Effectiveness of Intelligent Tutoring Systems in STEM

An intelligent tutoring system (ITS) is generally defined as a computer learning environment that helps students master knowledge by implementing artificial intelligent algorithms. Such algorithms are usually adapted to the individual student’s mastery level and unique learning types in presenting content, asking questions, and giving feedback [3]. More recently, ITS can now engage in conversational dialogues using natural language between an animated talking head agent and the human learner [4]. This advancement would allow for the next generation of ITS to mimic human speech, facial expressions, and even gestures. Keywords—intelligent tutoring systems,

Currently, according to the U.S. Education Department’s National Center for Education Statistics 50% of STEM undergraduate students are at risk of leaving the field before completing a college degree. Therefore, there is an urgency for technology-based learning assistants tailored to the particular needs of individual students in STEM majors. Computer tutoring systems that are deemed as “intelligent” enough to meet the standards of expert human tutors have become a reality[2]. The current study draws from the previous meta-analysis literature to conduct a systematic review of the learning theories, scaffolding methods, and effectiveness. The main goal of building an ITS is to simulate a human tutoring environment [5]. The basic architecture of these systems follows the principles from classical learning theories such as social and cognitive constructivism, where the students learn by discovering and assimilating new information into their existing knowledge structure and interacting with their peers

For over 40 years, researchers have been trying to develop ‘intelligent’ computer technology to improve the delivery of instruction to learners and help them learn faster and more efficiently. The general feeling among educators and researchers alike is that these efforts have not been successful, at least in terms of their widespread adoption in education.

Intelligent tutoring systems (ITS) differ from the previous generation of computer-aided instructional tools because they try to model the domain being taught and the student’s likely mastery of its content. An adaptive and completely personalized path through the content is dynamically and continually constructed by the system, based on continuous assessment and feedback.

A large number of studies suggest that students taught using ITS in well-defined domains, such as mathematics, can learn to the same level of mastery as those taught by traditional classroom teaching and even one-to-one human tutoring. In addition, the time to reach that level of competency is generally shorter when compared with more traditional methods.

Kulik and Fletcher [20] conducted a more rigorous analysis on ITS that delved deeper into types of experiments conducted, particularly the appropriate usage of control groups and assessment instruments used for evaluating ITS effectiveness as a learning aid. They found that overall, ITS had moderate to large effects on improving learning. However, the effectiveness varies in accordance to the following: (1) when the test instruments were locally constructed rather than using standardized test items, ITS had a substantially larger effect size in learning, (2) when a large sample size was used in the evaluations, the effect size shrinks, (3) when participants were either of lower grade levels or were unfamiliar with using the system, the effect sizes were smaller, but can improve substantially over a long period of usage, and (4) ITS had significantly less effect on learning outcomes when multiple choice tests were used to measure outcomes. In addition, Kulik and Fletcher [20] found that when the ITS Cognitive Tutor was used in the meta-analysis, the effect size decreased. This is likely due to the large-scale nature of the evaluation used in Cognitive Tutor, and standardized test items used almost exclusively for the system.

An important question to answer is whether or not ITSs are really effective in providing the learning outcomes they claim to obtain. There have been a number of meta-analysis efforts to investigate the effectiveness of ITSs. The following present a few recent such efforts with their findings to answer the question. Ameta-analysis was conducted by VanLehn in 2011 for the purpose of comparing effectiveness of computer tutoring, human tutoring and no tutoring [20]. In this analysis, computer tutors were characterized based on the granularity of the user interface interactions, including answer-based, step-based, and sub step-based tutoring systems. Their analysis included studies published between 1975 and 2010. 10 comparisons were presented from 28 evaluation studies. The study found that human tutoring raised test scores by an effect size of 0.79 compared to no tutoring; thus, it is not as effective as 2.0 found by Bloom earlier [2]. Moreover, it was found that step-based tutoring (0.76) was as almost effective as human tutoring whereas sub step-based tutoring was only 0.40 as effective compared to no tutoring. VanLehn’s findings suggest that tutoring researchers should focus on ways to improve computer tutoring to reach up to Bloom’s finding that human tutoring has 2.0 multiplicative effect compared to no tutoring. The meta-analysis conducted by Steenbergen and Cooper in 2013 analyzed the effectiveness of ITSs on k-12 students’ math learning [21]. This empirical research examined 26 reports comparing the effectiveness of ITSs with that of regular classroom instruction. Their finding was that ITSs did not have a significant effect on student learning outcomes when used for a short period. However, the effectiveness appeared to be greater when ITS was used for one full school year or longer. In addition, the effects appeared to be greater on general students than on low achievers. The meta-analysis by Ma et al. [22] was conducted in 2014 for the purpose of comparing the learning outcomes for those who learn by using ITSs and those who learn in non-ITS learning environments. Their goal was to verify the effect sizes of ITSs taking into account factors such as type of ITS, type of instruction (individual, small, large human instruction etc.), and subject domain (chemistry, physics, mathematics etc.), and other factors. Ma et al., analyzed 107 effect size findings from 73 separate studies. The ITS environment was associated with greater learning achievement compared to teacher-led and large group instruction with an effect size of 0.42, 0.57 for non-ITS computer based instruction, and 0.35 for text books or workbooks. On the other hand, there was no considerable difference between learning outcomes from ITSs and from individualized human tutoring (-0.11) or small group instruction (0.05). Ma et al., reported that ITSs achieved higher education outcome than other forms of instructions except for small group human tutoring. In addition, the ITS effect varied as features and characteristics of ITSs, student attributes, domain knowledge, and other factors varied. Finally, the meta-analysis produced by Kulik and Fletcher in 2015 [23] compared the learning effectiveness of ITSs with conventional classes from 50 studies. 92% of the studies indicated that students who interacted with ITSs outperformed those who received traditional class instructions. In 39 of the 50 studies, performance improvement gains were up to 0.66 median effect sizes, which is considered to be moderate to strong. However, the effect was weak for standardized tests as the effect size was 0.13. Because of the fact that there is no general agreement on the effectiveness of ITSs, questions come up for researchers to answer. How effective are ITSs really?, What are the critical reasons that affect learning in ITSs?, What possible changes can be made to improve ITSs? Intelligent Tutoring Systems

Reference

Mathews, M. (2006). Investigating the effectiveness of problem templates on learning in intelligent tutoring systems.

Krouska, A., Troussas, C., Giannakas, F., Sgouropoulou, C., & Voyiatzis, I. (2021). Enhancing the effectiveness of intelligent tutoring systems using adaptation and cognitive diagnosis modeling. In *Novelties in Intelligent Digital Systems* (pp. 40-45). IOS Press.