Estimating the Effect of Web-Based Homework

Kim Kelly[1], Neil Heffernan[1], Cristina Heffernan[1],
Susan Goldman[2], James Pellegrino[2], and Deena Soffer Goldstein[2]

[1] Worcester Polytechnic Institute
[2] University of Illinois –Chicago
{kkelly,nth}@wpi.edu

Abstract. Traditional studies of intelligent tutoring systems have focused on their use in the classroom. Few have explored the advantage of using ITS as a web-based homework (WBH) system, providing correctness-only feedback to students. A second underappreciated aspect of WBH is that teachers can use the data to more efficiently review homework. Universities across the world are employing these WBH systems but there are no known comparisons of this in K12. In this work we randomly assigned 63 thirteen and fourteen year olds to either a traditional homework condition (TH) involving practice without feedback or a WBH condition that added correctness feedback at the end of a problem and the ability to try again. All students used ASSISTments, an ITS, to do their homework but we ablated all of the intelligent tutoring aspects of hints, feedback messages and mastery learning as appropriate to the two practice conditions. We found that students learned reliably more in the web-based homework condition and with an effect size of 0.56. Additionally, teacher use of the homework data lead to a more robust and systematic review of the homework. Future work will further examine modifications to WBH to further improve learning from homework and the role of WBH in formative assessment.

Keywords: intelligent tutoring systems, immediate feedback, homework, effect size, formative assessment.

1 Introduction

Several studies have shown the effectiveness of intelligent tutoring systems when used in the classroom [7], [9], reporting effect sizes up to 0.78. The few studies that have explored the effectiveness of ITS when used as homework were very encouraging [9]. Yet, complex tutoring systems are not suited for nightly homework. Computer aided instruction (CAI), which gives all students the same questions with immediate end-of-question feedback is more applicable as teachers can easily create the content from textbooks or worksheets. Kulik and Kulik's [4] meta-analysis reviewed CAI and reported an effect size of 0.3 for simple computer based immediate feedback systems. However, these studies were not in the context of homework use and did not focus on how teachers use the data to respond to student performance.

Despite the relatively low effect sizes reported in Kulik and Kulik [4], web-based homework (WBH) holds promise for improving learning from homework by tailoring

K. Yacef et al. (Eds.): AIED 2013, LNAI 7926, pp. 824–827, 2013.

practice to individual performance. Doing so enables individuals to get corrective feedback so they can focus on areas where they are not successful. Shute [6] reviews the plethora of studies and theoretical frameworks developed around understanding the role of feedback for students as well as teachers. Black and William [1] have focused on formative assessments, with an eye on informing the teacher and giving feedback to students. The cognitive science literature suggests that letting students practice the wrong skill repeatedly on their homework is detrimental to learning. In this study we look to measure the effect on learning by comparing simple WBH to a traditional homework (TH) condition representing the type of practice that millions of students perform every night in America and probably around the world. Additionally, we explore how the teacher can use the data to modify and improve instruction.

2 Experimental Design

Participants were 63 seventh grade students, who completed the activities included in the study as part of their regular math class and homework. Students were assigned to conditions by blocking on prior knowledge. All students were given a pre-test and lesson on negative exponents. That night, students completed their homework using ASSISTments. The assignment was designed in triplets, with three morphologically similar questions in a row. Additional challenge questions were included to maintain ecological validity.

Students in the WBH condition were given correctness-only feedback at the end of the problem. If a student answered a question incorrectly, he/she was given unlimited opportunities to self-correct, or he/she could press the "show me the last hint" button to be given the answer. It is important to emphasize that this button did **not** provide a hint; instead it provided the correct response, which was required to proceed to the next question. Students in the TH condition used ASSISTments in "test mode" to simulate traditional homework practice without any feedback.

The following day all students took PostTest1 and then participated in the homework review process. Students in the WBH condition left the room and completed an unrelated assignment. Students in the TH condition reviewed their homework in a very prevalent and traditional fashion. They were given the answers to the homework, time to check their work, and the opportunity to ask questions. The groups of students switched and the teacher used the item report, generated by ASSISTments to review the homework with students in the WBH condition. Common wrong answers and obvious misconceptions guided the discussion. The next day, all students took PostTest2. All of the study materials, data and videos are available in Kelly [3].

3 Results

Several scores were derived from the data collected by the ASSISTments system. Student's HW Average was calculated based on the number of questions answered correctly on the first attempt divided by the total number of questions on the assignment (20). Partial Credit HW Score was calculated by dividing the number of

questions answered without being given the answer by the number of total questions on the homework assignment (20). Time Spent was calculated using the problem log data generated in ASSISTments and is reported in minutes. Times per action are truncated at five minutes.

Learning Gains from Homework: An ANCOVA showed that students in the WBH condition reliably outperformed those in the TH condition on both PostTest1 ($F(1,60)=4.14$, $p=0.046$) and PostTest2 ($F(1,60)=5.92$, $p=0.018$) when controlling for pre-test score. See Table 1 for means and standard deviations. If the difference was reliable we computed a Hedge corrected effect size [2]. The effect sizes do not take into account pretest. The key result for posttest2 of 0.56 effect size had a confidence interval of between 0.07 and 1.08. Unexpectedly, correctness-only feedback was found to be time efficient. Students in both conditions spent the same amount of time to complete their homework ($F(1,60)=0.002$, $p=0.96$).

Table 1. Means, standard deviations (in parenthesis), and effect size for each measure by condition. *Notes a reliable difference.

	TH	WBH	*p*-value	Effect Size
Pre-Test	9% (17)	7% (14)	0.78	NA
PostTest1	58% (27)	69% (21)	0.046*	0.52
PostTest2	68% (26)	81% (22)	0.018*	0.56
HW Average	61% (20)	60% (15)	0.95	NA
Partial Credit HW Score	61% (20)	81% (18)	0.0001*	1.04
Time Spent (mins)	22.7 (9.6)	23.2(6.2)	0.96	NA

Learning Gains from Homework Review: To address the second research question of the effectiveness of using the data to support homework review, a paired t-test revealed that students in both conditions did reliably better on PostTest2 than on PostTest1 ($t(62)=3.87$, $p<0.0001$). However, an ANCOVA revealed that when accounting for PostTest1 scores, there is not a reliable difference by condition in the gains from PostTest1 to PostTest2 ($F(1,60)=2.18$, $p=0.15$). This suggests that both methods of reviewing the homework lead to substantially improved learning.

Observational Results: In addition to examining the effects of immediate feedback on learning, this study explored the qualitative changes to the homework review process the following day in class. An observational analysis of the video recordings of the teacher reviewing the homework revealed that while the time spent in the WBH condition was often longer than the TH, it was also far more focused than the TH. Specifically, when students were in the TH condition, on average two minutes passed before any meaningful discussion took place. Whereas, when students were in the WBH condition, homework review began immediately with the teacher reviewing what she perceived to be the most important learning opportunities. Additionally, students in the TH condition reviewed fewer questions than the WBH condition and they tended to ask the same types of questions or even the same exact question that

was already reviewed. In the WBH condition, the teacher was able to ensure that a variety of question types and mistakes were addressed.

4 Contributions and Future Work

In this fast-paced educational world, it is important to ensure that time spent in class and on homework is as beneficial as possible. The randomized-controlled study presented here provides some strong evidence that web-based homework systems that provide correctness-only feedback and data to teachers are useful tools to improve learning on homework without additional time, suggesting a new use for ITS.

Acknowledgment. The authors would like to acknowledge support from the Bill and Melinda Foundation via EDUCAUSE as well as IES grants R305C100024 and R305A120125.

References

1. Black, P., Wiliam, D.: Inside the black box: Raising standards through classroom assessment. Granada Learning (2006)
2. CEM (2013), http://www.cemcentre.org/evidence-based-education/effect-size-calculator (accessed January 28, 2013)
3. Kelly, K.: (2012), http://www.webcitation.org/6E03PhjrP contains the materials Browse, http://web.cs.wpi.edu/~nth/PublicScienceArchive/Kelly.htm to view the materials
4. Kulik, C.C., Kulik, J.A.: Effectiveness of computer-based instruction: An updated analysis. Computers in Human Behavior 7, 75–94 (1991)
5. Rochelle: The IES Grant (2013), http://ies.ed.gov/funding/grantsearch/details.asp?ID=1273 (accessed January 28, 2013)
6. Shute, V.: Focus on Formative Feedback. Review of Educational Research 78(1), 153–189 (2008), http://www.ets.org/Media/Research/pdf/RR-07-11.pdf
7. Singh, R., Saleem, M., Pradhan, P., Heffernan, C., Heffernan, N.T., Razzaq, L., Dailey, M.D., O'Connor, C., Mulcahy, C.: Feedback during Web-Based Homework: The Role of Hints. In: Biswas, G., Bull, S., Kay, J., Mitrovic, A. (eds.) AIED 2011. LNCS, vol. 6738, pp. 328–336. Springer, Heidelberg (2011)
8. VanLehn, K.: The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist 46(4), 197–221 (2011)
9. VanLehn, K., Lynch, C., Schulze, K., Shapiro, J.A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., Wintersgill, M.: The Andes physics tutoring system: Lessons Learned. International Journal of Artificial Intelligence and Education 15(3), 1–47 (2005)