the latter completed 20% more Matific episodes. However, there were no significant differences at the aggregate classroom level (box plot in Figure 7), signaling variation in Matific usage was individualized at the student level between the control and intervention groups.

Figure 6. Matific Use by HTHT Intervention Groups ("Soft" and "Hard") – Panel A



Note: In the bar graph (right), Y-axis presents students' Matific episodes completion by intervals of 10 episodes (individual). Box plot (left) presents Matific episodes completed at the class level (group).

In each series of tables, we show connectivity to Matific (school work started/completed, homework started/completed, bonus started/completed, number of days connected, matific use by groups) showing student-level usage. In addition, we also present classroom-level Matific usage (weeks > 30% and 50% of class finished Matific episodes). Overall, results show significantly greater Matific use in the treatment group. When we look at the highest Matific usage group, those who completed more than 40 episodes, 30% of students in treatment group belong to the highest-usage group, whereas only 10.5% students in control group belong to this group (Table 10.2). Results also show significant association between greater Matific usage at school and holme with SEA+ score gains (4.3.1c). These results are consistent across both Panels A and B. We stratify matific usage by Panel A (see Table 10) and Panel B (see Table 11).

Table 10. Descriptive Trends – Panel A

10.1. Matific Use by HTHT Treatment Status

Matific Use	Control (n = 591)		Treatment (n = 1,411)		*p*-value
	Mean	(SD)	Mean	(SD)	
# Matific episodes: School Work Started	28.81	(55.92)	58.61	(61.60)	< .001
# Matific episodes: School Work Completed	**15.81**	**(31.63)**	**31.93**	**(31.61)**	**< .001**
# Matific episodes: Homework Started	9.41	(28.55)	11.26	(16.69)	< .1
# Matific episodes: Homework Completed	5.02	(11.30)	6.77	(10.13)	< .001
# Matific episodes: Bonus Started	3.11	(12.85)	12.64	(29.36)	< .001
# Matific episodes: Bonus Completed	1.50	(6.32)	7.14	(17.59)	< .001
# Days Connection to Matific	96.86	(92.07)	94.87	(86.08)	0.645
Matific Use group (Group 1~Group 4)	1.65	(1.02)	2.54	(1.19)	< .001
Weeks > 30% class finished Matific episode	3.66	(3.99)	9.40	(5.88)	< .001
Weeks > 50% class finished Matific episode	2.28	(3.02)	6.92	(5.47)	< .001

10.2. Matific Use Group Distribution by HTHT Treatment Status

Matific Use-Group	Group 1 (< 10 eps.)	Group 2 (< 20 eps.)	Group 3 (< 40 eps.)	Group 4 (40+ eps.)	Total
Control	386 65.31%	88 14.89%	55 9.31%	62 10.49%	591 100%
Treatment	393 27.85%	286 20.27%	307 21.76%	425 30.12%	1,411 100%
Total	779 38.91%	374 18.68%	362 18.08%	487 24.33%	2,002 100%

10.3. Matific Use: Association with SEA+ Gains

Matific Use	Control		Treatment	
	Correlation	*p*-value	Correlation	*p*-value
# Matific episodes: School Work Started	–.04	.367	.09	.001***
# Matific episodes: School Work Completed	–.03	.467	.09	< .001***
# Matific episodes: Homework Started	–.06	.165	.10	< .001***
# Matific episodes: Homework Completed	–.04	.312	.12	< .001***
# Matific episodes: Bonus Started	–.03	.435	.04	.111
# Matific episodes: Bonus Completed	–.03	.438	.05	.080
# Days Connection to Matific	–.04	.389	.09	.016**
Weeks > 30% class finished Matific eps	–.02	.644	.08	.002***
Weeks > 50% class finished Matific eps	–.02	.646	.11	< .001***

*** p<0.01, ** p<0.05, * p<0.1

Table 11. Descriptive Trends – Panel B

11.1. Matific Use by HTHT Treatment Status

Matific Use	Control (n = 434)		Treatment (n = 1,116)		*p*-value
	Mean	(SD)	Mean	(SD)	
# Matific episodes: School Work Started	30.07	(57.17)	62.08	(64.03)	< .001
# Matific episodes: School Work Completed	16.43	(29.88)	33.91	(34.00)	< .001
# Matific episodes: Homework Started	9.58	(31.82)	11.85	(17.27)	< .1
# Matific episodes: Homework Completed	4.97	(11.67)	7.15	(10.50)	< .001
# Matific episodes: Bonus Started	3.57	(14.67)	13.98	(31.73)	< .001
# Matific episodes: Bonus Completed	1.74	(7.17)	7.89	(18.95)	< .001
# Days Connection to Matific	101.25	(93.36)	98.76	(87.33)	0.621
Matific Use group (Group 1~Group 4)	1.67	(1.05)	2.60	(1.18)	< .001
Weeks > 30% class finished Matific episode	3.56	(4.22)	9.60	(6.02)	< .001
Weeks > 50% class finished Matific episode	2.36	(3.22)	7.14	(5.59)	< .001

11.2. Matific Use Group Distribution by HTHT Treatment Status

Matific Use-Group	Group 1 (< 10 eps.)	Group 2 (< 20 eps.)	Group 3 (< 40 eps.)	Group 4 (40+ eps.)	Total
Control	285 65.67%	58 13.36%	41 9.45%	50 11.52%	434 100%
Treatment	284 25.45%	228 20.43%	247 22.13%	357 31.99%	1,116 100%
Total	569 36.71%	286 18.45%	288 18.58%	407 26.26%	1,550 100%

11.3. Matific Use: Association with SEA+ Gains

Matific Use	Control		Treatment	
	Correlation	*p*-value	Correlation	*p*-value
# Matific episodes: School Work Started	–.07	.161	.08	.006***
# Matific episodes: School Work Completed	–.05	.311	.09	.003***
# Matific episodes: Homework Started	–.06	.223	.09	.004***
# Matific episodes: Homework Completed	–.03	.542	.10	< .001***
# Matific episodes: Bonus Started	–.04	.411	.05	.120
# Matific episodes: Bonus Completed	–.04	.413	.05	.092
# Days Connection to Matific	–.04	.366	.06	.035**
Weeks > 30% class finished Matific episode	–.05	.302	.09	.004***
Weeks > 50% class finished Matific episode	–.05	.254	.10	< .001***

*** p<0.01, ** p<0.05, * p<0.1

For more in-depth analysis examining the impact of Matific usage with SEA+ gains, we stratify students into quintiles of SEA+ score gains in Panel A. Using these quintiles representing five groups of increasing SEA+ score cohorts, we show SEA+ scores for different categories of Matific score use by HTHT treatment groups (see Table 12). We also examined teacher characteristics by class-level Matific use within HTHT intervention groups (see Table 13). In this analysis, we found that teacher age, teaching experience, and preparation time for instruction were factors contributring to grater Matific episore completion.

Predictor of Matific Use by schoolwork/homework/bonus. We further examined which factors affect students' more frequent use of Matific, measured by episodes completed, as part of schoolwork, homework, and bonus (voluntary activities initiated by students during spare time). For the use of Matific as schoolwork, students' participation in HTHT intervention and baseline math score were positive and statistically significant predictors for more frequent use of Matific, while students who are eligible for Afempe were likely to complete fewer Matific episodes than their peers (Table 14 and 15). For the use of Matific as homework and bonus, the students' baseline math score was a positive predictor, whereas students' eligibility for TUS predicted less use of Matific at home. Similarly, as shown in Table 16, HTHT intervention was a positive and statistically significant predictor for teachers' more frequent use of Matific in their classrooms. Teacher's age also has a modest effect on the frequent use of Matific at school.

Table 12: Descriptive Statistics: Matific Use by Student's SEA+ Baseline Score (Panel A)

12.1. Total (n=2,002)

Matific Use	Students' Baseline: Q1		Students' Baseline: Q2		Students' Baseline: Q3		Students' Baseline: Q4		Students' Baseline: Q5	
	M	SD	M	SD	M	SD	M	SD	M	SD
SEA+ Baseline score	845.92	(23.84)	891.39	(9.86)	925.78	(11.35)	971.54	(14.50)	1041.705	(38.24)
# Days Connection to Matific	**77.34**	**(78.33)**	**87.78**	**(85.03)**	**94.98**	**(85.08)**	**103.43**	**(91.60)**	**113.82**	**(94.37)**
Weeks > 30% class finished Matific episode	7.70	(5.73)	7.14	(5.90)	8.00	(6.00)	7.94	(6.08)	7.76	(6.23)
Weeks > 50% class finished Matific episode	5.45	(5.07)	5.09	(5.14)	5.76	(5.22)	5.72	(5.48)	5.72	(5.64)
# Matific episodes: School Work Started	**44.19**	**(60.98)**	**42.41**	**(54.08)**	**49.44**	**(57.78)**	**56.74**	**(66.24)**	**56.30**	**(66.31)**
# Matific episodes: School Work Completed	**22.18**	**(29.11)**	**22.63**	**(27.74)**	**27.33**	**(34.54)**	**30.89**	**(34.78)**	**32.83**	**(37.30)**
# Matific episodes: Homework Started	**9.52**	**(31.27)**	**8.92**	**(16.11)**	**9.85**	**(14.39)**	**11.25**	**(16.60)**	**14.06**	**(21.28)**
# Matific episodes: Homework Completed	4.55	(10.52)	5.17	(8.96)	5.99	(9.27)	6.94	(10.60)	8.64	(12.43)
# Matific episodes: Bonus Started	**6.58**	**(18.53)**	**7.65**	**(22.08)**	**8.34**	**(20.78)**	**12.63**	**(32.11)**	**13.94**	**(32.34)**
# Matific episodes: Bonus Completed	3.25	(9.94)	4.03	(12.94)	4.59	(12.38)	7.31	(19.21)	8.21	(19.44)

12.2. Treatment (n=1,411)

Matific Use	Students' Baseline: Q1		Students' Baseline: Q2		Students' Baseline: Q3		Students' Baseline: Q4		Students' Baseline: Q5	
	M	SD	M	SD	M	SD	M	SD	M	SD
SEA+ Baseline score	844.75	(25.39)	891.70	(9.59)	926.34	(11.31)	971.98	(14.75)	1041.76	(38.11)
# Days Connection to Matific	**73.31**	**(74.10)**	**91.22**	**(88.09)**	**98.36**	**(84.10)**	**99.78**	**(88.55)**	**113.83**	**(90.99)**
Weeks > 30% class finished Matific episode	8.95	(5.74)	8.86	(5.88)	9.69	(5.70)	9.69	(5.99)	9.87	(6.06)
Weeks > 50% class finished Matific episode	6.48	(5.24)	6.43	(5.33)	7.08	(5.23)	7.22	(5.66)	7.43	(5.86)
# Matific episodes: School Work Started	**48.72**	**(54.99)**	**52.28**	**(58.66)**	**57.17**	**(54.65)**	**66.47**	**(68.36)**	**69.82**	**(68.73)**
# Matific episodes: School Work Completed	**24.67**	**(28.62)**	**27.76**	**(29.52)**	**31.21**	**(30.00)**	**35.76**	**(33.89)**	**41.26**	**(38.25)**
# Matific episodes: Homework Started	7.94	(16.25)	9.74	(16.65)	11.49	(15.42)	12	(15.29)	15.55	(18.85)
# Matific episodes: Homework Completed	4.15	(9.40)	5.82	(9.31)	7.07	(9.83)	7.32	(9.25)	9.80	(11.93)
# Matific episodes: Bonus Started	7.78	(19.54)	9.97	(25.24)	10.77	(23.69)	16.85	(37.37)	18.58	(36.57)
# Matific episodes: Bonus Completed	4.09	(11.33)	5.28	(14.84)	5.95	(14.14)	9.74	(22.26)	11.13	(22.38)

12.3. Control (n=591)

Matific Use	Students' Baseline: Q1		Students' Baseline: Q2		Students' Baseline: Q3		Students' Baseline: Q4		Students' Baseline: Q5	
	M	SD	M	SD	M	SD	M	SD	M	SD
SEA+ Baseline score	849.27	(18.42)	890.64	(10.49)	924.29	(11.36)	970.61	(13.98)	1041.6	(38.64)
# Days Connection to Matific	**88.86**	**(88.70)**	**79.55**	**(76.95)**	**85.92**	**(87.41)**	**111.19**	**(97.66)**	**113.78**	**(101.26)**
Weeks > 30% class finished Matific episode	4.15	(3.93)	3.03	(3.43)	3.46	(4.15)	4.23	(4.39)	3.45	(3.93)
Weeks > 50% class finished Matific episode	2.51	(2.99)	1.88	(2.66)	2.23	(3.16)	2.53	(3.31)	2.25	(2.94)
# Matific episodes: School Work Started	31.25	74.31	18.82	(30.19)	28.72	(61.01)	36.05	(56.38)	28.86	(51.30)
# Matific episodes: School Work Completed	15.06	(29.46)	10.38	(17.81)	16.94	(42.92)	20.53	(34.50)	15.72	(28.58)
# Matific episodes: Homework Started	14.03	(54.87)	6.96	(14.62)	5.43	(9.95)	9.64	(19.06)	11.04	(25.31)
# Matific episodes: Homework Completed	5.66	(13.20)	3.63	(7.89)	3.09	(6.86)	6.13	(13.01)	6.27	(13.12)
# Matific episodes: Bonus Started	3.14	(14.88)	2.09	(9.45)	1.83	(5.34)	3.68	(11.82)	4.5	(18.07)
# Matific episodes: Bonus Completed	0.85	(2.67)	1.05	(5.37)	0.94	(3.20)	2.13	(7.91)	2.27	(8.81)

Table 13. Descriptive Statistics: Teachers' Characteristics by Class-level Matific Use within HTHT Intervention group (Panel A)

	Weeks in which at least 30% class finished a Matific episode			Weeks in which at least 50% class finished a Matific episode		
	HTHT+ Low Completion (Less than 10 weeks)	HTHT+ High Completion (10 weeks and above)	Diff (t-test)	HTHT+ Low Completion (Less than 10 weeks)	HTHT+ High Completion (10 weeks and above)	Diff (t-test)
	$n = 50$	$n = 35$		$n = 62$	$n = 23$	
Teacher's age	**42.04 (7.50)**	**44.25 (9.00)**	**2.22**	**42.20 (7.48)**	**44.96 (9.72)**	**2.74**
Teacher's education levels	2.1 (0.42)	2.1 (0.53)	0.01	2.08 (0.38)	2.17 (0.65)	0.09
Teacher's teaching experience	**4.42 (3.60)**	**5.97 (5.93)**	**1.55**	**4.87 (4.01)**	**5.57 (6.37)**	**0.69**
Teacher's preparation time for instruction	**8.68 (4.76)**	**9.62 (5.09)**	**0.95**	**8.53 (4.61)**	**10.52 (5.43)**	**1.99***
Teacher's self-efficacy	3.97 (0.53)	4.08 (0.43)	0.11	3.98 (0.53)	4.12 (0.36)	0.15
Teacher's HTHT perception	3.93 (0.71)	4.04 (0.52)	0.11	3.93 (0.66)	4.12 (0.55)	0.18
School data system	3.69 (0.90)	3.74 (0.79)	0.06	3.70 (0.83)	3.75 (0.19)	0.05
Tech personalization	3.56 (0.90)	3.79 (0.78)	0.23	3.60 (0.87)	3.80 (0.81)	0.19
Technology curriculum	3.64 (0.90)	3.51 (0.91)	-0.13	3.58 (0.95)	3.60 (0.77)	0.02
Curriculum instruction	3.70 (0.73)	3.83 (0.56)	0.17	3.74 (0.68)	3.80 (0.62)	0.06
Professional environment	4.23 (0.57)	4.14 (0.48)	-0.09	4.22 (0.55)	4.14 (0.49)	-0.08
Access to technology	2.34 (0.37)	2.54 (0.36)	0.19	2.39 (0.37)	2.55 (0.37)	0.17

Note: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Table 14. Predictor of Matific Use by schoolwork/homework/bonus (Categorical): Panel A

	(1)		(2)		(3)	
VARIABLES	Matific use: School		Matific use: Homework		Matific use: Bonus	
HTHT	0.92***	0.83***	0.32*	0.26	0.60***	0.26
Treatment	(0.20)	(0.27)	(0.18)	(0.24)	(0.14)	(0.24)
Baseline math	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
score	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Living in	-0.18	-0.27	0.13	0.08	0.12	0.08
capital	(0.28)	(0.27)	(0.25)	(0.24)	(0.28)	(0.24)
School quintile	0.02	0.05	0.06	0.04	0.03	0.04
	(0.07)	(0.07)	(0.07)	(0.07)	(0.06)	(0.07)
Eligible for	-0.09	-0.14*	-0.07	-0.08	-0.18***	-0.08
Afampe	(0.07)	(0.08)	(0.07)	(0.07)	(0.06)	(0.07)
Eligible for	-0.05	-0.01	-0.23***	-0.24***	0.09	-0.24***
TUS	(0.09)	(0.08)	(0.08)	(0.08)	(0.10)	(0.08)
Teacher's		-0.00		0.09		0.09
highest edu.		(0.11)		(0.09)		(0.09)
Teacher's		0.03		0.00		0.00
experience		(0.02)		(0.02)		(0.02)
Teacher's		0.02		0.00		0.00
class prep time		(0.01)		(0.02)		(0.02)
Constant	-0.17	-0.46	-0.50	-0.48	-0.86	-0.48
	(0.57)	(0.74)	(0.54)	(0.75)	(0.62)	(0.75)
Observations	1,954	1,591	1,954	1,591	1,954	1,591

Note: To ease the interpretation of the results, regression to the mean was used instead of multivariate logistic analysis.
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Table 15. Predictor of Matific Use by schoolwork/homework/bonus (Categorical): Panel B

	(1)		(2)		(3)	
VARIABLES	Matific use: School		Matific use: Homework		Matific use: Bonus	
HTHT	0.98***	0.88***	0.44**	0.40*	0.66***	0.40*
Treatment	(0.21)	(0.29)	(0.19)	(0.24)	(0.15)	(0.24)
Baseline math	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
score	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Living in	-0.22	-0.31	0.08	0.05	0.08	0.05
capital	(0.28)	(0.29)	(0.24)	(0.24)	(0.28)	(0.24)
School quintile	0.01	0.03	0.03	0.01	0.03	0.01
	(0.07)	(0.08)	(0.07)	(0.08)	(0.07)	(0.08)
Eligible for	-0.17**	-0.25***	-0.15**	-0.18**	-0.22***	-0.18**
Afampe	(0.08)	(0.09)	(0.07)	(0.07)	(0.07)	(0.07)
Eligible for	-0.03	-0.01	-0.20**	-0.21**	0.12	-0.21**
TUS	(0.10)	(0.11)	(0.08)	(0.09)	(0.10)	(0.09)
Teacher's		0.01		0.07		0.07
highest edu.		(0.12)		(0.11)		(0.11)
Teacher's		0.02		0.00		0.00
experience		(0.02)		(0.02)		(0.02)
Teacher's		0.01		-0.00		-0.00
class prep time		(0.01)		(0.02)		(0.02)
Constant	0.13	0.08	-0.41	-0.28	-0.85	-0.28
	(0.55)	(0.74)	(0.55)	(0.77)	(0.68)	(0.77)
Observations	1,510	1,238	1,510	1,238	0.66***	0.40*

Note: To ease the interpretation of the results, regression to the mean was used instead of multivariate logistic analysis.
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Table 16. Predictor of Teacher's Class-level Matific Use: Panel A – Logistic Regression

	Weeks in which at least 30% class finished a Matific episode (Low completion (0)/ High completion (1))			Weeks in which at least 50% class finished a Matific episode (Low completion (0)/ High completion (1))		
VARIABLES	Odds ratio	Odds ratio	Odds ratio	Odds ratio	Odds ratio	Odds ratio
HTHT Treatment	6.40***	6.01**	6.34**	5.46**	5.97*	6.65*
	(4.14)	(4.31)	(4.62)	(4.25)	(5.86)	(6.71)
Teacher's age	1.00*	1.00	1.00	1.00*	1.00	1.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
T's highest edu.		0.75	0.75		0.96	1.01
		(0.31)	(0.29)		(0.40)	(0.40)
T's teaching experience		1.05	1.05		1.01	1.00
		(0.05)	(0.05)		(0.05)	(0.05)
T's prep time for instruction		1.01	1.00		1.03	1.03
		(0.03)	(0.03)		(0.03)	(0.04)
T's self-efficacy (baseline)		1.71	1.74		1.95	2.09
		(0.73)	(0.78)		(0.89)	(0.94)
School in capital			0.50			0.55
			(0.27)			(0.36)
School socio-cultural quintile			1.01			0.89
			(0.17)			(0.18)
Constant	0.04***	0.01**	0.01**	0.02***	0.00***	0.00***
	(0.03)	(0.02)	(0.02)	(0.02)	(0.00)	(0.00)
Observations	111	111	111	111	111	111

Note: Class-level Matific Use category: 0 (less than 10 weeks in which at least 30% or 50% class finished a Matific episode); 1 (10 weeks or above weeks in which at least 30% or 50% class finished a Matific episode).
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

4.2. HTHT Intervention and Teachers' CREA use

There is a statistically significant difference in the number of days of connection to CREA between HTHT and control groups (Mean 101.5 vs. 77.23, $p < 0.05$). Strong correlations were observed between Teacher's CREA connection (# of log-in days) and student's Matific use in both HTHT and control groups. However, there is a weak correlation between Teacher's CREA connection and teacher characteristics (age, teaching experience, and instruction preparation time), except for teacher's education level. There is also a weak correlation between Teacher's CREA connection/activities and students' math score gains during the intervention period. See Table 17 and 18.

Table 17. Descriptive Statistics & *t*-test: Teacher's CREA use (Panel A)

CREA Use	Total (n=139)		Control (n = 39)		Treatment (n = 100)		*p*-value
	Mean	(SD)	Mean	(SD)	Mean	(SD)	
# Days Connection (Login) to CREA	94.69	(58.12)	101.5	(59.03)	77.23	(52.49)	< .05**
# of posted comments	28.00	(44.28)	29.36	(41.20)	24.51	(51.77)	> .1
# of actions	530.92	(617.47)	570.91	(646.99)	428.38	(528.23)	> .1
# of course content creations	211.09	(305.78)	228.71	(33.06)	165.90	(227.75)	> .1
# of created events	.417	(3.17)	.58	(3.73)	0	0	> .1
# of message sent	8.86	(22.14)	10	(25.56)	5.92	(8.84)	> .1
# of unique users	.96	(.20)	.95	(.22)	.97	(.16)	> .1
# of active users	.949	(.219)	.95	(.219)	.948	(.223)	> .1

Note: (1) Unique users: Users who logged in to the platform at least 1 day and performed at least one activity on it; (2) active users: Users who performed an action on at least 10 different days during the academic year.

Table 18. Association: CREA connection/use

18.1. Teacher's CREA connection & Student's Matific use

CREA Connection	Control		Treatment	
	Correlation	*p*-value	Correlation	*p*-value
# Days Connection to Matific	0.19***	<.01	0.06**	<.05
Weeks > 30% class finished Matific episode	0.53***	<.01	.0.30***	<.01
Weeks > 50% class finished Matific episode	0.37***	<.01	0.24***	<.01
# Matific episodes: School Work Started	0.30***	<.01	0.20***	<.01
# Matific episodes: School Work Completed	0.27***	<.01	0.22***	<.01
# Matific episodes: Homework Started	0.09**	<.05	0.08***	<.01
# Matific episodes: Homework Completed	0.11***	<.01	0.08***	< .01
# Matific episodes: Bonus Started	0.01	>.1	0.09***	< .01
# Matific episodes: Bonus Completed	-0.01	>.1	0.09***	< .01

18.2. Teacher's CREA connection & Teacher's characteristics

CREA Connection	Control		Treatment	
	Correlation	*p*-value	Correlation	*p*-value
Teacher's age	0.19***	<.01	-0.07*	<.1
Teacher's education level	-0.32***	<.01	0.45***	<.01
Teacher's total teaching experience	0.03	>.1	0.03	>.1
Teacher's teaching experience in current school	-0.001	>.1	0.002	>.1
Teacher's instruction preparation time	0.03	>.1	0.06**	<.05

18.3. Teacher's CREA use & Student's SEA+ math score gains

SEA+ math score gains	Control		Treatment	
	Correlation	*p*-value	Correlation	*p*-value
# Days Connection (Login) to CREA	0.03	>.1	0.01	>.1
# of posted comments	0.004	>.1	-0.03	>.1
# of actions	0.07	>.1	-0.006	>.1
# of course content creations	0.08	>.1	0.02	>.1
# of created events	-	-	-0.08***	<.01
# of message sent	-0.10	>.1	0.06	>.1

4.3. Main results: Effect of HTHT on Math Achievement

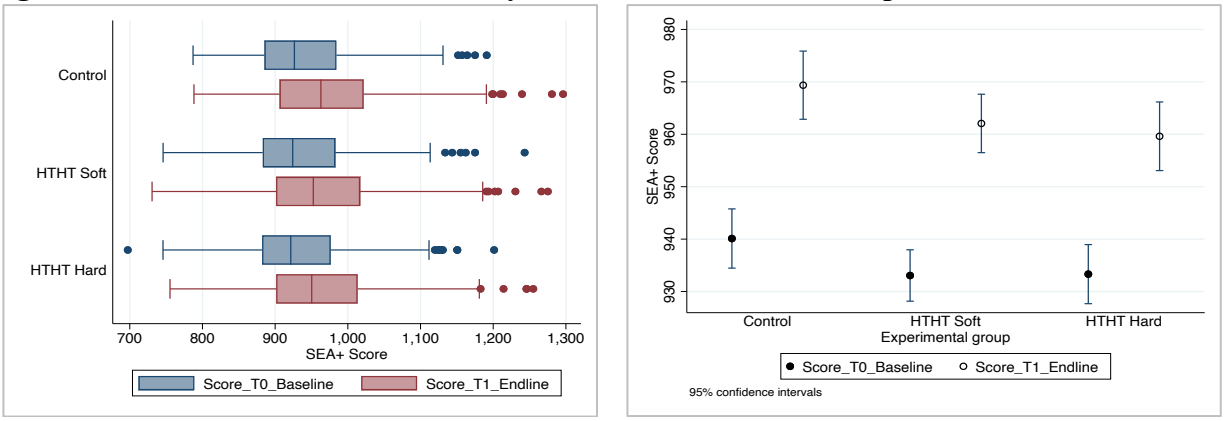
The offer of the intervention had a null effect (of about 0.02 SDs) on the math achievement of the average students, regardless of whether we account for students' performance at baseline (Table 19). The results remain consistent after controlling for student and school characteristics (models 2 and 5) and introducing school fixed effects to account for school-level variance (models 3 and 6). Figure 7 shows the visualization of descriptive statistics on average SEA+ math scores at baseline and endline by "hard" and "soft" HTHT intervention groups and control groups, which reveals no significant differences in math achievement across the three groups.

Table 19. Effect of HTHT Intervention on math achievement at endline – Panel A[2]

Variables	SEA+ Math IRT Score			SEA+ Math Standardized Score		
	(1)	(2)	(3)	(4)	(5)	(6)
HTHT Treatment	-1.51	-1.44	-1.76	-0.02	-0.02	-0.02
	(3.14)	(3.13)	(3.99)	(0.04)	(0.04)	(0.05)
Baseline math score	0.97***	0.95***	0.95***	0.84***	0.83***	0.82***
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Living in capital		-5.40*	-7.11		-0.07*	-0.09
		(3.16)	(4.76)		(0.04)	(0.06)
School quintile		1.83	1.37		0.02	0.02
		(1.23)	(1.17)		(0.02)	(0.01)
Eligible for Afampe		-4.62**	-5.05**		-0.06**	-0.06**
		(2.24)	(2.31)		(0.03)	(0.03)
Eligible for TUS		-2.79	-3.39		-0.03	-0.04
		(2.67)	(2.57)		(0.03)	(0.03)
Constant	60.33***	70.57***	82.24***	0.01	-0.00	0.04
	(18.80)	(20.12)	(14.44)	(0.03)	(0.07)	(0.06)
Observations	2,002	1,954	1,954	2,002	1,954	1,954
Number of Groups	-	-	99	-	-	99
R-squared	0.71	0.72	0.72	0.71	0.72	0.72
Fixed-effects	NO	NO	YES	NO	NO	YES

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure 7. SEA+ Mathematics Scores by HTHT Intervention Groups – Panel A



Note: Figures show box plot of SEA+ scores by data collection phase (top) and by error bar plots (bottom).

[2] See Appendix for the results of Panel B.

4.3.1. Heterogenous Effects by Matific Use

Student Individualized Use of Matfic and Interaction Effects. The null average effects, however, masked important heterogeneous impacts. We investigated whether the effect of HTHT intervention differed across students' use of Matific. We demonstrate this analytically in two ways: by accounting for students' Matific use (continuous, i.e., number of episodes completed) and interacting it with the treatment indicator (Table 20), and by interacting this indicator with categorical variables for students' Matific use (group 1 to group 4) (Table 21).

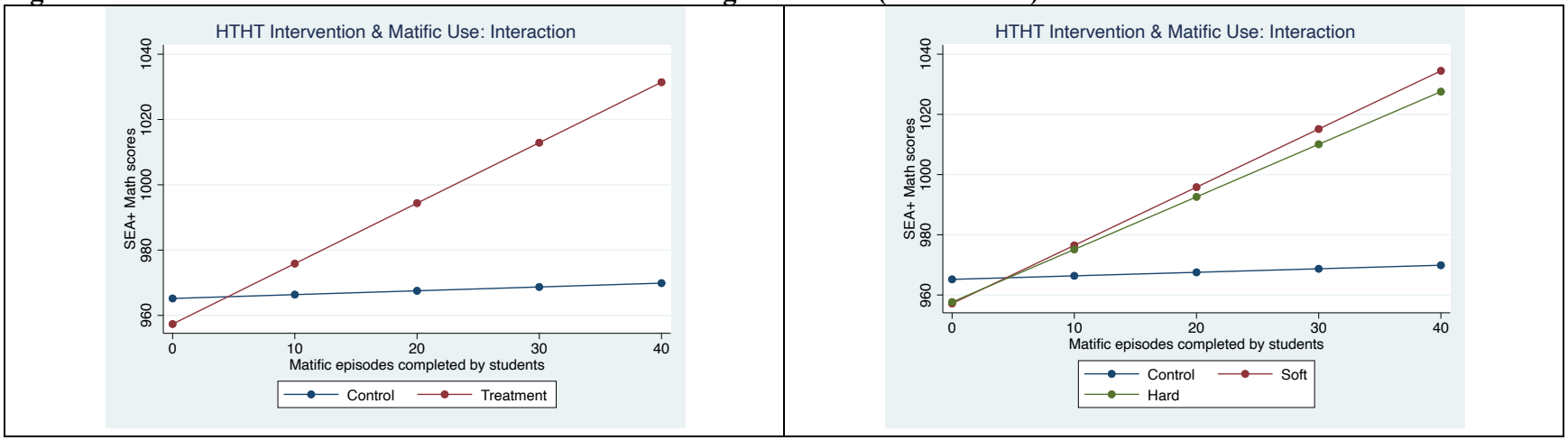
Notably, we found that the intervention had a medium-to-large positive effect for students with higher usage of Matific. Given that the coefficient of interaction term is not analogous to the coefficient in linear regression without interaction term, we show this graphically by plotting the treatment effects by students' Matific use (Figure 8). For student group who completed the highest number of Matific episodes (Matific use > 40), HTHT intervention had a significant effect of 0.76 SD, followed by 0.54 SD (Matic use 31~40), 0.33 SD (Matific use 21~30), 0.12 SD (Matific use 11~20), and -0.09 SD (Matific use 0~10). These findings maintain their statistical significance at the 10% level, after accounting for student and school characteristics. In addition, when we introduced interaction terms with categorical variables for Matific use (Table 21), it shows consistently significant results for students who completed higher number of Matific episodes (0.25 SD).

Table 20. The Interaction Effect of HTHT and Matific Usage at School (Continuous) on Students' Math Score : Panel A

Variables	SEA+ Math Score			SEA+ Math standardized Score		
	(1)	(2)	(3)	(4)	(5)	(6)
HTHT * Matific Use	1.80**	1.79**	1.73**	0.02**	0.02**	0.02**
	(0.83)	(0.85)	(0.85)	(0.01)	(0.01)	(0.01)
HTHT Treatment	-6.74*	-6.58*	-7.84*	-0.08*	-0.08*	-0.10*
	(3.84)	(3.85)	(4.46)	(0.05)	(0.05)	(0.05)
Matific Use	-0.35	-0.37	0.12	-0.00	-0.00	0.00
	(0.47)	(0.44)	(0.71)	(0.01)	(0.01)	(0.01)
Baseline math score	0.96***	0.95***	0.93***	0.83***	0.82***	0.81***
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Controls	NO	YES	YES	NO	YES	YES
Constant	69.28***	78.80***	91.27***	0.02	-0.00	0.03
	(18.86)	(19.99)	(14.60)	(0.03)	(0.07)	(0.07)
Observations	2,002	1,954	1,954	2,002	1,954	1,954
Number of groups	-	-	99	-	-	99
R-squared	0.71	0.72	0.72	0.71	0.72	0.72
Fixed Effects	NO	NO	YES	NO	NO	YES

Note: Robust standard errors clustered at the school level in parentheses. Matific episodes completion by intervals of 10 episodes.
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Figure 8. The Interaction Effect of HTHT and Matific Usage at School (Continuous) on Students' Math Score : Panel A



Note: Matific episodes completion by intervals of 10 episodes.

Group	HTHT intervention	Matific Use at school	Margin (Math IRT score)	SD	95% CI	
1	Control	0~10	965.19	(3.58)	958.16	972.21
2	Control	11~20	966.36	(6.87)	952.88	979.83
3	Control	21~30	967.53	(3.55)	940.96	994.10
4	Control	31~40	968.70	(20.55)	928.42	1008.99
5	Control	>40	969.88	(27.62)	915.73	1024.02
6	Treatment	0~10	957.35	(2.66)	952.13	962.57
7	Treatment	11~20	973.86	(3.79)	968.42	983.31
8	Treatment	21~30	995.38	(8.13)	978.44	1010.32
9	Treatment	31~40	1012.89	(12.73)	987.93	1037.85
10	Treatment	>40	1031.40	(17.38)	997.31	1065.5

Table 21. The Interaction Effect of HTHT and Matific Usage at School (Categorical) on Students' Math Score : Panel A

	SEA+ Math Score			SEA+ Math standardized Score		
	(1)	(2)	(3)	(4)	(5)	(6)
HTHT Treatment*Matific Use	-	-	-	-	-	-
Group 1 (*ref*)	-	-	-	-	-	-
HTHT Treatment*Matific Use	10.40	9.24	8.91	0.13	0.11	0.11
Group 2	(6.29)	(6.22)	(6.19)	(0.08)	(0.08)	(0.08)
HTHT Treatment*Matific Use	19.75**	15.62*	13.22*	0.24**	0.19*	0.16*
Group 3	(8.93)	(8.91)	(7.57)	(0.11)	(0.11)	(0.09)
HTHT Treatment*Matific Use	22.93***	22.32**	20.04**	0.28***	0.27**	0.25**
Group 4	(8.35)	(8.52)	(8.68)	(0.10)	(0.10)	(0.11)
Matific Use Group 2	-1.29	-2.51	-2.05	-0.02	-0.03	-0.03
	(4.57)	(4.49)	(5.11)	(0.06)	(0.06)	(0.06)
Matific Use Group 3	-2.06	-1.19	2.10	-0.03	-0.01	0.03
	(7.95)	(7.90)	(6.61)	(0.10)	(0.10)	(0.08)
Matific Use Group 4	-4.79	-5.86	0.69	-0.06	-0.07	0.01
	(7.12)	(7.11)	(7.82)	(0.09)	(0.09)	(0.10)
HTHT Treatment	-13.64***	-12.07***	-13.20***	-0.17***	-0.15***	-0.16***
	(4.23)	(4.40)	(4.80)	(0.05)	(0.05)	(0.06)
Baseline math score	0.96***	0.95***	0.93***	0.83***	0.82***	0.81***
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Controls	NO	YES	YES	NO	YES	YES
Constant	71.14***	80.07***	93.44***	0.02	0.02	0.04
	(18.82)	(19.89)	(14.53)	(0.04)	(0.07)	(0.07)
Observations	2,002	1,954	1,954	2,002	1,954	1,954
Number of groups	-	-	99	-	-	99
R-squared	0.72	0.73	0.73	0.72	0.73	0.73
Fixed Effects	NO	NO	YES	NO	NO	YES

Note: Matific Use category: 1 (less than 10 episodes); 2 (less than 20 episodes); 3(less than 40 episodes); 4 (40 episodes or above). *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Classroom-level Matific Use. The HTHT intervention, which includes teacher training through workshops, instructional materials, and mentoring, was implemented largely as intended. Teachers who participated in the intervention were encouraged to use the Matific ("High Tech") to provide more personalised learning for students combined with project-based learning activities ("High Touch") to deepen students' understanding of math concepts and foster their non-cognitive skills such as collaboration and GRIT. Implementing the HTHT learning models requires both minimum use of Matific (episodes completion) and teacher's dedication to allocate appropriate time, resources and attention to project-based learning offered by HTHT instructional materials. Figure 9 shows the distribution of classroom-level Matific use.

Table 22 presents the effect of HTHT on math achievement by class-level Matific Usage. For students who belong to the class using Matific more frequently, the HTHT intervention is associated with a math test score increase by 0.15 standard deviation (weeks in which 30% of the class finished a Matific episode) and 0.17 standard deviation (weeks in which 50% of the class finished a Matific episode) after accounting for student and school characteristics. When we investigate interaction between treatment and class-level Matific use, students in the class with more frequent use of Matific scored 0.19 SD higher in math assessment (Table 23).

Figure 9. Distribution of classroom-level Matific use

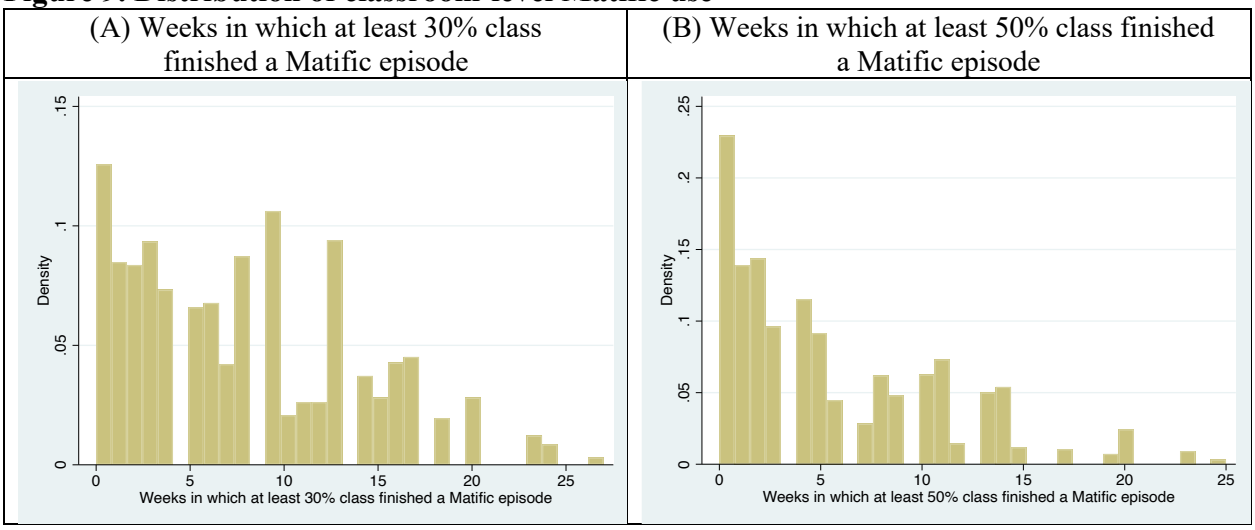


Table 22: The Effect of HTHT on Students' Math Score by Class-level Matific Usage - Panel A

22.1. Weeks in which at least 30% class finished a Matific episode

	SEA+ Math IRT Score				SEA+ Math Standardized Score			
VARIABLES	Less than 10 weeks		10 weeks and above		Less than 10 weeks		10 weeks and above	
	(1) Math Score	(2) Math Score	(3) Math Score	(4) Math Score	(5) Math Score	(6) Math Score	(7) Math Score	(8) Math Score
HTHT Treatment	-5.15	-5.56	9.11**	12.11***	-0.06	-0.07	0.11**	0.15***
	(3.68)	(3.49)	(3.60)	(4.03)	(0.05)	(0.04)	(0.04)	(0.05)
Baseline math score	0.96***	0.95***	0.97***	0.96***	0.84***	0.82***	0.85***	0.83***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)
Controls	NO	YES	NO	YES	NO	YES	NO	YES
Constant	65.89***	74.28***	47.64	68.37**	0.02	-0.04	-0.06***	-0.02
	(23.30)	(24.47)	(30.42)	(27.73)	(0.03)	(0.07)	(0.02)	(0.09)
Observations	1,394	1,357	608	597	1,394	1,357	608	597
R-squared	0.70	0.72	0.72	0.73	0.70	0.72	0.72	0.73

Note: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

22.2. Weeks in which at least 50% class finished a Matific episode

	SEA+ Math IRT Score				SEA+ Math Standardized Score			
VARIABLES	Less than 10 weeks		10 weeks and above		Less than 10 weeks		10 weeks and above	
	(1) Math Score	(2) Math Score	(3) Math Score	(4) Math Score	(5) Math Score	(6) Math Score	(7) Math Score	(8) Math Score
HTHT Treatment	-2.97	-2.93	9.47**	13.87**	-0.04	-0.04	0.12**	0.17**
	(3.46)	(3.32)	(3.63)	(6.58)	(0.04)	(0.04)	(0.04)	(0.08)
Baseline math score	0.97***	0.96***	0.95***	0.93***	0.84***	0.83***	0.83***	0.81***
	(0.02)	(0.02)	(0.04)	(0.04)	(0.02)	(0.02)	(0.04)	(0.04)
Controls	NO	YES	NO	YES	NO	YES	NO	YES
Constant	58.04***	66.11***	66.04	99.10**	0.02	-0.04	-0.08***	0.04
	(20.29)	(21.99)	(42.67)	(36.44)	(0.03)	(0.07)	(0.01)	(0.08)
Observations	1,610	1,569	392	385	1,610	1,569	392	385
R-squared	0.71	0.72	0.71	0.72	0.71	0.72	0.71	0.72

Note: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Table 23. The Interaction Effect of HTHT and Class-level Matific Use (Categorical) on Students' Math Score : Panel A

Variables	SEA+ Math Score			SEA+ Math Standardized Score		
	(1)	(2)	(3)	(4)	(5)	(6)
HTHT Treatment*Class	12.63***	11.74**	15.45**	0.16***	0.14**	0.19**
Matific Use	(4.73)	(5.33)	(17.21)	(0.06)	(0.07)	(0.21)
HTHT Treatment	-3.01	-2.83	-3.81	-0.04	-0.03	-0.05
	(3.47)	(3.37)	(4.20)	(0.04)	(0.04)	(0.05)
Matific Use	-8.19***	-7.56**	-9.36	-0.10***	-0.09**	-0.12
	(2.51)	(3.02)	(16.72)	(0.03)	(0.04)	(0.21)
Baseline math score	0.97***	0.95***	0.95***	0.84***	0.83***	0.82***
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Living in capital		-5.58*	-7.14		-0.07*	-0.09
		(3.15)	(4.76)		(0.04)	(0.06)
School quintile		1.85	1.43		0.02	0.02
		(1.21)	(1.17)		(0.01)	(0.01)
Eligible for Afampe		-4.37*	-4.91**		-0.05*	-0.06**
		(2.24)	(2.31)		(0.03)	(0.03)
Eligible for TUS		-2.82	-3.37		-0.03	-0.04
		(2.64)	(2.57)		(0.03)	(0.03)
Controls	NO	YES	YES	NO	YES	YES
Constant	61.49***	71.26***	82.30***	0.02	0.00	0.04
	(18.53)	(19.87)	(14.46)	(0.03)	(0.07)	(0.07)
Observations	2,002	1,954	1,954	2,002	1,954	1,954
Number of groups	-	-	99	-	-	99
R-squared	0.71	0.72	0.72	0.71	0.72	0.72
Mixed Effects	NO	NO	YES	NO	NO	YES

Note: Class-level Matific Use category: 1 (less than 10 weeks in which at least 50% class finished a Matific episode); 2 (10 weeks or above weeks in which at least 50% class finished a Matific episode). *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

4.3.1. Sub-group analysis by student initial achievement, gender, and father's education

We investigated whether the effect of HTHT intervention differed across three pre-specified student characteristics recorded in our data: student initial achievement (Tables 24 and 25), gender, and father's education (Table 26). There were no significantly different effects by students' quartile groups based on their initial test scores in math assessment, as well as gender and father's education. Although not presented here, we observed similar patterns (i.e., no differential effects among sub-groups) by school quintile, school types, Afempe/TUS cash transfer status, and residency in the capital.

Table 24: The Effect of HTHT on Students' Math Score by Students' Baseline Score Quartile - Panel A

VARIABLES	Students' Baseline: Q1		Students' Baseline: Q2		Students' Baseline: Q3		Students' Baseline: Q4	
	(1) Math Score	(2) Math Score	(3) Math Score	(4) Math Score	(5) Math Score	(6) Math Score	(7) Math Score	(8) Math Score
HTHT Treatment	0.96	3.39	-8.22	-7.90	3.76	3.36	-3.35	-4.58
	(5.37)	(5.32)	(5.12)	(5.10)	(4.19)	(3.77)	(4.88)	(4.83)
Baseline math score	0.66***	0.68***	1.23***	1.19***	1.10***	1.07***	1.04***	1.02***
	(0.09)	(0.08)	(0.17)	(0.17)	(0.11)	(0.12)	(0.07)	(0.07)
Living in capital		-9.26		-4.13		-7.71*		1.45
		(5.89)		(5.19)		(4.37)		(4.62)
School quintile		2.57		1.59		1.96		0.77
		(1.60)		(1.51)		(1.78)		(2.23)
Eligible for Afampe		-5.55		3.40		-5.08		-9.82
		(4.46)		(4.57)		(4.29)		(6.18)
Eligible for TUS		-0.15		-8.15		-3.99		-1.47
		(3.88)		(5.32)		(5.96)		(7.26)
Constant	318.90***	303.76***	-178.04	-146.65	-68.47	-39.65	-8.96	9.80
	(75.53)	(73.33)	(153.29)	(155.10)	(107.36)	(111.69)	(68.48)	(69.48)
Observations	501	493	500	493	501	486	500	482
R-squared	0.13	0.16	0.13	0.14	0.15	0.17	0.42	0.44

Note: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Table 25: The Effect of HTHT on Students' Math Score by Students' Baseline Score Quartile - Panel B

VARIABLES	Students' Baseline: Q1		Students' Baseline: Q2		Students' Baseline: Q3		Students' Baseline: Q4	
	(1) Math Score	(2) Math Score	(3) Math Score	(4) Math Score	(5) Math Score	(6) Math Score	(7) Math Score	(8) Math Score
HTHT Treatment	-3.48	-0.32	-11.55**	-18.63***	-0.79	2.75	1.53	-1.72
	(5.98)	(6.71)	(5.19)	(6.56)	(4.57)	(4.90)	(5.38)	(5.18)
Baseline math score	0.76***	0.68***	1.10***	1.05***	0.98***	0.93***	1.12***	1.13***
	(0.09)	(0.11)	(0.17)	(0.21)	(0.14)	(0.15)	(0.08)	(0.08)
Living in capital		-6.82		-4.27		-13.09**		6.93
		(4.96)		(8.01)		(5.13)		(5.64)
School quintile		1.20		-2.72		1.43		-2.73
		(2.08)		(2.74)		(1.80)		(3.05)
Eligible for Afampe		4.64		-2.71		-2.41		-10.25
		(5.88)		(6.42)		(6.80)		(8.32)
Eligible for TUS		-3.18		-22.09***		0.78		8.52
		(5.17)		(6.75)		(7.77)		(10.07)
Female		2.22		-3.30		-10.86*		-0.11
		(5.51)		(4.92)		(6.05)		(5.75)
Age		-0.30		-0.58**		-0.49*		0.08
		(0.22)		(0.28)		(0.28)		(0.41)
Mother edu secondary		4.36		6.51		6.05		-9.82
		(8.72)		(6.08)		(6.06)		(7.30)
Farther edu secondary		7.23		2.17		8.70		15.75**
		(8.50)		(5.54)		(6.28)		(7.33)
Living with father		-4.04		0.93		4.52		6.38
		(5.70)		(5.31)		(5.53)		(5.88)
Living with mother		2.37		-3.91		5.42		-0.34
		(8.88)		(8.37)		(7.96)		(10.29)
Household size		8.86		10.75		-9.37		-0.95
		(7.23)		(10.69)		(10.24)		(11.82)
Wealth index tercile		0.59		-1.70		-3.29		10.62*
		(2.92)		(3.65)		(3.97)		(5.44)
Constant	241.22***	329.11***	-61.61	76.76	46.28	153.67	-97.65	-132.10
	(77.62)	(113.03)	(153.38)	(184.75)	(131.29)	(140.85)	(79.72)	(99.46)
Observations	388	215	387	236	388	266	387	271
R-squared	0.20	0.23	0.13	0.22	0.13	0.19	0.42	0.52

Note: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Table 26: The Effect of HTHT on Students' Math Score by Students' Gender and Father's Edu - Panel B

VARIABLES	Boys		Girls		Father didn't complete secondary education		Father completed secondary education	
	(1) Math Score	(2) Math Score	(3) Math Score	(4) Math Score	(5) Math Score	(6) Math Score	(7) Math Score	(8) Math Score
HTHT Treatment	-3.75	-7.61	-2.19	-2.47	-4.89	-4.17	-8.02	-6.91
	(4.53)	(5.96)	(3.51)	(3.86)	(3.63)	(3.78)	(6.85)	(6.95)
Baseline math score	0.95***	0.99***	1.01***	0.96***	0.98***	0.95***	1.05***	1.04***
	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)
Living in capital		-4.38		-7.21*		-6.24*		0.72
		(5.12)		(4.18)		(3.74)		(8.69)
School quintile		-1.43		-0.47		0.30		-6.06
		(2.16)		(1.86)		(1.48)		(4.16)
Eligible for Afampe		-10.97**		3.93		-3.57		-4.33
		(4.54)		(4.88)		(3.36)		(9.12)
Eligible for TUS		1.89		-8.32*		-4.68		2.42
		(5.85)		(4.52)		(3.95)		(11.25)
Female		-		-		-5.37		1.72
		-		-		(3.40)		(5.82)
Age		-0.30		-0.26		-0.30*		-0.47
		(0.19)		(0.18)		(0.16)		(0.56)
Mother edu secondary		-6.26		9.49**		-0.19		6.60
		(4.59)		(4.55)		(3.49)		(7.70)
Farther edu secondary		9.21*		11.11*		-		-
		(5.10)		(5.78)		-		-
Living with father		6.60		-2.13		2.07		1.35
		(4.03)		(3.68)		(3.04)		(7.06)
Living with mother		-7.22		6.68		0.77		-0.34
		(6.07)		(5.64)		(5.05)		(11.44)
Household size		7.82		-2.08		4.83		-17.26
		(7.31)		(5.60)		(4.77)		(18.90)
Wealth index tercile		-0.72		3.75		0.51		4.50
		(3.04)		(2.97)		(2.18)		(5.87)
Constant	74.28***	85.56*	18.07	79.48**	49.65**	109.51***	-8.89	61.75
	(27.41)	(45.85)	(26.69)	(34.93)	(24.84)	(36.64)	(40.87)	(77.58)
Observations	769	503	734	485	924	798	205	190
R-squared	0.71	0.75	0.73	0.77	0.73	0.73	0.77	0.79

Note: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

4.5. Translational Impact of Teacher Factors on Student Outcomes: SEA+, Grit, and Student-Reported Teaching Quality.

We examined factors that may have translational impact of teacher factors on student outcomes: (1) SEA+ Mathematics Score, (2) Grit Score, and (3) Student-Report Teaching Quality. Overall, we found that teacher factors – self-efficacy, perceptions of data, and curricular access – do not mediate SEA+ scores or student grit scores. However, we found that teachers' self-efficacy in instruction has a modest effect on student-reported teaching quality ($\beta = .07$, $p < .05$). In addition, curriculum instruction also had a modest effect on student-reported teaching quality ($\beta = .09$, $p < .05$). Details of the translational outcome analyses are presented in Table 27.

Table 27. Mediated Effect of Teacher Measures on Student Outcomes: SEA+ scores

Predictors Contributing to Translational Outcomes	Outcome Measures (SE)		
	SEA+ Math Score	GRIT	Student-Reported Teaching Quality
Teacher Self-Efficacy (standardized)	1.96 (3.83)	–0.02 (0.04)	0.05 (0.04)
Teacher SE: Engagement	1.33 (3.66)	–0.01 (0.04)	0.03 (0.04)
Teacher SE: Instruction	–1.81 (3.78)	–0.02 (0.04)	0.07* (0.04)
Teacher SE: Management	3.77 (2.95)	–0.03 (0.03)	0.02 (0.03)
HTHT Perception	3.24 (3.02)	0.02 (0.04)	–0.01 (0.03)
School Data System	–0.59 (2.44)	–0.00 (0.03)	0.00 (0.03)
Personalised Learning	1.61 (2.66)	–0.02 (0.03)	0.05* (0.03)
Technology Curriculum Access	–0.88 (1.76)	–0.01 (0.02)	–0.03 (0.02)
Non-Tech Curriculum Access	1.48 (1.81)	–0.01 (0.02)	–0.01 (0.02)
Curriculum Instruction	–0.11 (3.85)	–0.03 (0.04)	0.09** (0.04)
Professional Environment	3.10 (3.16)	–0.04 (0.04)	–0.00 (0.03)
Access to Technology	0.96 (4.64)	0.06 (0.05)	0.04 (0.05)

Note: Values in parenthesis are standard errors from mixed-effects regression models. * $p < .05$, ** $p < .01$

This is the first large-scale experimental study demonstrating the effectiveness of the HTHT paradigm. Interventions that leverage both HTHT components in large-scale implementation through a controlled experimental design are uncommon. Educational policy may be designed to maximize interventions that combine HTHT components synergistically, while allowing flexibility in large-scale implementation of teacher interventions.

IV. REFERENCES

Alan, S., Boneva, T., & Ertac, S. (2019). Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit, *The Quarterly Journal of Economics, 134* (3), 1121–1162

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, *333*, 1034-1036

Allen, J. P., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. C. (2013). Observations of effective teaching in secondary school classrooms: Predicting student achievement with the CLASS-S. *Journal of School Psychology*, *42*(*1*), 76-98.

Andelfinger, J. A., Bray, L. A., Kruis, N. E., Smith, M. O., & Hawkins, P. M. (2017). *Perceptions of the Summit Learning Platform*. Indiana, PA: Applied Research Lab.

Ball, D. L., & Cohen, D. K. (1999). Developing practices, developing practitioners: Toward a practice-based theory of professional development. In G. Sykes & L. Darling-Hammonds (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 30–32). San Francisco, CA: Jossey-Bass.

Bill and Melinda Gates Foundation, Measures of Effective Teaching (MET). (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Seattle, WA: Bill and Melinda Gates Foundation.

Blank, R. K. (2009, June). *Effects of teacher professional development on gains in student achievement: How meta analysis provides scientific evidence useful to education leaders*. Washington, DC: Council of Chief State School Officers.

Goldstein, H. (1995). Hierarchical Data Modeling in the Social Sciences. *Journal of Educational and Behavioral Statistics, 20(*2), 201–204.

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting education? Experimental evidence on technology-aided instruction in India. *American Economic Review, 109*(4), 1426-60.

Murnane, R., & Willett, J. (2010). *Methods Matter: Improving Causal Inference in Educational and Social Science Research*. Oxford University Press.

Pane, J. F., Steiner, E. D., Baird, M. D., & Hamilton, L. S. (2015). *Continued progress: Promising evidence on personalized learning*. Seattle, WA: Bill and Melinda Gates Foundation.

Park, Y. S., Xing, K., & Lee, Y.-S. (2018). Explanatory cognitive diagnostic models: Incorporating latent and observed predictors. *Applied Psychological Measurement*, *42*(*5*), 376-392.

Park, Y. S., & Lee, Y.-S. (2014). An extension of the DINA model using covariates: Examining factors affecting response probability and latent classification. *Applied Psychological Measurement*, *38*(*5*), 376-390.

Pianta, R. C. (2011, November). *Teaching children well: New evidence-based approaches to teacher professional development and training*. Washington, DC: Center for American Progress.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Wilson, S. M. & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of Research in Education, 24*, 173–209.

Yoon, K. S., Duncan, T., Lee, S., W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement.* [Issues & Answers Report, REL 2007-No. 033]. Washington, DC: U.S. Department of Education, Institute of Educational Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest.

APPENDIX. Supplemental Information

Table A1. Target Recruitment of Schools by Sampling Design

School Classification	Number	Study Condition Groups			Overall
		Control	Intervention		
			"Hard"	"Soft"	
Regular schools ("Urban" schools with 4-hour school days)	Schools	8	8	14	30
	Teachers	16	15	27	58
	Students	363	351	696	1,410
Appender schools (vulnerable schools in Sociocultural Quintiles 1 to 2)	Schools	6	6	11	23
	Teachers	11	9	17	37
	Students	224	209	405	838
Full-time schools in Sociocultrual Quintiles 1 to 3 (schools with longer 6- to 8-hour school days)	Schools	6	6	13	25
	Teachers	8	9	17	34
	Students	204	171	407	782
Full-time schools in Sociocultrual Quintiles 4 to 5 (schools with longer 6- to 8-hour school days)	Schools	5	5	9	19
	Teachers	7	8	15	30
	Students	168	191	388	747
Total	Schools	25	25	47	97
	Teachers*	42	41	76	159
	Students	959	922	1,896	3,777

Note: *For number of teachers, 159 reflects teaching positions, with 158 unique teachers (one teacher serves two schools).

Table A2. Overview of HTHT Data Collection

A2.1. Panel A: Student Assessment and School Data (without Student Survey)

Instrument	Description	Number of Observation
SEA+ Math Assessment	Baseline (T_0)	2,709
SEA+ Math Assessment	Endline (T_1)	2,018
SEA+ Math Assessment	Both waves	2,018
Student Information CSC	Total	2,776
Student Database	**Total**	**2,002**
School Information CSC	Total	176
Student+School Database	Total	2,002
Teacher Survey	Baseline (T_0)	165
Teacher Survey	Endline (T_1)	118
Teacher Survey	Both Waves	114
Teacher Database	Total	114
Student+School+Teacher	Total	1,628

A2.2. Panel B: Student Assessment and School Data (with Student Survey)

Instrument	Description	Number of Observation
SEA+ Math Assessment	Baseline (T_0)	2,709
SEA+ Math Assessment	Endline (T_1)	2,018
SEA+ Math Assessment	Both waves	2,018
Student Information CSC	Total	2,776
Student Survey	Baseline (T_0)	2,238
Student Survey	Endline (T_1)	2,415
Student Survey	Both waves	1,587
Student (Math+Survey)	Total	1,550
School Information CSC	Total	176
Student Database	Total	1,550
Teacher Survey	Baseline (T_0)	165
Teacher Survey	Endline (T_1)	118
Teacher Survey	Both Waves	114
Teacher Database	Total	114
Student + Teacher	All information	1,268

Note: Panel A (without student surveys) represents 72% (n = 2,002) of students with CSC information (n = 2,776). Panel B (with student surveys), represents 56% (n = 1,550) of students with CSC information (n = 2,776) and 77% of data from Panel A.

Table A3. The Effect of HTHT on Students' Math Score (SEA+ Assessment): Panel B

Variables	SEA+ Math Score				
	(1)	(2)	(3)	(4)	(5)
HTHT Treatment	-3.28	-2.84	-2.79	-4.81	-4.15
	(3.00)	(3.06)	(3.12)	(3.38)	(4.36)
Baseline math score	0.99***	0.98***	0.97***	0.97***	0.96***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Living in capital		-6.22**	-4.74	-5.09	-6.42
		(2.86)	(2.90)	(3.35)	(5.11)
School quintile		0.77	0.45	-0.94	-0.33
		(1.40)	(1.48)	(1.68)	(1.41)
Eligible for Afampe		-5.81**	-5.74**	-3.39	-4.38
		(2.57)	(2.74)	(3.16)	(3.45)
Eligible for TUS		-1.53	-0.74	-3.51	-4.16
		(2.91)	(3.05)	(3.75)	(3.94)
Female			-1.95	-3.64	-3.88
			(2.35)	(2.93)	(2.58)
Age			-0.17	-0.28*	-0.25*
			(0.11)	(0.15)	(0.14)
Mother complete secondary				1.46	2.26
				(3.07)	(3.26)
Farther complete secondary				10.44**	9.01**
				(4.02)	(3.61)
Living with father				2.46	1.57
				(2.86)	(2.68)
Living with mother				0.08	-0.03
				(4.40)	(4.31)
Household size				2.62	3.00
				(4.69)	(4.41)
Wealth index tercile				1.51	1.57
				(2.01)	(2.02)
Constant	43.75**	51.77**	79.64***	87.88***	92.00***
	(19.19)	(20.64)	(25.29)	(30.62)	(26.80)
Observations (schools)	1,550	1,510	1,464	988	988 (95)
R-squared	0.73	0.74	0.73	0.76	0.76
Mixed effects	NO	NO	NO	NO	YES

Table A4. The Interaction Effect of HTHT X Matific Usage (Continuous) on Students' Math Score: Panel B

Variables	SEA+ Math Score				
	(1)	(2)	(3)	(4)	(5)
HTHT Treatment*Matific Use	1.89**	1.93**	1.87**	1.35**	1.54**
	(0.85)	(0.86)	(0.87)	(0.96)	(1.05)
HTHT Treatment	-8.61**	-8.11**	-7.87**	-8.76*	-9.01*
	(3.68)	(3.78)	(3.93)	(4.47)	(5.03)
Matific Use	-0.66	-0.76*	-0.76*	-0.44	-0.31
	(0.46)	(0.41)	(0.41)	(0.52)	(0.91)
Baseline math score	0.98***	0.97***	0.96***	0.97***	0.96***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Living in capital		-6.49**	-5.00*	-5.55	-6.58
		(2.94)	(3.01)	(3.49)	(5.12)
School quintile		0.92	0.63	-0.78	-0.13
		(1.34)	(1.44)	(1.65)	(1.42)
Eligible for Afampe		-5.38**	-5.31*	-2.97	-4.22
		(2.52)	(2.71)	(3.16)	(3.44)
Eligible for TUS		-1.72	-0.98	-3.57	-4.08
		(2.93)	(3.06)	(3.79)	(3.93)
Female			-1.82	-3.60	-3.85
			(2.36)	(2.95)	(2.57)
Age			-0.16	-0.26*	-0.24*
			(0.11)	(0.15)	(0.14)
Mother complete secondary				1.63	2.17
				(3.10)	(3.25)
Farther complete secondary				10.59***	9.20**
				(3.99)	(3.60)
Living with father				2.20	1.27
				(2.88)	(2.68)
Living with mother				0.33	0.07
				(4.45)	(4.30)
Household size				2.43	2.84
				(4.67)	(4.40)

Wealth index tercile				1.36	1.39
				(2.03)	(2.01)
Constant	51.72***	58.88***	84.23***	91.46***	97.02***
	(18.97)	(20.26)	(25.13)	(30.43)	(26.84)
Observations (# of schools)	1,550	1,510	1,464	988	988 (95)
R-squared	0.73	0.74	0.73	0.76	0.76
Mixed Effects	NO	NO	NO	NO	YES

Note: Robust standard errors clustered at the school level in parentheses. Matific episodes completion by intervals of 10 episodes.
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Figure A1. The Interaction Effect of HTHT X Matific Usage (Continuous) on Students' Math Score: Panel B

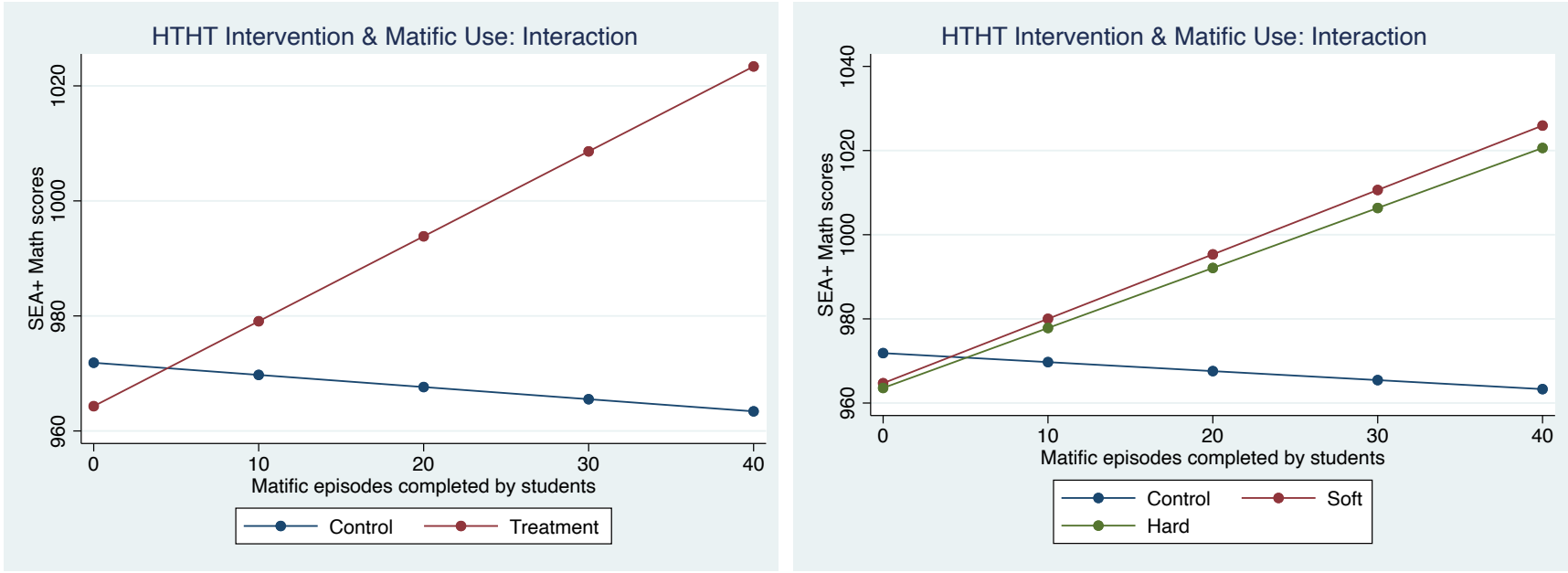


Table A5. The Interaction Effect of HTHT X Matific Usage at School (Categorical) on Students' Math Score : Panel B

Variables	SEA+ Math Score				
	(1)	(2)	(3)	(4)	(5)
HTHT Treatment*Matific Use	8.65***	8.50***	8.34***	5.82**	6.32**
	(2.43)	(2.52)	(2.57)	(2.76)	(3.16)
HTHT Treatment	-23.36***	-22.09***	-21.73***	-18.87***	-20.33***
	(5.54)	(5.87)	(6.18)	(7.09)	(7.62)
Matific Use	-2.76	-3.19	-3.10	-1.42	-0.61
	(1.94)	(1.99)	(1.98)	(2.21)	(2.76)
Baseline math score	0.98***	0.97***	0.97***	0.97***	0.96***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Living in capital		-5.82*	-4.39	-5.18	-5.73
		(3.11)	(3.21)	(3.62)	(5.15)
School quintile		0.61	0.36	-0.99	-0.40
		(1.31)	(1.41)	(1.64)	(1.42)
Eligible for Afampe		-5.18**	-5.14*	-2.72	-3.95
		(2.48)	(2.67)	(3.08)	(3.43)
Eligible for TUS		-1.68	-0.85	-3.39	-4.04
		(2.94)	(3.07)	(3.81)	(3.91)
Female			-2.05	-3.87	-4.12
			(2.40)	(2.97)	(2.56)
Age			-0.16	-0.26*	-0.24*
			(0.11)	(0.15)	(0.14)
Mother complete secondary				1.79	2.21
				(3.11)	(3.23)
Farther complete secondary				10.67***	9.28***
				(3.97)	(3.58)
Living with father				1.84	0.98
				(2.86)	(2.67)
Living with mother				0.11	-0.12
				(4.52)	(4.29)
Household size				2.43	2.92
				(4.77)	(4.38)

Wealth index tercile				1.31	1.40
				(2.04)	(2.00)
Constant	57.39***	63.71***	88.80***	94.14***	99.38***
	(19.23)	(20.32)	(25.38)	(30.92)	(27.00)
Observations (# of schools)	1,550	1,510	1,464	988	988 (95)
R-squared	0.73	0.74	0.73	0.76	0.76
Mixed Effects	NO	NO	NO	NO	YES

Note: Robust standard errors clustered at the school level in parentheses. Matific Use category: 1 (less than 10 episodes); 2 (less than 20 episodes); 3(less than 40 episodes); 4 (40 episodes or above). *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Model (5) Effect size: 0.08

Table A6. The Interaction Effect of HTHT X Matific Usage at School (Categorical) on Students' Math Score : Panel B

Variables	SEA+ Math Score				
	(1)	(2)	(3)	(4)	(5)
HTHT Treatment*Matific Use Group 1 (*ref*)	-	-	-	-	-
	-	-	-	-	-
HTHT Treatment*Matific Use Group 2	13.78*	12.32	12.51	13.59	12.04
	(7.74)	(7.85)	(8.43)	(8.36)	(9.52)
HTHT Treatment*Matific Use Group 3	17.23**	14.31*	12.92	6.99	5.20
	(7.66)	(7.85)	(7.93)	(9.03)	(9.83)
HTHT Treatment*Matific Use Group 4	28.27***	27.97***	28.15***	21.60***	24.30**
	(7.12)	(7.06)	(7.29)	(8.13)	(10.47)
HTHT Treatment	-16.56***	-14.70***	-14.60***	-14.96***	-14.91***
	(3.94)	(4.23)	(4.52)	(4.99)	(5.68)
Matific Use	-4.17	-5.15	-4.57	-4.19	-3.67
	(6.41)	(6.55)	(6.95)	(6.44)	(8.28)
Baseline math score	0.98***	0.97***	0.97***	0.97***	0.96***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Living in capital		-5.50*	-4.09	-4.80	-5.70
		(3.17)	(3.27)	(3.68)	(5.15)
School quintile		0.52	0.30	-1.04	-0.40
		(1.28)	(1.38)	(1.60)	(1.42)
Eligible for Afampe		-5.52**	-5.49**	-3.15	-4.26
		(2.48)	(2.66)	(3.13)	(3.44)
Eligible for TUS		-1.59	-0.76	-3.43	-4.19
		(2.94)	(3.06)	(3.82)	(3.91)
Female			-2.04	-3.98	-4.18
			(2.40)	(2.96)	(2.56)
Age			-0.16	-0.26*	-0.24*
			(0.11)	(0.15)	(0.14)
Mother complete secondary				2.13	2.47
				(3.15)	(3.24)
Farther complete secondary				10.53***	9.24**
				(3.84)	(3.59)

Living with father				1.90	1.04
				(2.80)	(2.67)
Living with mother				-0.00	-0.19
				(4.47)	(4.29)
Household size				2.72	3.13
				(4.83)	(4.38)
Wealth index tercile				1.23	1.26
				(2.03)	(2.01)
Constant	52.53***	59.53***	85.12***	92.99***	99.02***
	(18.86)	(20.04)	(25.22)	(30.31)	(26.69)
Observations (# of schools)	1,550	1,510	1,464	988	988 (55)
R-squared	0.73	0.74	0.74	0.76	0.76
Mixed Effects	NO	NO	NO	NO	YES

Note: Matific Use category: 1 (less than 10 episodes); 2 (less than 20 episodes); 3(less than 40 episodes); 4 (40 episodes or above).
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Model (5) Effect size: 0.29

Table A7. The Interaction Effect of HTHT X Class-level Matific Use (Categorical) on Students' Math Score : Panel B

Variables	SEA+ Math score			
	(1)	(2)	(3)	(4)
HTHT Treatment*Class Matific Use	14.67***	14.64**	14.98**	15.15**
	(4.80)	(5.58)	(5.72)	(16.84)
HTHT Treatment	-4.86	-4.28	-4.17	-3.40
	(3.24)	(3.24)	(3.29)	(4.26)
Matific Use	-10.87***	-11.45***	-12.07***	-11.73
	(2.24)	(3.01)	(3.02)	(16.32)
Baseline math score	0.98***	0.98***	0.97***	0.96***
	(0.02)	(0.02)	(0.02)	(0.02)
Living in capital		-6.36**	-4.90*	-5.39
		(2.86)	(2.92)	(4.76)
School quintile		0.73	0.40	0.42
		(1.38)	(1.46)	(1.25)
Eligible for Afampe		-5.58**	-5.52**	-5.86**
		(2.58)	(2.75)	(2.67)
Eligible for TUS		-1.67	-0.91	-1.94
		(2.93)	(3.07)	(3.06)
Female			-1.85	-2.45
			(2.35)	(2.18)
Age			-0.17	-0.15
			(0.11)	(0.11)
Constant	44.99**	52.67**	79.63***	86.81***
	(18.81)	(20.44)	(25.30)	(21.69)
Observations (# of schools)	1,550	1,510	1,464	1,464 (97)
R-squared	0.73	0.74	0.73	0.73
Mixed Effects	NO	NO	NO	YES

Note: Class-level Matific Use category: 1 (less than 10 weeks in which at least 50% class finished a Matific episode); 2 (10 weeks or above weeks in which at least 50% class finished a Matific episode). *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Model (4) Effect size: 0.18