Adaptive Learning Goes to China

Mingyu Feng,[1[0000-1111-2222-3333]] Wei Cui,[2[1111-2222-3333-4444]] and Shuai Wang [1[1111-2222-3333-4444]]

[1] SRI International, Menlo Park, CA 94025, USA
[2] Shanghai Yixue Education Technology Inc., Shanghai, China
mingyu.feng@sri.com, cuiwei@classba.cn, sam.wang@sri.com

Abstract. In online education products, adaptive learning, by definition, adjusts the content and guidance offered to individual learners. Systems that offer adaptive learning grew out of computer science research that aimed to replicate the dynamic interactions between human tutors and learners. Studies have shown that these systems can be effective learning tools. This paper introduces an adaptive learning system, "Yixue," that was developed and deployed in China. It diagnostically assesses students' mastery of fine-grained skills and presents them with instructional content that fits their characteristics and abilities. The Yixue system, first developed in 2016, has been used by over 10,000 students in 17 cities in China for learning 12 subjects in middle school in 2017. The hypothesis is that the Yixue adaptive learning system will improve student learning outcomes compared to other learning systems. This paper describes major features of the Yixue system and its implementation model in class. A learning analysis of 1355 students indicates that students learned from using the Yixue system and the results can generalize across students and skills. We also report a study that evaluates the efficacy of the Yixue math program in 8th and 9th grade, compared with whole class instruction by expert human teachers. The result suggests that students learned more from using Yixue adaptive learning system than whole classroom instruction by teachers.

Keywords: Adaptive Learning, Mastery-based Learning, Diagnostic Assessment, Efficacy.

1 Introduction

Through machine learning algorithms and data analytics techniques, adaptive learning systems offer learning personalized to students' characteristics and abilities. The intent is to determine what a student really knows and to accurately, logically move the student through a sequential path to prescribed learning outcomes and skill mastery. Many learning products with adaptive features have been developed, such as Cognitive Tutors®, i-Ready®, DreamBox® Learning, Achieve3000®, Knewton®, RealizIt®, ALEKS®. Such systems constantly collect and analyze students' learning and behavior data and update learner profiles. As students spend more time in it, the system knows their ability better and can personalize the course to best fit their talents (Triantallou, Pomportsis, & Demetriadis, 2003; van Seters et al., 2012)

The widespread availability of computers and network connections inside and outside schools has drawn greater attention to technology-based learning systems. Studies have shown that such systems can be effective learning tools (VanLehn, 2011) and can promote student engagement. An analysis of learning data from 6,400 courses, 1,600 of which were adaptive, revealed that the adaptive courses were more effective in improving student performance than the 4,800 nonadaptive courses (Bomash & Kish, 2015). In general, meta-analyses have found positive impacts from technology-based interventions for mathematics and other subjects (e.g., Cheung & Slavin, 2013; Steenbergen-Hu & Cooper, 2013, 2014). Pane et al. (2014) conducted a large-scale effectiveness study of Cognitive Tutor Algebra I, a representative intelligent tutoring system, in diverse real-world school contexts without any extraordinary effort to optimize implementation. The study was conducted in 73 high schools and 74 middle schools in 52 school districts in seven U.S. states. The study found strong evidence in support of a positive effect in the second year after two years of use of the intervention. The magnitude was sufficient to improve the average student's performance by approximately 8 percentile points. More recently, a study by RAND funded by the Bill & Melinda Gates Foundation on personalized learning (Pane et al., 2017) found that over 2 years, personalized learning had positive effects on student mathematics and reading performance and that the lowest performing students made substantial gains relative to their peers. Additionally, growth continued to accumulate in the third year after schools began implementing personalized learning. All 62 schools involved in the Pane et al. (2017) study implemented personalized learning schoolwide, as opposed to class- or teacher-level implementation.

Many schools in the United States have adopted adaptive learning systems of one kind or another. Schools see a number of advantages to online or blended learning supported by technology. Adaptive systems are known for "meeting students where they are" and adjust to learner differences. Adaptive systems help close performance gaps by allowing low-performing students to get more practice on the topics that they haven't grasped before advancing. Such systems often use assessment as a formative gauge for where students should move next based on what they know and they don't know. Adaptive systems provide teachers with real-time, frequent, and detailed information on individual students' needs and their progress, so that teachers can adjust instruction accordingly. Teachers may change curriculum pacing, group students differently, and give more focused attention to students who require individualized intervention. Adaptive systems are seen as having the potential to improve student achievement by providing more engaging and personalized instruction and more immediate feedback. Adaptive systems also help introduce variety into the classroom; while one group of students

Online education has developed rapidly in China in recent years. According to the China Internet Network Information Center (2017), by December 2016 the number of online education users in China had reached 138 million, accounting for 19% of total Internet users. Yet the development of learning systems, especially systems that adapt to students' needs, is still in the early stages in China, and little empirical evidence exists on their promise in improving learning outcomes.

Yixue Inc. was one of the first organizations to develop an adaptive learning system in China. With the objective of introducing effective learning systems to China, an initial version of the Yixue was created and tested in 2016. Yixue Inc. has developed instructional materials for middle school mathematics, English, physics, Chinese, and chemistry and is working on expanding content coverage to the whole spectrum of K–12 education settings. In 2017, Yixue was used by over 10,000 students in 17 cities in China, representing a broad range of student populations with respect to socioeconomic status, urbanicity, and performance levels. The key feature of Yixue is that it leverages a fine-grained domain model, a "knowledge map," to adapt instructional content based on students' diagnostic assessment results to improve learning efficiency. Students' ability on individual knowledge points is constantly assessed while they practice problem solving with instructional assistance.

This paper describes in further detail the functions and reasoning architecture of the Yixue system and reviews the theoretical basis for the Yixue approach. We also present an analysis of learning results from Yixue based on learning data from 1,355 students. Finally, an experiment on the efficacy of the system conducted in 2017 and its finding is described.

2 The Yixue Adaptive Learning System

The Yixue intelligent adaptive learning system is a computer-based learning environment that adjust the content and guidance to individual students at both macro- and micro-levels. It provides many opportunities for practice and feedback (Martin, Klein, & Sullivan, 2007). Overall, the system has the following features:

- Fine-grained knowledge map in which knowledge components are organized hierarchically based on a learning progression relationship
- Diagnostic pre-assessment
- Adaptive, automated differentiated instruction
- Immediate feedback and stepped explanations for students
- Rich, high-quality learning repository of various types of learning content
- In-class support and intervention by teachers.

As a **macro-adaptive strategy** supported by psychometric measurement models and artificial intelligence, the system implements competency-based learning and tracks students' mastery of knowledge over time. In competency-based, or *mastery-based*, learning (Park & Lee, 2003), students advance to a new learning objective only when they demonstrate proficiency in the current one. In Yixue, students are first given a pre-assessment that diagnoses which knowledge components they have mastered and which ones they have not, according to the predefined hierarchical knowledge structure map. Thus, the system identifies the student's position in the domain model, and the student model is updated accordingly. Then students enter a learning-by-doing stage. The knowledge they demonstrated mastery on during the pre-assessment is skipped during the instructional phase, while knowledge components they were weak on are arranged in an optimal order for learning. As students work, the system simultaneously updates (a) estimates of their competency on each individual knowledge components using a Bayesian statistical model and (b) estimates of overall student proficiency level using an item response theory (IRT) model (van der Linden, 2016) and delivers individualized learning content to each student, such as instructional videos, lecture notes, worked examples, embedded practice problems, or unit tests.

Yixue implemented a differentiated instruction model in which students who progress at different paces receive different supports (Subban, 2006; Tomlinson & McTighe, 2006). Students who progress quickly through the basic level of a topic and demonstrate a higher ability level are given more challenging problems to accelerate their preparation for advanced topics.

The **micro-adaptive strategy** in the Yixue design has to do with provision of just-in-time feedback to students. After a student submits his/her solution to a problem, the system provides immediate feedback on the correctness of the answer. Students may attempt a problem multiple times and request an elaborated explanation of the solution processes step by step if they encounter difficulty. To increase learning efficiency and prevent students from wasting time in overattempting (such as taking a guess-and-check approach and repeatedly entering incorrect answers), the system stops students from trying after three failed attempts and displays the explanation. For selected subjects where misconceptions are common (e.g., physics), if the system detects a student repeatedly making the same kind of error after a number of practices, it automatically plays an instructional video addressing the misconception associated with the error.

Fundamental to the Yixue design is automatic collection of student performance data and provision of feedback and reports to students and teachers. As students work, the Yixue system automatically collects their responses to questions. Students are constantly presented with summary information on the screen (Figure 1) including how long they have practiced, their rate of correctness, and their progress toward completing the learning goal of the current lesson. At the end of each session, the system presents students with a summary report on how they performed with direct links to problem solutions and instructional videos on each topic. While students practice in Yixue, the system uses their response data in real time to update the estimates of student ability on related knowledge components in the student model and adjusts subsequent content accordingly. In Yixue, data and feedback are continuously available to teachers through reports they can use to make instructional decisions, hence enabling formative assessment, another practice with a strong research basis (see review below).

2.1 Illustration of a Yixue Learning-by-Doing Math Problem

When using Yixue, students spend a substantial portion of classroom time practicing, corresponding to the principle of academic time on task (Stallings, 1980). Figure 1 is a screen shot of a mathematics problem in the Yixue system. At the center of the screen is a math problem that the student needs to solve. The top left corner of the screen shows the focal knowledge that the problem addresses. The top right corner shows measures of practice and the student's progress, including how long he or she has been learning within the system, the percentage of problems correct so far, and, of all the

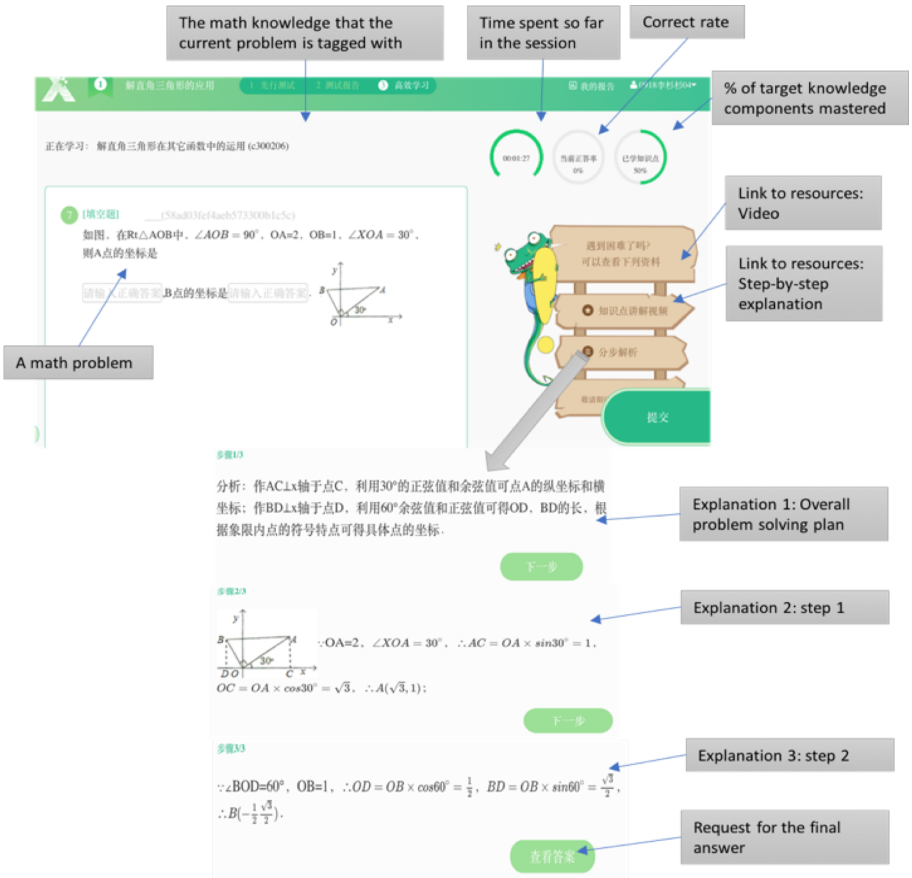

Fig. 1. Illustration of a Yixue math problem.

knowledge components that the student is weak on, what percentage of them the student has mastered through learning in Yixue. On the right side of the screen are buttons that, when clicked, will bring up resources the student may refer to if he/she has difficulty solving the problem. The student can ask to watch a short video (usually 3–5 minutes) in which a teacher explains the knowledge component and demonstrates how to solve similar problems, or the student can request step-by-step explanation of the problem. The bottom half of Figure 1 shows three messages. The first one provides a high-level problem-solving plan overall, but does not scaffold the problems or show single steps. The purpose of presenting the plan at first is to prompt the students to come up with their own solutions given the overview. If a student still has trouble moving on, he/she can click **Next** to request explanation 2, which shows the first step to solve the problem, elaborating on explanation 1. The explanation is available only upon the student's request, and the student can go back to attempting to solve the problem at any time. In the end, the student can click the **Show me the answer** button for the system to display the correct answer. In doing so, the student will be prompted that this problem will be evaluated as “incorrect” and asked if he/she wishes to continue with the request.

Note that even if a student provides the correct answer to the problem on the first attempt, the system still presents the full explanation, including explanations 1 to 3, plus all forms of correct answers to prompt the student to compare his/her solution with the system-provided one(s) to reinforce the student’s understanding.

2.2 Yixue Implementation Model

A Yixue classroom appears different from a typical classroom because students are expected to spend much of their time independently learning content and solving problems on a computer, with frequent feedback, instructional support, and use of features unique to the delivery of instruction on a computer (e.g., videos, animations, %correct) that are designed to keep their engagement high throughout the instructional period. In addition, teachers have important and defined roles in the Yixue classroom. Teachers receive easy-to-read reports on classroom and students’ individual progress, and they are expected to use the data to identify individual students who might be struggling and what they are struggling with and provide them with targeted support during class time. While students are engaging in learning with technology support, the teacher has more time available to respond to individual needs. In this way, Yixue supports a synergistic blending of the teacher’s and technology’s roles in delivering instruction and supports differentiated instruction. In addition to the offline working mode where teachers and students work together in person, Yixue also supports an online learning mode where teachers and students hold virtual meetings using videoconferencing tools.

3 Theoretical Basis for Yixue Approach and Features

Yixue was developed on a strong theoretical and research-based foundation.

Cognitive diagnostic assessment. In the groundbreaking book Knowing What Students Know (Pellegrino, Chudowsky, & Glaser, 2001), a compelling case is made to integrate the best from the worlds of learning theory (cognitive psychology) and assessment design and measurement (psychometrics). Leighton and Gierl (2007) described cognitive diagnostic assessments as “designed to measure specific knowledge structures and processing skills in students so as to provide information about their cognitive strengths and weaknesses” (p. 3).

Immediate and frequent feedback to the learner. An extensive set of studies (e.g., Clariana, 1990, 1993; Mendicino, Razzaq, & Heffernan, 2009; Morrison, Ross, Gopalakrishnan, & Casey, 1995; Shute, 2008; Whyte, et al., 1995) have shown significant learning gains in response to feedback in computer-based instruction. The effects reported were greater when the feedback was immediate and elaborated. Research has shown that frequent feedback increases student learning (Hattie & Timperley, 2007; Kluger & deNisi, 1996). Azevedo and Bernard (1995), based on their meta-analysis, stated that “immediate delivery of a feedback message provides the best instructional advantage to the student” (p. 122). The advantages are both a quick confirmation of a student's correct understanding and a rapid indication of where a misunderstanding has occurred and further clarification is needed.

Mastery learning. According to the learning principle of mastery learning (Park & Lee, 2003), the presentation of new content should be paced according to careful estimates of how much new content a learner can take on, based on the cognitive load theory (Sweller, 2003). The positive impacts of mastery learning adaptive learning systems have been established in empirical studies (Anderson, 2000; Gusky & Gates, 1986; Koedinger & Aleven, 2007; for review, see Kulik, C., Kulik, J., & Bangert-Drowns, R., 1990, and Durlach & Ray, 2011).

Individualized tutoring. Human one-on-one tutoring is highly effective, with experienced tutors achieving a 2 standard deviation improvement over classroom instruction (Bloom 1984). Interactive learning environments that mimic aspects of human tutors, such as intelligent tutoring systems, have been highly successful as well (Koedinger, Anderson, Hadley, & Mark, 1997; VanLehn et al., 2005). A meta-analysis of findings with 107 effect sizes involving over 14,000 participants (Ma, Adesope, Nesbit & Liu, 2014) showed that the use of intelligent tutoring systems was associated with greater achievement in comparison with teacher-led large-group instruction ($g = 0.42$). Steenbergen-Hu & Cooper (2014) reviewed 39 studies assessing the effectiveness of 22 types of intelligent tutoring systems in higher education settings. Overall, the systems had a moderate positive effect on college students' academic learning ($g = 0.32$ to $g = .37$).

Differentiated instruction. Differentiation means tailoring instruction to meet individual needs and teachers may differentiate content, process, products, or the learning environment (Tomlinson, 2000). Subban (2006) synthesized research on differentiated instruction from 1980 to 2005 and pointed out teachers need to know how to respond to the burgeoning diversity of contemporary classrooms as suggested by many studies. She noted that differentiated instruction presents an effective means to address learner variance (Tomlinson, 2000, 2001, 2003), avoids the pitfalls of the one-size-fits-all curriculum (McBride, 2004), incorporates current research into the workings of the human brain (Tuttle, 2000) while supporting the multiple intelligences and varying learning styles (Lawrence-Brown, 2004; Tuttle, 2000) within contemporary classrooms. Addressing student differences also appeared to enhance their motivation to learn while encouraging them to remain committed and stay positive (Stronge, 2004).

Formative assessment. Checking for student understanding is inherent to good instructional practice (Wiliam, 2016). The literature on mathematics education has recommended technology for formative assessment (Drijvers et al., 2016). The concept of formative assessment gained currency in the United States about 20 years ago and has received much attention in K–12 research and practitioner communities (Black & Wiliam, 1998a 1998b; Boston 2002; Heritage & Popham, 2013; Roediger & Karpicke, 2006). Leading researchers and practitioners characterize formative assessment as a process that uses student data to inform adaptive changes in instruction (Bennett, 2011; Black & Wiliam, 2009; Brookhart, 2007; Guskey, 2007; Heritage & Popham, 2013). Research documents modest to medium effect sizes of formative assessment on student learning (Black & Wiliam, 2009; Brookhart, 2007; Guskey, 2007; Hattie, 2009; Kingston & Nash, 2012; Shavelson, 2008; Speece, Molloy, & Case, 2003; Thum, et al.,

2015) for a variety of different modes, grade levels, content areas, and cultural settings. Frequent use of formative assessments can improve achievement, particularly when the results are used to adjust instruction (Bergan, Sladeczek, Schwarz, & Smith, 1991; Speece, Molloy, & Case, 2003).

4 Research Questions and Findings

In this paper, we aimed to answer two research questions (RQ):

1. Does Yixue effectively teach students?
2. Does Yixue have promise of improving student learning, compared with other online learning products and classroom teachers?

We addressed the research questions through analysis of system use data and experimental studies as described in the following two sections.

4.1 Analysis of Data to Determine Learning Effects

We addressed the first research question by looking at data from Yixue use to see if students demonstrated better performance after using the system. In Yixue, each item is tagged with a focal skill that it addresses. Learning was assessed by comparing students' performance on the first item they were given with their later performance on the second, third, and forth items on the same skill. If students tend to perform better on later opportunities at a skill, this indicates that they may have learned from the instructional assistance the prior items in the group provided. To see whether learning occurred and was generalized across students and skills, we conducted both a student-level analysis and a skill-level analysis. The hypothesis was that students were learning on groups of items that tapped similar skills. The particular dataset we analyzed was based on student use of Yixue mathematics programs in summer and fall 2017 across multiple grade levels.

For the student-level analysis, we set the criteria to include students who had worked on at least 10 skills and had at least four opportunities on each skill. A total of 1,355 students fit the criteria. We calculated average percentage correct on the four opportunities for all the sets of similar skills that they participated in and then conducted a t test to see if their performance was better at later opportunities. The results showed that the percentage correct increased significantly from students' first opportunity ($p = 0.03$) to the second opportunity and then continued to increase from the second to the third and the forth opportunities ($p < 0.01$).

For the skill-level analysis, we included only skills that had been studied by more than 20 students, with each student completing four problems addressing the skill. There were 662 different sets of skills that met the criteria for this analysis. We conducted t tests to see if the average percentage correct per skill increased significantly, and the increase was significant from each opportunity to the next ($p < 0.01$).

Overall, results from the student-level and item-level analyses suggested that students learned from using Yixue math products, and learning generalized across skills and students.

4.2 A Quasi-Experiment Comparing the Effects of Yixue with Classroom Instruction by Expert Teachers

In addition to examining learning within the system, we conducted studies to evaluate the efficacy of Yixue math middle school programs relative to other popular online learning programs, and classroom teachers. Primary purposes of the studies were to see if Yixue has the promise on improving student achievement and to generate estimates of the likely magnitude of effects of Yixue products, if they should exist. In Li et al. (in press), we reported a small scaled randomized controlled trial that was conducted in 2016. The experiment compared Yixue with Magic Grid (http://www.mofangge.com), another popular online learning platform, to examine the promise of the YiXue platform on improving student learning. 87 8th grade students participating the study were randomly assigned to treatment or control groups using stratified block randomization based on their grades in school (Trochim, Donnelly, & Arora, 2016, p229). A pre-test was administered and then students learned one English topic online for 100 minutes during the study, using Yixue (treatment) or Magic Grid (control). In the end, a post-test was given to all students. Analysis of the pre-test and post-test scores showed that after controlling for their pre-test scores, students who used YiXue on average scored 3.8 points higher on the post-test than students who used Magic Grid, and the difference was marginally significant ($F(2, 84) = 104.6$, $p = .09$, $r^2 = 0.71$). Post-study surveys showed that students had higher satisfaction with Yixue, especially regarding "ease of use" and "learning efficiency."

In this paper, we describe another study that was conducted in Oct. 2017, aiming at comparing efficacy of Yixue with whole group instruction provided by expert teachers.

Sample. 78 students from two grade levels (grades 8 and 9) were sampled from local schools and were assigned to conditions based on geographic convenience. Of the 78 students, 38 were assigned to the treatment condition where they used the Yixue system as the intervention, and the other 40 students were assigned to the control condition. Students in the control condition were then split into three groups that received whole group instruction from three teachers from local schools[1]. The teachers had taught middle or high school math for 8 to 18 years.

Study procedure. The study lasted for 4 days, during which each student received about 5 hours of instruction each day. The content covered during the instructional sessions included the Pythagorean theorem and its application, rational numbers, expressions, properties of a triangle, and line symmetry. These are representative of content covered in the grade 8 curriculum in local schools. Students in the treatment group were assigned user accounts in the Yixue system and worked on topics above individually during the study with no teacher assistance. In the control group, teachers taught the topics according to learning standards of the province. Students in the control group were not supported by any online learning programs.

Data source. Math pre- and post-test were administered to both group of students before and after the instructional sessions. Items in the tests were constructed by an experienced math teacher in a local school (not a part of the research team). When

[1] No students received instructions from their regular math teachers in school.

developing the tests, the teacher was only provided the math topics and the associated learning standards. Two independent, experienced subject matter experts reviewed the pre- and post-tests to ensure that they were comparable in their coverage, overall difficulty, types of items, and alignment with the learning standards. The experts also checked to make sure that the test items were not over-aligned with YiXue learning content. Each test composed of 14 multiple choice, short answer, or constructed responses problems, with a total score of 100 points. Students were given 60 minutes to complete the test. All students submitted their answers within the given time.

Analysis and results. First, we looked to see if students learned during the study by comparing their scores from the posttest to the pretest. We found that across all students, pretest scores were highly correlated with posttest scores ($r = 0.67$). The pretest mean score was 34.62, and the posttest mean score was 65.02. A paired t test showed that students' posttest scores were significantly higher than pretest scores ($p < 0.01$), suggesting that math achievement improved significantly after the experiment.

Differences in baseline characteristics for the students in the analytic sample were examined before statistical analysis. This step was especially critical considering that students were not randomly assigned to conditions. If students who used Yixue and those who did not use it had dissimilar prior achievement, then substantial error may have been introduced and statistical matching would be necessary before analyzing the treatment effect (What Works Clearinghouse [WWC], 2017a). We examined whether or not students in the treatment and control groups had similar pretest scores in grade 8 and grade 9 separately. The pretest total score was balanced between the treatment group and control group for grade 8 ($g = 0.13$) and grade 9 ($g = 0.24$). Since effect sizes on pretests in both grade levels were smaller than 0.25, statistical matching was not conducted (WWC, 2017a).

We then used generalized linear modeling to analyze student achievement, with the student posttest scores as the outcome variables, adjusting for pretest score, and treatment condition as a predictor at the student level. The resulting model was used to estimate effect size.

If Y_i is the posttest measurement of the i-th student, X_i is the pretest measurement, $I(t)$ is an indicator of treatment group membership, then the model is as follows:

$$Y_i = B_1 X_i + B_2 * I(t)_i + e_i$$

Here e_i is uncorrelated residual error terms at the student level.

The results showed that when student prior achievement was controlled, grade 8 Yixue students ($M = 69.96$, $SD = 22.34$) had significantly higher posttest scores than control students ($M = 61.40$, $SD = 20099$) ($b = 10.56$, $F\ (1, 32) = 3.35$, $p = 0.08$, $R^2 = 53.08$). There was a substantial effect size for the intervention, $g = 0.48$. This corresponds to an improvement index of 18 percentile points, indicating that the intervention would have led to an 18% increase in percentile rank for an average student in the control group and that 68% of the students in the intervention group scored above the control group mean (WWC, 2017b).

We also found that when student prior achievement was controlled for, grade 9 Yixue students ($M = 65.20$, $SD = 20.74$) had higher posttest scores than control students ($M = 61.79$, $SD = 22.56$), but the difference was not significant ($b = 7.24$, $F(1, 40) =$

2.02, $p = .16$, $R^2 = 49.38$). Yet there was a substantial effect size for the intervention, $g = 0.32$. This corresponds to an improvement index of 13 percentile points, indicating that the intervention would have led to a 13% increase in percentile rank for an average student in the control group and that 63% of the students in the intervention group would have scored above the control group mean.

Note that the statistical tests were underpowered and thus the *p*-values for the associated test were large partly because of the small sample sizes in this study. However, the large effect size did indicate that the Yixue students outperformed control students across grade 8 and grade 9.

5 Conclusion

The Yixue adaptive learning system was launched in 2016 and presently has over 100,000 users. In this paper, we introduced features of the system, its implementation model, and theoretical basis. Promising evidence was found that students learned during their use of the system, and a small-scale quasi-experiment demonstrated the efficacy of the system. Overall, the results suggested that the YiXue adaptive learning system has promise for improving student learning outcomes, potentially more effectively than whole group instruction by human teachers.

The study did have limitations; it was quasi experiment with relatively small sample. The study was conducted during a short time (4 days) and focused only on selected math topics. Additionally, we were not able to use an external standardized outcome measure. Thus, further research is warranted to examine the efficacy of the YiXue adaptive learning system. The team plans to carry out another efficacy study in the near future that that is appropriately powered to detect statistically significant differences between Yixue and business-as-usual classroom instructions or other technology-supported learning programs, if they should exist. Meanwhile, we are analyzing student learning log data to investigate the relationship between learning process and the learning outcome, and whether and how the learning gain differed across students of different incoming knowledge as measured by the pre-test. For instance, analysis is being conducted to see if higher learning gains in YiXue is associated with longer learning time, better performance within the system, or longer engagement time with videos.

As the technology infrastructure continues to develop in China, there are an increasing number of learning systems developed. However, there is little careful investigation of learning results from such systems. Very few experimental studies have been done in China to evaluate effectiveness of learning systems. With schools in China start to introduce educational technologies into classrooms, there is broad interest in how to select and use such systems and whether and which ones may lead to improvement. With these studies, we have the opportunity to contribute to much-needed knowledge about adaptive learning in K–12 instruction in China.

References

1. Anderson, J. R. (2000). *Learning and memory: An integrated approach* (2nd ed.). New York: John Wiley and Sons, Inc.
2. Azevedo, R., & Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research, 13*(2), 111–127.
3. Bergan, J., Sladeczek, I., Schwarz, R., Smith, A. (1991). Effects of a Measurement and Planning System on Kindergartners' Cognitive Development and Educational Programming. *American Educational Research Journal*, 28(3), 683-714.
4. Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, *5*, 7–74.
5. Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139–149.
6. Bloom, B. S. (1984). The 2-sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*(6), 4–16.
7. Bomash, I., & Kish, C. (2015). The improvement index: Evaluating academic gains in college students using adaptive lessons. New York, NY: Knewton.
8. Cheung, A.C. & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research review, 9,* 88-113.
9. China Internet Network Information Center (2017). *Statistical survey report on Internet development in China.* Retrieved March 15, 2017 from http://www.cnnic.cn/hlwfzyj/hlwxzbg/hlwtjbg/201701/P020170123364672657408.pdf
10. Clariana, R. B. (1990). A comparison of answer-until-correct feedback and knowledge-of-correct-response feedback under two conditions of contextualization. *Journal of Computer-Based Instruction, 17(4)*, 125–129.
11. Clariana, R. B. (1993). A review of multiple-try feedback in traditional and computer-based instruction. *Journal of Computer-Based Instruction*, *20*, 67–74.
12. Drijvers, P., Ball, L., Barzel, B., Heid, M.K., Cao, Y. & Maschietto, M. (2016). Uses of technology in lower secondary mathematics education: A concise topical survey. In *ICME-13 Topical Surveys*. New York: Springer International Publishing. 1-34. DOI: 10.1007/978-3-319-33666-4_1
13. Durlach, P. J., & Ray, J. M. (2011). *Designing adaptive instructional environments: Insights from empirical evidence* (Technical Report 1297). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. Retrieved from http://www.adlnet.gov/wp-content/uploads/2011/11/TR-1297.pdf
14. Guskey, T. R., & Gates, S. (1986). Synthesis of research on the effects of mastery learning in elementary and secondary classrooms. *Educational Leadership, 43*, 73–80.
15. Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77,* 88–112.
16. Heritage, M., & Popham, W. J. (2013). *Formative assessment in practice: A process of inquiry and action.* Cambridge, MA: Harvard Education Press.
17. Kluger, A. N., & deNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284.
18. Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8*, 30–43.

19. Koedinger, K., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with Cognitive Tutors. *Educational Psychology Review, 19*, 239–264.
20. Kulik, C., Kulik, J., & Bangert-Drowns, R. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research*, 60(2), 265–306.
21. Lawrence-Brown, D. (2004). Differentiated instruction: Inclusive strategies for standards based learning that benefit the whole class. *American Secondary Education*, 32(3), 34-62.
22. Leighton, J. P., & Gierl, M. J. (2007). Why cognitive diagnostic assessment? In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive assessment for education: Theory and applications* (pp. 3–18). New York, NY: Cambridge University Press.
23. Li, H., Cui, W., Xu, Z., Zhu, Z., Feng., M. (in press). Yixue Adaptive Learning System and Its Promise on Improving Student Learning. To appear in *Proceedings of the 10th International Conference on Computer Supported Education.* Funchal, Madeira-Portugal. March 2018.
24. Ma, W., Adesope, O., Nesbit, J., & Liu, Q. (2014). Intelligent Tutoring Systems and Learning Outcomes: A Meta-Analysis. *Journal of Educational Psychology. 106*(4). 901-918.
25. Martin, F., Klein, J., & Sullivan, H. (2007). The impact of instructional elements in computer-based instruction. British Journal of Educational Technology, 38(4), 623–636.
26. McBride, B. (2004). Data-driven instructional methods: "One-strategy-fits-all" doesn't work in real classrooms. *T.H.E Journal*, 31(11), 38-40.
27. Mendicino, M., Razzaq, L., & Heffernan, N. T. (2009). Comparison of traditional homework with computer supported homework: Improving learning from homework using intelligent tutoring systems. *Journal of Research on Technology in Education (JRTE). 41*(3), 331–359.
28. Morrison, G. R., Ross, S. M., Gopalakrishnan, M., & Casey, J. (1995). The effects of feedback and incentives on achievement in computer-based instruction. *Contemporary Educational Psychology, 20*, 32–50.
29. Pane, J., Griffin, B.A., McCaffrey, D.F., & Karam, R. (2014). Effectiveness of Cognitive Tutor Algebra I at Scale. *Educational Evaluation and Policy Analysis*, v. 36, no. 2, June 2014, p. 127 - 144.
30. Pane, J.., Steiner, E., Baird, M., Hamilton, L., and Pane, J. (2017). How Does Personalized Learning Affect Student Achievement? Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/research_briefs/RB9994.html.
31. Park, O., & Lee, J. (2003). Adaptive instructional systems. In D. H. Jonassen & M. Driscoll (Eds.), Handbook of research for educational communications and technology (2nd ed., pp. 651–684). Mahwah, New Jersey: Laurence Earlbaum. Retrieved from http://www.aect.org/edtech/ed1/25.pdf
32. Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy Press..
33. Shavelson, R., Young, D., Ayala, C., & Brandon, P., Furtak, E., Ruiz-Primo, M., Tomita, M., & Yin, Y. (2008) On the Impact of Curriculum-Embedded Formative Assessment on Learning: A Collaboration between Curriculum and Assessment Developers. *Applied Measurement in Education*, 21(4), 295-314, DOI: 10.1080/08957340802347647.
34. Shute V. (2008). Focus on formative feedback. *Review of Educational Research*. 78(1), 153-189.
35. Speece, D. L., Molloy, D. E., & Case, L. P. (2003). Responsiveness to general education instruction as the first gate to learning disabilities identification. *Learning Disabilities: Research & Practice, 18*(3), 147–156.
36. Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology*, *106*(2), 331-347.

37. Stronge, J. (2004). Teacher effectiveness and student achievement: What do good teachers do? Paper presented at the American Association of School Administrators Annual Conference and Exposition, San Francisco, California.
38. Subban, P. (2006). Differentiated instruction: A research basis. *International Education Journal*, *7*(7), 935–947.
39. Sweller, J. (2003). Evolution of human cognitive architecture. In B. H. Ross (Ed.) The psychology of learning and motivation: Advances in research and theory, Vol. 43 (pp. 216-261). New York, NY: Academic Press.
40. Roedinger, H., III., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention, *Psychological Science*, *17*(3), 249–255.
41. Thum, Y., Tarasawa, B., Hegedus, A., Yun, X., Bowe, B. (2015). *Keeping learning on track: A case study of formative assessment practice and its impact on learning in Meridian school district.* Portland, OR: NWEA. Retrieved from https://files.eric.ed.gov/fulltext/ED567844.pdf
42. Tomlinson, C. A., & McTighe, J. (2006). *Integrating differentiated instruction and understanding by design*. Alexandria, VA: Association for Supervision & Curriculum Development.
43. Tomlinson, C. A. (2004). Research evidence for differentiation. *School Administrator*, 61(7), 30.
44. Tomlinson, C. A. (2000). *The Differentiated Classroom: Responding to the Needs of all Learners*. Alexandria: Association for Supervision and Curriculum Development.
45. Tomlinson, C. A. (2001). Differentiated instruction in the regular classroom. *Understanding Our Gifted*, 14(1), 3-6.
46. Tomlinson, C. A. (2003). Deciding to teach them all. *Educational Leadership, 61*(2), 6-11.
47. Triantallou, E., Pomportsis, A., & Demetriadis, S. (2003). The design and the formative evaluation of an adaptive educational system based on cognitive style. *Computers & Education, 41*, 87–103.
48. Tomlinson, C. A. (August, 2000). Differentiation of Instruction in the Elementary Grades. *ERIC Digest.* ERIC Clearinghouse on Elementary and Early Childhood Education.
49. Trochim, W., Donnelly, J., Arora, K. (2016). *Research Methods: The Essential Knowledge Base*. Boston, MA: CENGAGE Learning.
50. Tuttle, J. (2000). *Differentiated Classrooms.* Woodbury: Cedar Mountain Academy.
51. vanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197–221.
52. van der Linden, W. (2016). *Handbook of Item Response Theory*. Boca Raton, FL: Chapman and Hall/CRC.
53. van Seters, J. R., Ossevoort, M. A., Tramper, J., & Goedhart, M. J. (2012). The influence of student characteristics on the use of adaptive e-learning material. *Computers & Education, 58*, 942–952.
54. What Works Clearinghouse (WWC) (2017a). *What Works Clearinghouse standards handbook (Version 4.0)*. Washington, DC: Institute of Education Sciences, US Department of Education.
55. What Works Clearinghouse (WWC) (2017b). *What Works Clearinghouse procedures handbook (Version 4.0)*. Washington, DC: Institute of Education Sciences, US Department of Education.
56. Whyte, M. M., Karolick, D. M., Neilsen, M. C., Elder, G. D., & Hawley, W. T. (1995). Cognitive styles and feedback in computer-assisted instruction. *Journal of Educational Computing Research, 12*(2), 195–203.

57. VanLehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., Treacy, D.,Weinstein, A., & Wintersgill, M. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence and Education, 15*(3), 1–47.