Rafael Laya

Jeannette Idiart

English& 102

17 June 2019

Intelligent Tutoring Systems: An Effective Complement to the Classroom

Have you ever imagined not having the grids that bring electricity to your home? not having a refrigerator to store food? Not having an email, phone or a computer with internet to communicate with your family and friends? or even the cars, planes and buses that allow you to move with ease? Science, technology, engineering, and mathematics (STEM) make it possible for us to have these beloved devices of the modern world. The US has long been a leader in STEM. However, the president's council of advisors on science and technology reported that American students are falling behind internationally in STEM education ("Prepare and Inspire" 1-2). This confirms that STEM is an important area for the wellness of the United States (US) and in need of improvement. In the same document, it was suggested to use technology to conduct education leadership ("Prepare and Inspire" 4). Therefore, it is reasonable to consider potential technological solutions to the stated problem. Given that mathematics is a tool that makes science and engineering possible, it is essential to worry about mathematics. A promising family of computer programs that may help students are Intelligent Tutoring Systems. An Intelligent Tutoring System (ITS) is a program that makes use of Artificial Intelligence (AI) to adapt to each student by choosing what material to teach and what strategy to use (Alkhatlan and Kalita 1). An ITS is then a program that tries to provide a personalized experience to each student. ITS programs must be able to figure out what type of student is using the program to identify the best learning path, combined with appropriate hints and feedback. While ITSs are currently as effective as human one-on-one tutoring and can benefit learning and outcomes in high school mathematics for US students, research shows that ITSs should be used as a complement in the classroom rather than a replacement. One way that ITSs can benefit student learning and outcomes in high school mathematics is when used for homework assignments.

Before comparing human tutors and ITSs, it is essential to understand a few terms. First, ITSs use Artificial Intelligence (AI). AI can be defined as the study of flexible systems that make proper choices to adapt in dynamic environments and circumstances, given past experiences and a set of limitations (Poole et al. 1). The most important part to remember from this definition is that these systems can adapt to different conditions, which is what most ITSs do to give a personalized experience to each student. In the present work, the system will be a regular computer running an ITS program. Some of the sources used in this paper to compare ITSs and human tutoring are meta-analyses. These are studies that integrate the findings of many other studies. Formally, Gene Glass defines the term as the analysis and combination of a large set of results from different independent sources (in Cheung 1). A meta-analysis does not study a single case, but rather, they compile a big number of results from other researchers and analyze the bigger picture to make conclusions about the typical case.

It is fundamental to understand tutoring. According to Chi, tutoring is all the segments of instruction where an agent called tutor and an agent called tutee interact constantly (1). Chi's definition of tutoring means that a tutor must collaborate with the student during the learning process. Particularly, the definition of human one-on-one tutoring used in this paper requires the agent to be an adult who is considered proficient in the field of study, and who must engage with the student in real time (VanLehn 1). The given definition allows making a fair comparison between tutors and ITSs, since the tutor must have good knowledge of the subject of study and is constantly communicating with the student. Without such a requirement, the comparison would be too broad and include unskilled tutors. Group tutoring will use the same definition, not enforcing the tutee to be a single student. It is of equal importance to understand what constitutes a traditional classroom. A traditional classroom is a classroom where a group of people is learning centered on an instructor, disregarding the different abilities and learning styles of each student (Li et al. 1-2). The definition provides a specific frame of reference to compare the effectiveness of ITSs. A traditional classroom, by definition, cannot have an ITS. This is because an ITS adapts to the style and abilities of each student, and it also grabs some attention from the main instructor. Therefore, a traditional classroom with an ITS would not center the learning around one instructor. What I need to evoke in the reader's mind when I mention traditional or conventional classrooms is that these types of classrooms cannot, by definition, have an ITS.

The performance of human tutors motivated AI researchers to develop ITSs. Because human tutors seem extremely effective, ITSs researchers have developed ITSs trying to mimic human tutoring behavior. In 1984, Bloom challenged researchers to find teaching strategies with effects similar to human tutoring, based on the experiments of his students that found human tutoring of groups of three or fewer students to be radically effective (Bloom in Bagheri 2; Kulik and Fletcher 4). Thus, Bloom believed, from experimentation, that human tutoring of small groups of students could drastically improve the learning outcomes of students compared to big groups. To illustrate, Ferester explains that solving Bloom's challenge is the equivalent of a student who increases from 500 to 700 points in the Scholastic Assessment Test (SAT) after tutoring (115). It is easy to imagine how this result sent an alert to researchers who were trying to improve education, since small group tutoring highly outperforms group instruction. One solution can be to assign a tutor to each student. However, individual tutoring has been impossible due to the extensive use of resources that one-on-one tutoring implies (Bagheri 2). Tutoring time costs money: if one wants to implement such a system in a classroom, then everyone would have to pay higher tuition to compensate for tutoring hours. The effectiveness of human tutors has incentivized researchers to develop successful ITSs as an alternative to replicate human tutoring (Alkhatlan and Kalita 13; Ferester 115; Koedinger and Aleven, "Exploring the Assistance Dilemma" 2). Bloom set up the challenge of replicating the success of human tutoring, and researchers have been trying for decades through ITSs. Knowing the motivation behind ITSs, it is necessary to assess the performance of human tutors, and then evaluate the performance of ITSs compared to their human role model.

Bloom's conclusion does not apply to the typical tutor, but ITS developers can make computers that perform like the typical tutor. Even though human tutors inspired researchers, these results are not relevant to students because the typical tutor cannot replicate such a great level of efficacy. It has been observed that few studies have had tutors who can show the level of learning described by Bloom (Graesser et al. 2). Since most experts in education have difficulties producing the levels described, it shows how uncommon it is for the typical tutor to meet those standards. Decades after Bloom's publication, VanLehn points out that only two studies support Bloom's results, and in his meta-analysis he concludes that human tutors have a smaller but still very significant effect than estimated in the challenge (14, 17-18). If it is such a problem to meet the levels reported by Bloom and only two different studies can support his claim, then his results are very limited. Furthermore, VanLehn used meta-analysis, which is more representative of the common case due to the number of studies reviewed. Moreover, current research suggests that human tutors don't correctly estimate what students don't understand, and many of their actions are not always related to better learning (Chi 3; Ohlsson et al. 2). By identifying which tutoring actions don’t improve learning outcomes, professionals can focus on the actions that improve learning and develop ITSs that can compete with human tutoring. Chi also notes that typical tutoring that produces learning outcomes results from replicable instructions (3). It is significant because if tutors frequently miss learning opportunities, and favorable learning outcomes come mostly from a standard set of rules, then experts can successfully automate many of these tasks through ITSs.

ITSs are as effective as human tutors in student learning. Given that ITSs were developed based on human tutors, it is reasonable to ask about the effectiveness of ITSs compared to human tutoring. VanLehn concludes that ITSs are as successful in expanding learning as one-on-one tutoring in science, technology, engineering and mathematics (18). In most cases, a student’s learning is the same if they receive tutoring from an ITS instead of a human. Another meta-analysis confirms there is no discernible difference in learning outcomes between ITSs, one-on-one tutoring, and small group tutoring of eight students or less (Ma et al. 7, 12). The importance is that a student’s mathematics learning will not diminish if they choose an ITS rather than a human tutor. Even a meta-analysis by Steenbergen-Hu and Cooper found that there might be a small to medium favorable effect from the use of ITSs relative to human tutoring, within the limited number of studies reviewed (13). The conclusion drawn from the evidence is that ITSs are at least as effective as the typical human tutor, and sometimes ITSs outperform the typical human tutor. The evidence has shown that ITSs have caught up with human tutors and can be as effective in the domain of mathematics.

Research shows that ITSs can benefit the mathematics outcomes of high school students. Since ITSs are as effective as one-on-one human tutoring in student learning, we proceed to evaluate the effectiveness of these systems regarding student outcomes in the classroom, relative to traditional classrooms without an ITS. The less optimistic of the spectrum results from the meta-analysis by Steenbergen-Hu and Cooper, who concluded that ITSs don't make a difference for K-12 students in mathematics relative to traditional classroom instruction, or the change is favorable but minimal (13). Even though this result is not very optimistic, it still sheds light. If ITSs don't hurt and may provide a small positive change in the outcomes of mathematics in high school students, it is still in our best interest to explore implementing ITSs. Nevertheless, Steenbergen-Hu and Cooper included systems that are neither classified as ITS by their developers nor by most experts (Kulik and Fletcher 30). These systems included might be less effective than ITSs, causing the results to be lower than the effects of ITSs.

More evidence supports the benefits of ITSs in the mathematics outcomes of high school students. It was found through another meta-analysis that ITSs have the most favorable effect when compared to other modes of instruction in high school mathematics, including conventional classroom instruction (Ma et al. 7, 9-12). Ma et al. also note that these results are significantly more favorable than Steenbergen-Hu and Cooper's analysis in K-12 mathematics (13). The analysis by Ma et al. shows that ITSs are producing the same outcomes of human tutoring, which are positive when compared to the outcomes of traditional instruction. This is consistent with the most recent data, meta-analysis shows in 46 out of 50 studies that ITSs can produce better outcomes than the traditional classroom, in the US and other countries (Kulik and Fletcher 28). The significance is that students in the US and other countries can improve in their classes by taking advantage of ITSs. Kulik and Fletcher observed a moderate-to-large favorable effect on the mathematics outcomes of high school students (30). Therefore, the outcomes of US high school students improve with the use of ITSs. Kulik and Fletcher mentioned that "Developers of ITSs long ago set out to improve on the success of CAI [Computer Assisted

Instruction] tutoring and to match the success of human tutoring" (28). We can then conclude that ITSs not only perform as good as human tutors, but they can provide a positive change in the mathematics outcomes of high school students in the US.

On the other hand, ITSs should be used to complement the classroom, not to replace it. Since ITSs are beneficial to the mathematics outcomes of high school students, it is tempting to make the mistake of attempting to replace classrooms and teachers with ITSs. It has been noted that computer technology is more favorable for students when used as a complement than when used for direct teaching (Tamim et al. 17). ITSs are part of computer technology; hence ITSs should be more effective when used as a support tool for the classroom. In fact, it has been shown that little and moderate technology usage is more favorable in terms of student performance than high technology usage, and the most benefits in students learning results when technology complements other methods (Schmid et al. in Steenbergen-Hu and Cooper 15). It is relevant because when technology replaces instruction, the amount of use increases significantly. Therefore, if moderate use of technology is more favorable for students, and technology is more suited to support instruction, it leads to thinking moderate use of ITS as a complement can be more beneficial than replacing the whole classroom. Koedinger, a doctor in Cognitive Psychology with more than 250 peer-reviewed publications ("Kenneth R. Koedinger"), mentioned in an interview that ITSs should aid instruction by allowing the instructor to perform more important tasks (Koedinger and Aleven, "An Interview Reflection" 8). Using an ITS allows the teacher enough time to focus on tasks that humans can do better. These more important activities can be the main component of instruction, guiding students, motivating discussion, and all the things humans are more suitable to perform, while an ITS complements instruction.

Teachers play an essential role in the mathematics classroom with ITSs. The benefits associated with ITSs are only present in the classroom when an experienced teacher is also present, making teachers necessary. For instance, when comparing the effectiveness of an ITS in high school mathematics with instructors that were new to ITS and instructors with experience, Koedinger and Anderson found that the classes with inexperienced teachers harmed students' outcomes while the experienced instructors produced a largely positive effect (in Kulik and Fletcher 27). This supports the point that teachers still have an important and central role in the mathematics high school classroom, even when an ITS is being used. Inexperienced teachers using ITSs can be worse than not using ITSs in the first place. ITSs should not completely replace teachers since they are still important. Opposite to experienced instructors, inexperienced instructors tended to use ITS as a replacement minimizing interaction with students, and the interactions that occurred between students and professors were distractive from the mathematics and concentrated in the technology (Koedinger and Anderson in Kulik and Fletcher 27). The meaning is that teachers still play an important position in student outcomes even after having implemented an ITS in the classroom. Only experienced instructors who understand how to implement ITSs in the classroom can harness the advantages of ITSs. ITS should only replace the parts of instruction where not much interaction is important, such as homework.

There is more evidence that also supports the importance of teachers in classrooms with ITSs. VanLehn pointed out in his meta-analysis that not even a single study that he analyzed replaced the whole classroom with ITSs (17). This is important because we have established that ITSs are as effective as one-on-one tutoring and can produce beneficial learning and outcomes for high school mathematics and even in other domains. However, it is rare for studies to replace traditional classrooms and teachers with an ITS. The reason we've seen such a big quantity of favorable results of ITSs can be in part because most studies implement ITSs as an aid to instruction, rather than a replacement for classroom instruction and teachers. It has been suggested that teachers must receive proper coaching for ITS to be effective in high school mathematics (Koedinger and Anderson 6). Accordingly, teachers must receive training because they still have a role even in classrooms that use ITS as a complement, the role is simply different. The role is different because teachers can now focus on the main components of instruction where a big amount of interaction is beneficial, such as direct instruction. ITS can then serve for other tasks that require less interaction, for example, homework and daily practice.

ITSs can be beneficial in student learning when used for homework assignments. Given that ITSs should support the classroom, it is important to explore one form in which an ITS can benefit students when used as a complement. It is essential to note that using ITSs for homework might seem like a small change. However, homework can represent a big portion of the students' learning (VanLehn 17). If homework is a significant part of learning, then ITSs can make a significant effect in the classroom when used for homework assignments. Researchers observed for five years that students who used a physics ITS for homework learned more than students who did conventional homework with paper and pencil, and also than students who used non-intelligent online homework services (VanLehn et al. 44). This is an example of a case where an ITS has been useful in the students' learning when used for homework assignments in an area where mathematics is fundamental.

In another four-year study, students scored higher in mathematics and statistics tests when they used an ITS for homework rather than conventional pen and pencil homework (Jonsdottir et al. 1, 7-8). Since tests try to measure student learning, we would expect that if students are doing better in tests when using ITSs for homework, they should also learn more.

Other two similar studies have found that students in different levels of mathematics who use an ITS to complete their homework assignments, learn more than students who turn in homework in the conventional form with paper and pencil (Mendicino et al. 13; Roschelle et al. 10). Therefore, research supports the claim that ITSs can be favorable in student learning and outcomes in mathematics when used in homework assignments. Certainly, Steenbergen-Hu and Cooper found ITSs beneficial to students with a small to moderate effect in the limited number of studies compared (13). Meta-analysis shows that ITSs can, in the typical case, benefit student learning and outcomes in the classroom. Another meta-analysis suggests that using ITSs for homework is associated with better learning (Ma et al. 12-13). It is important because it establishes that better learning and using ITSs for homework assignments occur simultaneously. Hence, ITSs can be beneficial to the classroom when used as a complement through homework assignments.

The biggest limitation of this paper is that it analyses the effectiveness of ITSs in the typical case. Of course, as it is the case with any computer program, no ITS program is just as effective as any other. Some studies and programs within the meta-analyses evaluated showed negative effects (Kulik and Fletcher 25; Ma et al. 7; Steenbergen-Hu and Cooper 7-8, VanLehn 22-25), and the meta-analyses excluded some studies for not meeting an inclusion criteria that each author designed to ensure a fair comparison (Kulik and Fletcher 9-10; Ma et al. 2, 6; Steenbergen-Hu and Cooper 2; VanLehn 8-9). But the positive effects outweigh by far the negative ones in all the studies reviewed, which leads the present work to assert that ITSs can be beneficial for students. Although it is true that most times that an ITS causes negative effects it can be due to a faulty implementation from an inexperienced instructor, one must take caution when one is evaluating whether to implement an ITS in a specific classroom. First, one must make sure that instructors receive appropriate training to implement an ITS into the classroom as a supporting tool. Second, ITS programs are unique, and each should be evaluated independently when deciding whether to implement such a program in a specific classroom, in a specific domain, and with a specific population of students. Therefore, one must not take this decision lightly. What the present work provides, is evidence that ITSs can be as effective as human tutors regarding the mathematics learning and outcomes of high school students in the US, and that ITSs should be used as a complement to the classroom. Furthermore, it provides one way that ITSs can be beneficial to aid instruction, which is for homework assignments.

As a consequence of the present work, the use of ITSs for homework assignments in the mathematics high school classroom may help minimize the existing gap between US students and other developed countries. Fascinating next areas of research can focus on domains other than mathematics or levels of studies other than high school students, the design of courses that educators can take to learn how to implement ITSs for support in the classroom, the design of guidelines that allow educators to choose which ITS is appropriate for their specific classroom's needs and goals, the development of a service that is updated regularly and shows the most recent evidence related to each individual ITS program, and the exploration of forms other than homework in which ITSs can support and benefit instruction.

Works Cited

Alkhatlan, Ali, and Jugal Kalita. "Intelligent Tutoring Systems: A Comprehensive Historical Survey with Recent Developments." *International Journal of Computer Applications*, vol. 181, no. 43, Mar. 2019, pp. 1-20, doi:10.5120/ijca2019918451. Accessed 7 May 2019.

Bagheri, Mehri M. "Intelligent and Adaptive Tutoring Systems: How to Integrate Learners." *International Journal of Education*, vol. 7, no. 2, Apr. 2015, pp. 1-16, doi:10.5296/ije.v7i2.7079. Accessed 6 May 2019.

Cheung, Mike W. *Meta-Analysis: A Structural Equation Modeling Approach*. John Wiley & Sons, 2015. *ProQuest*, ebookcentral-proquest-com.ezproxy.shoreline.edu/lib/shorelinecc-ebooks/detail.action?docID=1895808. Accessed 12 May 2019.

Chi, Michelene T. "Constructing Self-Explanations and Scaffolded Explanations in Tutoring." *Applied Cognitive Psychology*, S33-S49 ser., vol. 10, 1996, www.public.asu.edu/~mtchi/papers/selfex96.PDF. Accessed 11 May 2019.

Ferester, Bill. *Teaching Machines: Learning from the Intersection of Education and Technology*. John Hopkins UP, 2014. *EBSCOhost*, search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=778011&site=ehost-live. Accessed 7 May 2019.

Graesser, Arthur C., et al. "AutoTutor Improves Deep Learning of Computer Literacy: Is It the Dialog or the Talking Head?" *Artificial Intelligence in Education*, vol. 97, 2003, pp. 47-54. *EBSCO eBook Collection*, search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=110132&site=ehost-live. Accessed 11 May 2019.

Jonsdottir, Anna H., et al. "Difference in Learning among Students Doing Pen-and-Paper Homework Compared to Web-Based Homework in an Introductory Statistics Course." *Journal of Statistics Education*, vol. 25, no. 1, Apr. 2017, pp. 12-20. *Taylor & Francis*, doi:10.1080/10691898.2017.1291289. Accessed 7 May 2019.

"Kenneth R. Koedinger." *Carnegie Mellon University*, pact.cs.cmu.edu/koedinger.html. Accessed 30 May 2019.

Koedinger, Kenneth R., and Vincent Aleven. "Exploring the Assistance Dilemma in Experiments with Cognitive Tutors." *Educational Psychology Review*, vol. 19, no. 3, 21 July 2017, pp. 239-64. *Pittsburgh Advanced Cognitive Tutor Center*, doi:10.1007/s10648-007-9049-0. Accessed 16 May 2019.

---. "An Interview Reflection on 'Intelligent Tutoring Goes to School in the Big City.'" *International Journal of Artificial Intelligence in Education*, vol. 26, no. 1, Mar. 2016, pp. 13-24. *SpringerLink*, doi:10.1007/s40593-015-0082-8.

Koedinger, Kenneth R., and John R. Anderson. "Effective Use of Intelligent Software in High School Math Classrooms." *Proceedings of the Sixth World Conference on Artificial Intelligence in Education*, 1993, pp. 241-48. *Figshare*, figshare.com/articles/Effective_Use_of_Intelligent_Software_in_High_School_Math_Classrooms/6614585/1. Accessed 6 May 2019.

Kulik, James A., and J. D. Fletcher. "Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review." *Review of Educational Research*, vol. 86, no. 1, Mar. 2016, pp. 42-78, doi:10.3102/0034654315581420. Accessed 8 May 2019.

Li, Fei, et al. "Traditional Classroom vs E-learning in Higher Education: Difference between Students' Behavioral Engagement." *International Journal of Emerging Technologies in*

Learning, vol. 9, no. 2, 2014, pp. 48-51. *Online Journals*, doi:10.3991/ijet.v9i2.3268. Accessed 5 June 2019.

Ma, Wenting, et al. "Intelligent Tutoring Systems and Learning Outcomes: A Meta-Analysis." *Journal of Educational Psychology*, vol. 106, no. 4, Nov. 2014, pp. 901-18. *PsychArticles*, doi:10.1037/a0037123. Accessed 5 May 2019.

Mendicino, Michael, et al. "A Comparison of Traditional Homework to Computer-Supported Homework." *Journal of Research on Technology in Education*, vol. 41, no. 3, 2009, pp. 331-59. *ProQuest*, search.proquest.com/docview/274765745?accountid=1164. Accessed 7 May 2019.

Ohlsson, Stellan, et al. "Beyond the Code-and-count Analysis of Tutoring Dialogues." *Artificial Intelligence in Education*, 2007, pp. 349-56. *EBSCO eBook Collection*, search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=209234&site=ehost-live. Accessed 16 May 2019.

Poole, David, et al. "Computational Intelligence and Knowledge." Introduction. *Computational Intelligence: A Logical Approach*, Oxford UP, 1998. *Computer Science Department of the University of British Columbia*, www.cs.ubc.ca/~poole/ci/ch1.pdf. Accessed 12 May 2019.

"Prepare and Inspire: K-12 Science, Technology, Engineering, and Math (STEM) Education for America's Future." *The Education Digest*, vol. 76, no. 4, Dec. 2010, pp. 42-46. *ProQuest*, ezproxy.shoreline.edu/login?url=https://search-proquest-com.ezproxy.shoreline.edu/docview/819517160?accountid=1164. Accessed 12 May 2019.

Roschelle, Jeremy, et al. "Online Mathematics Homework Increases Student Achievement." *American Educational Research Association Open*, vol. 2, no. 4, 23 Oct. 2016, pp. 1-12. *SAGE Journals Online*, doi:10.1177/2332858416673968. Accessed 7 May 2019.

Steenbergen-Hu, Saiying, and Harris Cooper. "A Meta-Analysis of the Effectiveness of Intelligent Tutoring Systems on K–12 Students' Mathematical Learning." *Journal of Educational Psychology*, vol. 105, no. 4, Nov. 2013, pp. 970-87. *PsychArticles*, doi:10.1037/a0032447. Accessed 5 May 2019.

Tamim, Rana M., et al. "What Forty Years of Research Says about the Impact of Technology on Learning: A Second-Order Meta-Analysis and Validation Study." *Review of Educational Review*, vol. 81, no. 1, Mar. 2011, pp. 4-28. *JSTOR*, www.jstor.org/stable/23014286. Accessed 8 May 2019.

VanLehn, Kurt. "The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems." *Educational Psychologist*, vol. 46, no. 4, Dec. 2011, pp. 197-221. *Academic Search Complete*, doi:10.1080/00461520.2011.611369. Accessed 5 May 2019.

VanLehn, Kurt, et al. "The Andes Physics Tutoring System: Lessons Learned." *International Journal of Artificial Intelligence in Education*, vol. 15, no. 3, Aug. 2005, pp. 147-204. *Defense Technical Information Center*, apps.dtic.mil/dtic/tr/fulltext/u2/a521924.pdf. Accessed 8 May 2019.