

Intelligent Tutoring for a College Course

MASON NEWMAN, Colorado State University, USA

Intelligent tutoring is a process by which a user learns material for a course and are given personalized feedback tools to enhance their learning. The intelligent tutoring functions on a number of psychological concepts. The high level format explored is a check in tool provided to students after each unit to examine their level of learning. The students learning should examine previous content up to what they have just learned. This tool should provide better learning for the course and for long term learning as a whole.

CCS Concepts: • **Human-centered computing** → **Web-based interaction**; **HCI theory, concepts and models**.

Additional Key Words and Phrases: education, HCI, web-interfaces, deep learning

ACM Reference Format:

Mason Newman. 2021. Intelligent Tutoring for a College Course. *ACM Trans. Graph.* 37, 4, Article 111 (August 2021), 3 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 ABSTRACT

School curriculums are developed to generate deep learning in students [Zuo and Wang 2021]. Ebbinghaus theorized there exists a forgetting curve that states you serially forget items you have learned over time [Murre and Dros 2015]. That’s why in a school curriculum, you may have cumulative exams that force you to relearn content every so often so you can remember it better. That also may explain why a large portion of what you learn in school is forgotten once you graduate. Unless we actively use information, we lose it. Knowing this, we can develop a piece of software that like a school curriculum actively reminds and quizzes us over content we want to remember and don’t actively use better. My project seeks to imitate an intelligent tutor, which theoretically takes information from a student, a curriculum, and adapts it based on the student’s current knowledge and the user interface [Bradáč and Kostolányová 2017]. My project aims to implement psychology techniques of learning into a computer program that can be used to aid learning outside of school in meaningful ways. The major question we’re trying to ask is how can we use basic classroom input devices and computers to increase students long-term memory.

Author’s address: Mason Newman, mrnewman@colostate.edu, Colorado State University, 711 Oval Drive, Fort Collins, Colorado, USA, 80521.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Association for Computing Machinery.

0730-0301/2021/8-ART111 \$15.00

<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

2 INTRODUCTION

2.1 Intelligent Tutoring

Intelligent tutoring dictates two types of feedback for learners: inner loop and outer loop feedback [Tacoma et al. 2021]. Inner loop feedback refers to feedback on an individual task or concept [Tacoma et al. 2021]. This may be feedback on an individual term from a unit. Outer loop feedback refers to higher level study materials and tasks to do to enhance learning [Tacoma et al. 2021]. Essentially, inner loop determines individual areas where students could improve, and outer loop aggregates all areas into a comprehensive tool. However, the implementation of these tools largely depends on the characteristics of the individual learning. Therefore, intelligent tutoring by necessity has to adapt to students needs and should start with a strong understanding of what kind of student is being taught.

2.2 Inner Loop: Digital Learning Tools

Tools like Quizlet have existed for many years at this point. Prior research has observed a positive correlation between interaction with digital flashcards and performance [Yüksel et al. 2020]. Therefore, we seek to test a peripheral feature of Quizlet tools, testing and feedback. Our experiment differs however, as it includes greater feedback than indicating incorrect or correct answers. Quizlet is limited by the amount of data a user puts into it. Therefore to represent the inner loop of intelligent tutoring, our tool will provide content from the course proactively for participants. The inner loop deals with individual cases of feedback for users. Therefore, our experiment will provide a short-term analysis on learning for a single section of a curriculum.

3 MOTIVATION

One of the major themes of psychology courses at CSU is childhood development. There are many theories about how we can best aid children to become better adults in the world. One of the best predictors of occupational, socioeconomic, and income success is education [Becker et al. 2019]. In fact, education better predicts adult outcomes than family background and gender [Becker et al. 2019]. Therefore, this research asks how can education be made more available to children who may have limited access? Additionally, how can we utilize modern technology to aid education so that during critical learning periods, children learn more and remember more of what’s taught to them? Personally, I have a deep interest in psychology and childhood development, so I’d like to create a project that can utilize the psychology methods I’ve learned about learning. However, it isn’t feasible nor ethical to test on children for this project. So instead, my goal is to establish a fundamental concept of intelligent tutoring. The testing effect states that we learn better by making and taking tests than by rereading or restudying our notes [Su et al. 2021]. I would like to establish through the testing effect, are previously missed questions more important to go over than previous content, and does reviewing previous content influence learning as measured by a final exam. Specifically, if we

missed content from a certain unit, does reviewing what we missed further our learning.

4 METHOD

I developed a web-based tool that pairs with a curriculum. My project utilized the Harvard CS50 free online class. Participants were instructed to first state some demographic information; this information will include gender, age, and technical skill (1-10 Likert scale). Then, participants read an introduction paragraph that highlights key terms that you would likely see at the beginning of this course. Participants then took 2 tests. The first and last test is the same for all participants. The second test was then be one of two conditions. The assisted feedback condition instructed users to review questions they got wrong and explained the reason the correct answer was correct. The control condition will simply take two separate quizzes. Unfortunately, other factors interfere with exam performance. We believe the results may be influenced by some extraneous variables. Some examples are cramming right before a test to improve short term memory (not long term memory), stress, sleep, etc. We will discuss limitations in greater detail in the discussion.

4.1 Prototype

The focus of my prototype is on the inner loop feedback of intelligent tutoring. The prototype is a web based tool similar to Quizlet which allows users to read content, take a quiz, and record results. Users are given immediate feedback on their performance and participants in the assisted feedback condition are given a rationale for questions they answered incorrectly. This is to simulate an adjusted curriculum provided for students in an intelligent tutoring program. The importance of immediate corrective feedback will be discussed later. Prior research also suggests that use of digital flashcards significantly increases learning [Yüksel et al. 2020]. Therefore, this experiment seeks to test if performative tests similar to ones you can take on Quizlet are beneficial to learning. In order to get a better understanding of why intelligent tutoring is effective, it's important to understand

For this specific quiz, I used a partial passage from one of Harvard CS50 introductory courses. This means that this experiment will be aimed at people who are less technically savvy as more experienced CS majors will likely know the answers before reading the passage. I created 12 questions based off the passage and set the correct answer based on a random number generation. Additionally, for first quiz, I randomly picked the questions by a random number generator. I generated two sets of HTML code which took users through a web based program for the experiment. The JavaScript for the assisted feedback condition was slightly modified to include constructive feedback for incorrect questions for the first part. Users are not allowed to navigate back to the initial paper. The only reference and tool for learning is the feedback given on which questions are correct or not.

4.2 Participants

Participants were 10 adult family members and friends who I invited in my home to participate in this experiment. As this experiment was

aimed at less technically savvy people, I refrained from including any CS majors. Additionally, the content of the quizzes were beginner CS material, which may lead to skewed results due to prior knowledge. There were 5 male and 5 female participants. The average age of the participants was 33.8 years old. 5 participants were randomly selected for the control group and 5 participants were randomly selected for the assisted feedback condition. Participants did not complete the experiment in exchange for anything.

4.3 Evaluation

Participants were be split into two groups, assisted retest and normal retest. repeated standard testing, repeated content testing and repeated intelligent testing. Participants took a survey before the experiment for information on age, gender, and technical savvy. Participants were then instructed to read a short passage taken from one of the the Harvard CS50 pages. Participants then navigated to the first quiz when they were ready. Upon completion, both groups were given feedback for how many and which questions they got incorrect. However, participants in the assisted feedback condition were instructed to go back and read a short rationale for why the correct answer was correct. Then both groups completed the second quiz and their performance on both quizzes was recorded. Participants were then debriefed on the purpose of the experiment and which condition they were apart of verbally. Participants were allowed to leave at any time and completion was completely voluntary.

The experiment aims to test the short-term effect of immediate corrective feedback on performance. Some prior research shows spaced learning is more effective, but less enjoyable for those learning [Su et al. 2021]. Therefore, this experiment represents the most convenient, but not necessarily the most optimal way to learn.

5 RESULTS

A two-way repeated measures ANOVA with replication was conducted to compare test performance for participants in the control group and the assisted feedback condition. Overall, participants in the assisted feedback condition answered more questions correctly ($M = 3.8$, $SD = 0.84$) than participants in the control group ($M = 3.1$, $SD = 1.66$). Our two-way ANOVA revealed that there was not a statistically significant interaction between the conditions $F(3,16) = 1.00$, $p = .33$. Follow up t-tests indicated that there was a significant difference between the assisted feedback condition and the control condition for quiz 2 $t(8) = 1.90$, $p = .05$ and a significant difference between performance for question 1 and 2 for the control condition. However, there was not a significant difference in performance between average test performance for the assisted feedback condition and the control group $t(8) = 1.18$, $p = 0.13$, nor was there a significant difference in performance between quiz 1 and quiz 2 for the assisted feedback condition.

6 DISCUSSION

Although the initial two-way ANOVA didn't find any significant results, the follow up t-tests showed some significance. Although performance didn't significantly increase in the immediate feedback condition, it appears that performance significantly dropped in the control group. Therefore, feedback may not enhance learning, but

Table 1. Two-Way Repeated Measures ANOVA

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Sample	2.45	1	2.45	1.96	0.18	4.49
Columns	1.25	1	1.25	1	0.33	4.49
Interaction	1.25	1	1.25	1	0.33	4.49
Within	20	16	1.25			
Total	24.95	19				

help reduce the natural forgetting effect. However, this experiment cannot be used to determine long term effects on forgetting due to the testing window. The main focus of this experiment was on the effect of immediate corrective feedback. What we did not include in this study was a measure of user engagement. Prior research has shown that interaction with feedback is positively correlated to performance and getting the answer correct on the first try is negatively correlated with performance [Chen et al. 2018]. This result reflects that we are over confident learners, and our best learning comes when we review content we think we know well and content we don't know well. We also encouraged users to look at the feedback and engage with the rationale, but we have no measure of how much they used it. We could have used an introspective questionnaire to determine engagement, but we gaze measurement would be better. Eye fixation is commonly used to measure interest, desire, and likeness, so future research may utilize eye tracking technology to better understand how much users interact with feedback.

6.1 Implications for Intelligent Tutoring

If participants don't actively engage in their feedback, they won't see any benefit to intelligent tutoring. Computers can suggest feedback and study directions, but ultimately it's the users choice if they want to engage in the learning. Therefore, the next step for intelligent tutoring may be engaging users in meaningful ways. One large benefit of in person teaching is professor interaction. Prior research showed that depending on what kind of interaction students have with faculty and professors, they develop a positive or a negative opinion on the current racial diversity [Parker III and Trolan 2020]. This study reflects how important teacher student relationships are to perceptions for school and even desire to learn.

7 FUTURE DIRECTIONS FOR RESEARCH

We suggest future research should emphasize a longer period of learning and testing to better understand how feedback impacts learning and performance. Additionally, future research may utilize tools such as eye tracking to identify patterns of engagement in feedback. Additional research may include childhood development tools. One direction for early learning may be creating feedback through digital avatars if learning in a mixed environment. In the peak of covid-19, many kids became digital learners. Going forward, we should try to better understand how to teach digital learners as effectively as in person learners. However, as acknowledged in

prior research, this is a difficult, time consuming task to test, as the experience is highly individualized and the results depend on the user just as much as the technology [Tacoma et al. 2021]. That's why this will be the next step in education.

REFERENCES

- Michael Becker, Jürgen Baumert, Julia Tetzner, Kai Maaz, and Olaf Köller. 2019. Childhood intelligence, family background, and gender as drivers of socioeconomic success: The mediating role of education. *Developmental Psychology* 55, 10 (2019), 2231.
- Vladimír Bradáč and Kateřina Kostolányová. 2017. Intelligent tutoring systems. *Journal of Intelligent Systems* 26, 4 (2017), 717–727.
- Xin Chen, Lori Breslow, and Jennifer DeBoer. 2018. Analyzing productive learning behaviors for students using immediate corrective feedback in a blended learning environment. *Computers & Education* 117 (2018), 59–74.
- Jaap MJ Murre and Joeri Dros. 2015. Replication and analysis of Ebbinghaus' forgetting curve. *PLoS one* 10, 7 (2015), e0120644.
- Eugene T Parker III and Teniell L Trolan. 2020. Student perceptions of the climate for diversity: The role of student–faculty interactions. *Journal of Diversity in Higher Education* 13, 4 (2020), 333.
- Ningxin Su, Zachary L Buchin, and Neil W Mulligan. 2021. Levels of retrieval and the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 47, 4 (2021), 652.
- Sietske Tacoma, Paul Drijvers, and Johan Jeuring. 2021. Combined inner and outer loop feedback in an intelligent tutoring system for statistics in higher education. *Journal of Computer Assisted Learning* 37, 2 (2021), 319–332.
- H Gülrü Yüksel, H Güldem Mercanoğlu, and M Betül Yılmaz. 2020. Digital flashcards vs. wordlists for learning technical vocabulary. *Computer Assisted Language Learning* (2020), 1–17.
- Mei Zuo and Jixiang Wang. 2021. Higher Education Curriculum Evaluation Method Based on Deep Learning Model. *Computational Intelligence and Neuroscience* 2021 (2021).