Project Title: Evaluation of Generative AI Versus Human Feedback on Student Writing Performance

By Pranavsai Gandikota

Project Objective or Aim: The project investigates how generative AI feedback compares to human feedback in terms of improving a college student's writing skills throughout a semester. This study investigates which feedback type, AI compared to human, gives better writing results, how the students satisfaction and confidence in their writing ability is affected, and how many of the suggestions human vs AI, the students are more likely to implement.

Project Background and Significance: Recently people have begun utilizing AI to help with writing feedback. Steiss et al. (2024) and Wan & Chen (2023) in their studies showed how AI feedback has even started to match human quality in terms of clarity and grammar. However, past studies conducted on the performance of human and AI feedback didn't really investigate how students perceive feedback in terms of how confident they are, how satisfied and how much they feel they have progressed. This research aims to fill in the gap guided by the self-regulated learning theory by Zimmerman's (2002) and formative assessment models by Black & Wiliam (1998) to investigate not just examining a student's performance, but their overall opinion for their writing development. There will also be data collected as to whether the feedback is utilized, or not, which previous research has not yet been investigated.

Previous research had shown the pros and cons of AI generated feedback. Bouziane and Bouziane (2024) found GPT to struggle in consistency. This lack in reliability was also backed by Wetzler et al. (2024). Lin (2025) had noted that adult learners liked the speed of GPTs responses which they found really useful, but given feedback was not so useful on complicated tasks. But lacking in most of the past research was student perception and amount of feedback utilization which is planned to be a core part of this.

Research Methods: There will be recruiting of around 30 undergraduate students at UCF to participate in this study in a writing task taking place over a period of one semester. The students will be asked to submit an essay of 500-700 words each and will be given two types of feedback, one group given by ChatGPT 4 and the other given by a human writing tutor. There will be a total of 3 writing assignments each student will take part in.

First there will be a pre-study survey which investigates the confidence of the students and how their past experience with feedback was like. Then each group will be given the same writing assignment and will be asked to submit with feedback from either AI or a human, but this will not be disclosed to the participant. After this they will be asked to complete a survey regarding their satisfaction and feel regarding their overall writing development. This will be repeated 3 times with 3 different writing assignments to study the results. The writing papers will be graded by evaluators who are not disclosed to if AI or a human helped them.

Timeline (Fall 2025)

July – August(mid)

- Finalize study, submit IRB approval, create surveys and writing prompts, recruit graders and get participant recruitment started.

Aug 18–31

- Confirm participant recruitment, collect consent, begin pre-survey and split participants into groups, and hand out Writing Assignment 1.

Sept 1–21

- Collect and review Assignment 1, provide feedback, collect post feedback survey, students submit revised papers sent for final grading.

Sept 22–Oct 12

- Repeat same cycle for Assignment 2

Oct 13–Nov 2

- Same cycle for Assignment 3

 Nov 3–16

- Send post study survey and begin analyzing data from surveys and grading

Nov 17–Dec 6

- Complete analyzing the data, write the research article, edit and finalize the research poster.

Finals Week (till mid Dec)

- Submit article to Undergraduate Research Journal

This gives qualitative and quantitative data from the study with performance and rubric grades showing the effectiveness and the student opinion surveys showing how well the effects have been. A connection between these two will be drawn as well.

Expected Outcome: On completion of the study, a comparison and connection of all data will be made and presented in an article that compares AI and human feedback on student writing in terms of qualitative and quantitative data. This article after proper revision and edits is planned to be submitted to the undergraduate research journal. At the end of the project, it is expected that there will be a completed analysis article on the above-mentioned topic and there will also be a poster that visualizes the data collected for presentation purposes.

It is hoped to achieve a better understanding of the help of AI feedback for growth in writing. More knowledge on the topic as to whether using AI in contrast to human feedback has any benefits and finally, whether the students' opinion on their writing has been affected positively or not. From previous research trends, it is expected that the project findings show the quick feedback, grammar, spelling, structure and relevance from AI to be more beneficial, but the human feedback to be preferred in terms of context related issues, regarding consistency, understanding and deeper analysis of papers.

This research would help instructors know the best way to implement the use of AI along with their teaching for best results. It would help us understand where AI places its role in education (especially related to writing) and would help in reducing the workload of teachers who can use AI to help their students rather harm in certain areas where applicable. It would also be useful in research related to human computer interaction, and AI training to find areas of improvement to build more efficient models that would be able to account for emotion in writing and maintenance of overall themes as well, which is a major part of what is still in the progress of building in current AI models.

Literature Review

Black, Paul, and Dylan Wiliam. "Inside the Black Box: Raising Standards Through Classroom Assessment." Phi Delta Kappan, vol. 80, no. 2, 1998, pp. 139–148.

Bouziane, Karima, and Abdelmounim Bouziane. "AI versus Human Effectiveness in Essay Evaluation." Discover Education, vol. 3, no. 1, 2024.

Escalante, Juan, Austin Pack, and Alex Barrett. "AI-Generated Feedback on Writing: Insights into Efficacy and ENL Student Preference." International Journal of Educational Technology in Higher Education, vol. 20, no. 1, 2023.

Guo, Kai, et al. "Effects of an AI-Supported Approach to Peer Feedback on University EFL Students' Feedback Quality and Writing Ability." The Internet and Higher Education, vol. 63, 2024, Article 100962

Lin, Xi. "A Preliminary Study Investigating Adult Learners' Perspectives of Using ChatGPT to Seek Effective Assignment Feedback." Australian Educational Researcher, 2025.

Steiss, J., et al. "Comparing the Quality of Human and ChatGPT Feedback of Students' Writing." Learning and Instruction, vol. 91, 2024, Article 101894.

Wetzler, Elizabeth L., et al. "Grading the Graders: Comparing Generative AI and Human Assessment in Essay Evaluation." Teaching of Psychology, vol. 52, no. 3, 2024, pp. 298–304.

Zimmerman, Barry J. "Becoming a Self-Regulated Learner: An Overview." Theory into Practice, vol. 41, no. 2, Spring 2002, pp. 64–70.

Preliminary Work and Experience: I am a computer science major who has past experience especially in AI related research. I have worked as a research assistant in the VR and HCI lab for over a year now and have conducted user studies, have written related works for a paper and have assisted in making new AI related technologies related to LLMs. I have previously done English composition and am also currently enrolled in the technical writing course where I have had practice on writing assignments helping me in my general writing ability. I have also had past experiences in peer reviews and in feedback analysis from my last English composition research paper. All this will be helpful to me in creating surveys for my writing assignments, and for understanding different writing styles and student feedback. I have previously conducted a user study as well about gamification and student opinions on it being used in the educational field where I got experience as regards to going through surveys to find trends and conclusions as well, which I believe will be fruitful for this project.

IRB/IACUC statement: IRB approval will be required as human participants will be involved with the use of survey and the writing assessments. A provided consent form will be given and data will be collected anonymously with participants given the

Budget: The required budget will be needed for the following:

To pay the student participants as compensation for their participation through gift cards ($15*30) – $450

Professional essay grading services (for human feedback): $10 * 3 essays * 30 participants = $900

Miscellaneous/printing/presentation/advertising for study expenses: $60

TOTAL REQUESTED: 450+60+900 = **$1410**

(Extra required amounts will come from my side if funds are not sufficient.)