Implementing and Evaluating ASSISTments Online Math Homework Support At large Scale over Two Years: Findings and Lessons Learned

Mingyu Feng[1[0000-0001-9635-1611]], Neil Heffernan[2[0000-0002-3280-288X]], Kelly Collins[1], Cristina Heffernan[3], and Robert F. Murphy[4]

[1] WestEd, San Francisco, California 94107, USA
[2] Worcester Polytechnic Institute, Worcester, Massachusetts 01609, USA
[3] The ASSISTments Foundation, Auburn, Massachusetts 01501, USA
[4] LFC Research, California 94043, USA

mfeng@wested.org, nth@wpi.edu, kelly.collins@wested.org, cristina.heffernan@assistmetns.org, bfmurph@icloud.com

Presented at the The 24th International Conference on Artificial Intelligence in Education, July 3-7, 2023 in Tokyo, Japan.

Abstract. Math performance continues to be an important focus for improvement. The most recent National Report Card in the U.S. suggested student math scores declined in the past two years possibly due to COVID-19 pandemic and related school closures. We report on the implementation of a math homework program that leverages AI-based one-to-one technology, in 32 schools for two years as a part of a randomized controlled trial in diverse settings of the state of North Carolina in the US. The program, called "ASSISTments," provides feedback to students as they solve homework problems and automatically prepares reports for teachers about student performance on daily assignments. The paper describes the sample, the study design, the implementation of the intervention, including the recruitment effort, the training and support provided to teachers, and the approaches taken to assess teacher's progress and improve implementation fidelity. Analysis of data collected during the study suggest that (a) treatment teachers changed their homework review practices as they used ASSISTments, and (b) the usage of ASSISTments was positively correlated with student learning outcome.

Keywords: ASSISTments, math learning, effective teaching, AI-based program, school implementation

1 Introduction

Math performance continues to be an important focus for improvement in the United States. Due to the promise of technology as a tool for improving mathematics education and closing the achievement gap, the use of educational technology in K-12 education has expanded dramatically in recent years, accelerated by the COVID-19 pandemic.

The AIED and intelligent tutoring systems researchers and developers have built numerous technology-based learning platforms and programs, and many have been shown to be effective in a lab setting, or with a small number of closely monitored classrooms. Yet few of these products have been implemented at large scale in authentic school settings over an extended period. The challenges of wide adoption and effective implementation in schools come from several aspects, such as understanding school settings and meeting school priorities, availability of technology infrastructure to guarantee sufficient student/teacher access to equipment, integration of the program into established classroom practices, as well as training and continuous support for users to ensure sustained use with fidelity. In this paper we report on a large-scale efficacy randomized controlled trial (RCT) in diverse settings of the state of North Carolina in the U.S. In particular, we focus on the implementation of the ASSISTments platform and discuss how we have addressed each challenge to ensure faithful implementation.

The ASSISTments platform [1] is a technology-based, formative assessment platform for improving teacher practices and student math learning outcomes. As students work through problems and enter their answers into ASSISTments, the system provides immediate feedback on the correctness of answers and offers additional assistance in the form of hints or scaffolds. Students' performance on ASSISTments problems serves as an assessment of proficiency, enabling teachers to adjust classroom instruction and pacing to match the knowledge base of the class. ASSISTments was identified as effective at improving student learning and changing teacher's homework review practices during an efficacy study in Maine ([2,3], effect size g=.22, p<.01). We recently conducted a large scale RCT to replicate the Maine study and see whether the found effects replicate in a heterogeneous population that more closely matches national demographics. The replication study to test the impact of the ASSISTments platform on students' math learning was guided by the following research questions:

- Student learning: *Do students in schools that use ASSISTments for homework learn more than students in schools that do homework without ASSISTments?*
- Impacts on classroom instruction: *Does using ASSISTments lead to adjustment in teachers' homework practices, and, if so, how they have done this?*
- Relationship between usage and student learning outcome: *What are the effects of implementation fidelity and dosage on learning?*

2 Background

2.1 The ASSISTments Program

Over the past two years, ASSISTments use in schools increased significantly, going from supporting 800 teachers to supporting 20,000 teachers and their 500,000 students. A rapid review that synthesizes existing evidence in online programs that promoted learning recommended ASSISTments as one of the few digital learning programs for use in response to the COVID pandemic [4]. A review of 29 studies that met rigorous standards of randomization [5] indicated that ASSISTments was one of only "Two interventions in the United States [that] stand out as being particularly promising."

The platform uses technology to give teachers new capabilities for assigning and reviewing homework and to give students additional support for learning as they do homework. Content in ASSISTments consists of mathematics problems with answers and hint messages. These mathematics problems are bundled into problem sets which teachers can use ASSISTments to assign to students in class or as homework. Students first do their assigned problems on paper and then enter their answers into ASSISTments to receive immediate feedback about the correctness of their answers, and/or hints on how to improve their answers or help separate multi-step problems into parts. One type of problem set is mastery-oriented "Skill Builders". Each skill builder provides opportunities for students to practice solving problems that focus on a targeted skill, until they reach a teacher-defined "mastery" threshold of proficiency (e.g., a streak of three correct answers on similar math problems). ASSISTments also implements the research-based instructional strategy of spaced practice [6] by automatically re-assessing students on skills and concepts that were "mastered" earlier at regular intervals, and providing further opportunities for students to hone those skills.

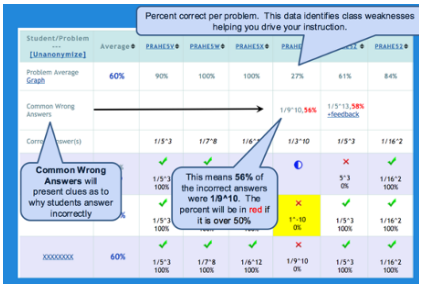

Fig. 1. An item report for teachers.

ASSISTments provides teachers with real-time, easily accessible reports that summarize student work for a particular assignment in a grid format (Fig. 1), which teachers can use to target their homework review in class and tailor instruction to their students' needs. These reports inform teachers about the average percent correct on each question, skills covered in the assignment, common wrong answers, and each student's answer to every question. The color-coded cells provide information on each student's performance on each problem, in a format that enables teachers to focus on the problems that were difficult for many students. Teachers use information about individual students to form instructional groups and address common difficulties.

2.2 Theoretical Framework

The design and development of ASSISTments aligns to theory- and empirically-based instructional practices of formative assessment [7] and skill development [8], and is built upon the following theoretical foundations and empirical research.

Immediate Feedback to Learners. Feedback has been identified as a powerful way to increase student learning [9]. An extensive set of studies (e.g., [10, 11, 12]) has found significant learning gains in response to feedback within computer-based instruction. A common form of feedback is correctness feedback, in which feedback informs the learner if their response is correct or incorrect, with no other information. Studies have shown positive results of correctness feedback [13, 14].

Formative Assessment. Formative assessment provides feedback to both teachers and students by informing teachers of student learning and progress and students of their own performance in relation to learning goals. Several decades of research have

shown formative assessment, also known as assessment for learning, to be an effective way of improving both teaching and student learning [15, 16, 17, 18, 19].

Education technology can facilitate formative assessment and have a significant effect on improving math learning [20, 21]. Teachers benefit from the information and analytics on student learning and performance that these tools can provide, which help them modify instruction and better adapt to student learning needs [21, 22]).

Mastery learning and spaced practice. In mastery learning, students only advance to a new learning objective when they demonstrate proficiency with the current one. Mastery learning programs lead to higher student achievement than more traditional forms of teaching [23,24]. An Institute of Education Sciences (IES) Practice Guide [6] recommends spacing practice over time. Research has demonstrated that retention increases when learners are repeatedly exposed to content (see [25, 26, 27]).

Homework to boost learning. Homework is often required in middle schools and expected by students and parents, and a meaningful amount of instructional time is allocated to homework [28, 29]. Yet it is also somewhat controversial and perceived as needing improvement (e.g. [30, 31]). Homework provides a key opportunity for individual practice, which is undeniably important to mathematics learning. To "make homework work," teachers are recommended to streamline assignments and be intentional with each assignment, taking into account factors such as whether students understand the purpose of the assignment and the feedback the teacher will provide [32].

2.3 Research Design

Sample and Settings. The study[1] took place in North Carolina (NC), a state more demographically representative of the U.S. than Maine. Sixty-three schools from 41 different districts enrolled in the study. Demographic characteristics of the schools are shown in Fig. 2, comparing to the Maine study and the population in the U.S.. These schools served several different grade levels (6-8, 8-12, K-8) and were distributed across rural, town, suburban, and city communities (33 rural, 11 town, 8 suburban, and 11 city). Of the 63 schools, 18 were charter schools, 45 were public schools, and 48 received Title 1 funding. 102 7th grade math teachers and their classrooms enrolled in the study.

Baseline School Characteristics	Treatment Schools	Control Schools	NC Population	US Population	Sample in Original Study in Maine
Avg. % of White students	58	58	51	53	93
Avg. % Hispanic students	13	16	14	23	2
Avg. % Black students	20	21	27	16	2
Avg. % socioeconomically disadvantaged	53	48	59	54	39
% of Rural Schools	47	58	49	38	Over 50%

Fig. 1. Demographics of Participating Schools

[1] The study has been pre-registered on Registry of Efficacy and Effectiveness Studies (REES) https://sreereg.icpsr.umich.edu/framework/pdf/index.php?id=2064. The study has been approved by the Institutional Review Board at Worcester Polytechnic Institute and WestEd. Participating teachers all signed consent forms. Parents received a notification letter and opt-out form for their children.

Study Design. Schools within each district were paired based on their demographic characteristics and student prior performance on state math and English language arts (ELA) tests and then randomly assigned to a treatment or business-as-usual control condition. Overall, 32 schools and 60 teachers were assigned to the treatment condition and 31 schools and 52 teachers to the control condition. All seventh-grade math teachers in a school had the same assignment of condition.

The study was carried out during the 2018-19 and 2019-20 school years. During the study, ASSISTments was conceptualized as consisting of both teacher professional development (PD) and the use of the platform by teachers and students as illustrated by the theory of change in Fig 3.

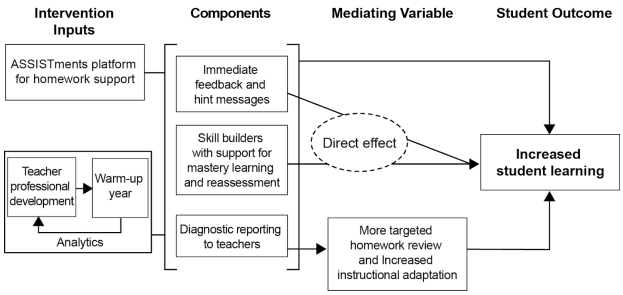

Fig 3. ASSISTments Theory of Change

The program was implemented by all Grade 7 teachers in the treatment schools over two consecutive years—the first year was a ramp-up year, where teachers in the treatment condition learned and practiced using the tool with their students. The second year of implementation was the measurement year when treatment teachers continued to use ASSISTments with a new cohort of 7th grade students whose math learning outcome was measured in spring 2020. Treatment teachers received 2 days of PD during summer 2018 and 2019, and additional coaching and technical assistance distributed across the school year via webinars, video conferencing, and 2-3 times in-person visits each school year. Teachers had the full discretion to either select pre-built content from ASSISTments that aligns with the learning standards they teach in class, or request to have their own homework problems built in the system for them to assign.

Teachers in the control condition did not have access to ASSISTments or to the PD and coaching till summer 2020 ("delayed treatment"). They continued with their typical homework practices ("business-as-usual") including the possible use of other online homework platforms or supplemental programs in class. We used an end-of-year survey and daily instructional logs to collect information on their homework assignment and reviewing practices as well the online platforms control teachers have used to assign homework. Teachers reported using a variety of technology programs, including Kahoot, Schoolnet, iXL, Khan Academy, iReady, Quizlet, Desmos, MobyMax, Prodigy, BrainPoP, SuccessMaker, Freckle, etc., as well as online resources provided by curriculum providers.

Influence of COVID. The onset of COVID-19 forced the closure of study schools March 2020 and the move to remote instruction until the end of the school year across

all participating districts. This forced unplanned changes to the implementation of the study. It shortened the expected duration and lessened intensity of the use of ASSISTments by 2-3 months, which potentially could have affected the study's measured effect on student learning. ASSISTments use records indicated a significant drop in use coinciding with school closures. By the middle of April, 43% teachers in the treatment group resumed using ASSISTments to support their remote instruction, although at a reduced level compared to prior to school closures. Fig. 4 summarizes usage before and after the outbreak of the pandemic during the 2019-20 school year. Additionally, the state End-Of-Grade test (EOG), which was planned to serve as the primary outcome measure of student learning, was canceled for spring 2020. We had to request teachers to administer a supplemental online assessment as a replacement.

	Pre.Covid/In Person				Post.Covid/Distance			
Metrics	mean	sd	min	max	mean	sd	min	max
% Problems Completed by Students out of the problems they started	0.94	0.12	0	1	0.91	0.23	0	1
% of Assignments Started by Students	0.63	0.27	0	1	0.45	0.38	0	1
Average % of Assignments Completed by Students out of the ones they were assigned to	0.49	0.29	0	1	0.38	0.38	0	1
Average # of days per week students worked in ASSISTments	1.54	0.60	0	3.7	0.80	0.88	0	5
Average minutes per week students worked in ASSISTments	17.92	12.18	0	92.6	11.53	17.79	0	208.4
% of assignment reports viewed (avg per teacher)	0.46	0.41	0	1	0.62	0.33	0	1

Fig. 4. Summary of ASSISTments Usage during 2019-20

3 Implementation of ASSISTments at Scale

The adoption and successful implementation of a technology-based program in authentic school settings for an extended period is influenced by numerous factors. Below we discuss in detail how we addressed some of them during multiple stages of the study.

3.1 Recruitment

Recruitment participants for large school-based studies, particularly RCTs, poses a significant challenge. Schools are often engaged in various initiatives at any given time. Given the dynamic nature of district and school-level initiatives, gaining buy-in for a research-based program like ASSISTments can be difficult, even though it is a free platform that has demonstrated rigorous empirical impact. Some schools may view ASSISTments as less sophisticated than similar commercial tools and may already be using paid tools for similar purposes. Moreover, schools may be reluctant to adopt a new free tool due to associated costs and the learning curve involved.

To address these challenges, researchers treated recruitment as an advertising campaign and identified factors that supported or inhibited participation, responding accordingly. Uncertainty associated with randomization hindered schools' willingness to participate, which was mitigated by offering control schools free access to ASSISTments in a non-study grade level plus delayed treatment for control schools. North Carolina schools varied in their technology adoption and teacher preferences for learning management systems (LMSs). To overcome these concerns, the recruitment was confined to districts that have implemented 1:1 access to computers. ASSISTments was developed as a versatile tool that can be used on any device without internet access and

easily integrated with most common LMSs in North Carolina. The team examined district priorities based on public information and targeted recruitment efforts toward schools with a history of low math performance that prioritize the use of data to personalize student learning—a fit for the ASSISTments program.

3.2 Understanding school context

Early during the study, we conducted interviews with school principals to identify factors that could affect the fidelity of implementation. Participants were asked about: their concerns regarding students' performance in math; the school's major priorities or areas of focus; the technology available to students at school, specifically for math, and at home; and their homework and data-use policy. Principals in the treatment group were also asked whether they foresaw of any obstacles to the use of ASSISTments or any additional supports that would help teachers use ASSISTments. The interviews helped us discover what could facilitate the use of ASSISTments (e.g., school initiative focusing on improvement and student identity; common planning time within grade level; established homework policy) and what could inhibit its use in the schools (e.g., limited access to technology). We then customized the communication and support for each school to ensure the program align with the school' context for easier implementation.

3.3 Training and continuous support

Research suggested when both PD and formative assessment technology are provided, teachers can learn more about their students and adapt instruction, resulting improvements in student outcomes. In this study, a former high school teacher and remediation program organizer coached treatment teachers on integrating ASSISTments into instruction and provided ongoing coaching and technical support throughout the school year via webinars, video conferencing, and in-person visits. During the PD, treatment teachers learned techniques such as encouraging students to rework problems they initially got wrong, focusing on incorrect homework problems, reviewing solution processes for difficult problems, and addressing common misconceptions. The PD training helped teachers to utilize the technology's reports to make informed decisions, such as identifying which problems or students to prioritize. As a result, teachers were able to personalize instruction for individual students, groups, or the entire class; for example, they could assign additional practice to individual students (e.g., by assigning skill builders) or adding other problems or hints to existing problem sets.

3.4 Specifying a use model and expectation

To aid teachers in integrating ASSISTments into their existing instructional practices, the intervention was framed into a 4-step loop that aligns with common homework routines, as illustrated in Fig. 5. This approach helped teachers avoid the burden of figuring out how to utilize the program in their instruction and ensured consistency in program implementation. Per the use model, treatment teachers were expected to assign approximately 20-23 minutes of homework via the platform twice per week using a combination of textbook-based problems and pre-built content. Teachers were also expected to open reports for at least 50% of assignments. Opening the reports is a precursor to

adapting instruction. Teachers were given autonomy in determining the amount and type of homework to assign. The clearly specified expectation served as an objective for sufficient dosage of implementation, which is crucial for the effectiveness of any intervention.

3.5 Monitoring dosage and evaluating quality of implementation

Corresponding to the four-step loop, the ASSISTments team created a rubric (Fig. 6) that was used to assess teacher's progress in implementing the four steps during the school year. During the in-person school visits, the local coach evaluated each teacher's status according to the rubric and recorded their evaluation in a coach log, which was used to track a teacher's progress over time and gauge whether additional support needed to be provided to the teacher. The color-coded log (one row per teacher, Fig. 7) made it easy for the coach to identify teachers who have improved from visit 1 to visit 3 (e.g., rows 4 or 5), or teachers who regressed over time (e.g., rows 1 or 6).

	2	1	0
Finding and Assigning Content	Teacher knows which content they want to use and where to find it.	Teacher is somewhat clear on what content to use and where to find it.	Teacher is unable to find content they want to use.
Students Doing Content	Students are completing at least 9 problems per week.	Students are doing content some of the time. (1-2 times per week)	Students are not doing content. (Less than 1 time per week)
Teachers Viewing Report Data	Teacher is checking reports for at least 75% of problem sets.	Teacher reviewed some data in the item report.	Teacher is not reviewing data in the report.
Teachers Sharing Report Data with Students	Teacher accomplishes the following during a report review: 3. Overviewing data on problems with which students struggled 4. Reteaching/direct instruction, student discussion or other method used to engage with the data.	Teacher is reviewing reports with students some of the time. (1-2 times per week)	Teacher is not reviewing the report with students. (Less than 1 time per week)

Fig. 6. 4-Step Rubric for Evaluating Progress

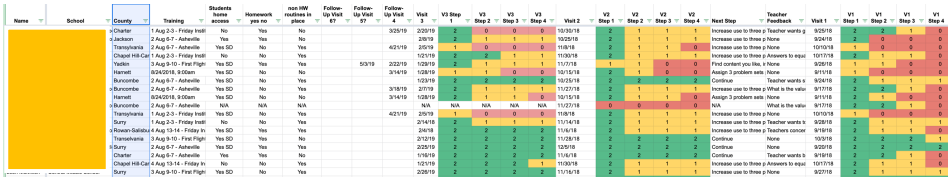

Fig 7. A Part of Coach Logs of School Visits during the 2019-20 School Year

4 Data Collection

Student Learning. We obtained student demographic, enrollment and prior grade state test performance data from the state-wide database hosted by the North Carolina Education Research Data Center (NCERDC)[2]. Student demographic data accessed include gender, race/ethnicity, free and reduced-price lunch status, and individualized education program status. Students' 6th grade EOG math scores served as a measure of prior

[2] Our agreement with NCERDC doesn't permit sharing of obtained data with any third parties. Other data collected during the study has been deposited to the Open ICPSR data repository (https://www.openicpsr.org/openicpsr/project/183645/version/V1/view).

achievement. As a replacement for the state test, the study team requested teachers to administer the online Grade 8 Mathematics Readiness Test (MRT) early May during remote learning time. The test is aligned with the Common Core State Standards. An analysis found a strong alignment between MRT and the NC Course of Study Standard Topics. Analysis of the response data suggested a correlation of 0.68 between students' MRT scores and their 6th grade EOG test scores from spring 2019, and the reliability of the test (KR20 for dichotomous items) was 0.94.

Teacher Homework Practice Measure. An interview was conducted with a sample of treatment and control teachers to their homework practices and the kinds of instructional supports the school and district provide to teachers. A quantitative measure of teacher homework review practice-Targeted Homework Review-was created from teacher self-reports on weekly instructional logs. All teachers were asked to complete daily 10-minute online logs during 5 consecutive instructional days across 3 collection windows across the school year for a total 15 instructional days. On average, the log completion rate was 81 percent for control teachers and 96 percent for treatment teachers. Teachers were asked if they reviewed assignments with the whole class and the ways they select which problems to review with the class (e.g., *I reviewed all the problems, I reviewed a sample of problems that I think my students might have difficulty on, I asked my students what problems they wanted to go over*).

Student and Teacher Use of ASSISTments. The primary source of ASSISTments use data was the electronic use records collected by the ASSISTments system. All student and teacher actions on the platform are time-stamped. Student system use data includes their login time, duration of each session, the number of problems attempted, and answered correctly and incorrectly, whether an assignment was completed on time, and student-to-system interactions during problem-solving, such as requesting hint or commenting on a question. Teacher use data includes the dates each assignment was created, the kind of problems assigned, the type of assignment (e.g., whole class, or individualized), assignment due date, and whether the teacher opened a report. These data allowed us to assess the extent to which students used ASSISTments to complete their math homework and the extent to which teachers assigned problems and monitored students' homework performance according to the implementation model.

5 Analysis and results

Student Learning Outcomes. For the impact analysis, we used a two-level hierarchical linear model (i.e., students nested within schools), controlling for school-level and student-level characteristics and students' prior achievement on their sixth-grade math state tests. The model didn't detect a significant difference between treatment and control students in their performance on the MRT test ($p > 0.05$).

Teacher Practice Outcomes. Analysis of teacher interview transcripts showed, all teachers in the treatment group noticed some changes in their instructional practices since they began using ASSISTments. Some of the changes identified by the teachers included: more thoughtful selection of problems; more one-on-one teaching; and being

more efficient and more targeted in their instruction. Most of them have found the ASSISTments reports to be valuable and have shared them with students by projecting the anonymized reports on a screen in front of the class.

An HLM model with teacher and school levels was posited to look at the effect of ASSISTments on teacher practices, adjusting for the same school level covariates as in the student achievement analysis. The results indicated that after adjusting for school characteristics, treatment teachers reported a significantly percentage (17.4) of times when they applied targeted homework review practice ($p < 0.05$).

Relationship between Usage and Learning Outcome. We conducted an exploratory analysis using the three-level hierarchical linear regression model using usage indicators to predict student achievement. Two usage indicators, *total number of problems completed by students*, and *total number of assignments at the classroom level*, are positively related to student performance on MRT. The results indicate that more homework assigned by the teacher through the ASSISTments platform is associated with higher math achievement ($p = .029$). More problems completed by a student is also associated with higher math achievement ($p < .001$).

6 Conclusion

Improving mathematics education is an important educational challenge. Many technology solutions have been developed and demonstrated evidence of promise during pilot studies, but many stayed in the research stage and very few have been effectively implemented in authentic school settings on a large scale for extended periods of time. In this paper, we reported on a large-scale efficacy study that explicitly adopted planned variation to test the impact of the ASSISTments program under a variety of setting and implementation characteristics. Particularly, we shared the procedures the team took to introduce the researcher-developed, free program to middle school teachers, and approaches to ensure consistency and faithful implementation of the program.

The results of our analyses suggested ASSISTments helped teachers practice the key aspects of rich in-depth formative assessment-driven instruction, and more use of the program was associated with higher student math achievement but didn't provide further evidence of the replicability of student learning impact on the selected measure.

The reported results are bound by several limitations related to the disruption of classroom instruction brought on by the onset of the COVID pandemic, including the unexpected alteration to the implementation model for ASSISTments and a shortening of the expected duration of the use of the intervention. The replacement outcome measure was not part of the state accountability assessment program and was administered at home during a time when teachers and schools were struggling to put together an instruction plan to engage students. Thus, the study experienced significant attrition in outcome data collection. Further analysis is ongoing to compare the sample, school context, implementation dosage, and counterfactuals between the original Maine study and the current study, and to better understand the nature of use of ASSISTments before and during the COVID pandemic.

Acknowledgement

This material is based on work supported by the Institute of Education Sciences of the U.S. Department of Education under Grant R305A170641. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the funders.

References

1. Heffernan, N. & Heffernan, C. (2014). The ASSISTments Ecosystem: Building a Platform that Brings Scientists and Teachers Together for Minimally Invasive Research on Human Learning and Teaching. *International Journal of Artificial Intelligence in Education. 24 (4),* 470-497. *DOI 10.1007/s40593-014-0024-x.*
2. Roschelle, J., Feng, M., Murphy, R., & Mason, C. (2016). Online mathematics homework increases student achievement. *AERA Open Oct 2016, 2 (4)* 2332858416673968; DOI: 10.1177/2332858416673968.
3. Murphy, R., Roschelle, J., Feng, M., & Mason, C. (2020). Investigating Efficacy, Moderators and Mediators for an Online Mathematics Homework Intervention, *Journal of Research on Educational Effectiveness*. DOI: 10.1080/19345747.2019.1710885
4. Sahni, S. D., Polanin, J. R., Zhang, Q., Michaelson, L. E., Caverly, S., Polese, M. L., & Yang, J. (2021). A What Works Clearinghouse Rapid Evidence Review of Distance Learning Programs. U.S. Department of Education.
5. Escueta, M., Quan, V., Nickow, A. J., & Oreopoulos, P. (2017). Education Technology: An Evidence-Based Review. NBER Working Paper No. 23744. August 2017 JEL No. I20,I29,J24. Retrieved from http://tiny.cc/NBER
6. Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M., et al. (2007). *Organizing instruction and study to improve student learning*. IES practice guide (NCER 2007-2004). Washington, D.C.: National Center for Education Research.
7. Heritage, M., & Popham, W. J. (2013). Formative assessment in practice: A process of inquiry and action. Cambridge, MA: Harvard Education Press.
8. Koedinger, K. R., Booth, J. L., & Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science, 342*(6161), 935–937. https://doi.org/10.1126/science.1238056.
9. Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, *77*(1), 81-112.
10. DiBattista, D., Gosse, L., Sinnige-Egger, J. A., Candale, B., & Sargeson, K. (2009). Grading scheme, test difficulty, and the immediate feedback assessment technique. *The Journal of Experimental Education*, *77*(4), 311-338.
11. Mendicino, M., Razzaq, L., & Heffernan, N. T. (2009). A comparison of traditional homework to computer supported homework. *Journal of Research on Technology in Education (JRTE)*. 41(3), 331–358.
12. Fyfe, E. R., & Rittle-Johnson, B. (2016). The benefits of computer-generated feedback for mathematics problem solving. *Journal of Experimental Child Psychology*, *147*, 140-151.
13. Kehrer, P. & Kelly, Kim & Heffernan, Neil. (2013). Does immediate feedback while doing homework improve learning? FLAIRS 2013 - Proceedings of the 26th International Florida Artificial Intelligence Research Society Conference. 542-545.
14. Kelly, K., Heffernan, N., Heffernan, C., Goldman, S., Pellegrino, J., & Soffer-Goldstein, D. (2014). Improving student learning in math through web-based homework review. In Liljedahl, P., Nicol, C., Oesterle, S., & Allan, D. (Eds.). (2014). Proceedings of the Joint Meeting of PME 38 and PME-NA 36 (Vol. 3). Vancouver, Canada: PME. pp. 417-424.

15. Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in education: Principles, policy & practice*, *25*(6), 551-575.
16. Bennett, R.E. (2011). Formative assessment: A critical review**.** Assessment in Education: Principles, Policy & Practice, 18 (1) (2011), pp. 5-25.
17. Dunn, K.E, & Mulvenon, S.W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research and Evaluation*, 14 (7) (2009), pp. 1-11
18. OECD (2005), *Formative Assessment: Improving Learning in Secondary Classrooms*, OECD Publishing, Paris, https://doi.org/10.1787/9789264007413-en.
19. Wiliam, D., & Thompson, M. (2007). Integrating assessment with learning: what will it take to make it work? In C.A. Dwyer (Ed.), The future of assessment: Shaping teaching and learning (pp. 53–82). Lawrence Erlbaum Associates.
20. Faber, J. M., Luyten, H., & Visscher, A. J. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. Computers & Education, 106 , 83-96.
21. Koedinger, K. R., McLaughlin, E. A., & Heffernan, N. T. (2010). A quasi-experimental evaluation of an on-line formative assessment and tutoring system. Journal of Educational Computing Research, 43(4), 489e510. http://dx.doi.org/10.2190/EC.43.4.d
22. Sheard, M., Chambers, B., & Elliott, L. (2012). Effects of technology-enhanced formative assessment on achievement in primary grammar. Retrieved from the University of York website https://www.york.ac.uk/media/iee/documents/QfLGrammarReport_Sept2012.pdf.
23. Anderson, J. R. (2000). *Learning and memory: An integrated approach* (2nd ed.). New York: John Wiley and Sons, Inc.
24. Koedinger, K., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with Cognitive Tutors. *Educational Psychology Review, 19*, 239–264.
25. Rohrer, D. (2009). The effects of spacing and mixing practice problems. *Journal for Research in Mathematics Education*, *40*, 4–17.
26. Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics practice problems improves learning. *Instructional Science*, *35*, 481–498.
27. Rohrer, D., & Pashler, H. (2007). Increasing retention without increasing study time. *Current Directions in Psychological Science, 16*, 183-186.
28. Fairman, J., Porter, M., Fisher, S. (2015). Principals Discuss Early Implementation of the ASSISTments Online Homework Tutor for Mathematics: ASSISTments Efficacy Study Report 2. Menlo Park, CA: SRI International.
29. Loveless, T. (2014). *The 2014 Brown Center Report on American Education: How well are American students learning?* The Brown Center on Education Policy. Washington, DC: The Brookings Institution.
30. Bennett, S., & Kalish, N. (2006). *The case against homework: How homework is hurting our children and what we can do about it*. New York, NY: Crown Publishers.
31. Kohn, A. (2006). *The homework myth: Why our kids get too much of a bad thing*. Cambridge, MA: Da Capo Press.
32. Pope, D (2020, August 27). *Making Homework Work*. ASCD. Retrieved October 20, 2021.