

CS6460: Final Paper

A Study on Student Engagement and Learning Outcomes in Online Learning

Cleo Zhang

yzhang3761@gatech.edu

Abstract—This paper documents the research investigating students' engagement and corresponding learning outcomes in an online learning context. Employing controlled experiments with 18 participants, the study applies quantitative and qualitative methodologies to conclude the participant feedback. Utilizing visual aids like combined area and line charts and Word Clouds aids in comparative analysis. The discussions highlight challenges in evaluating diverse language learning curves and acknowledge limitations in data scope and potential biases. Future work suggestions encompass interventions' impact, implicit assessment of learning outcomes, and the effects of repetitive behaviours on learning, aiming to refine educational technology's efficacy and pedagogical approaches.

1 INTRODUCTION

My journey into educational technology began during my undergraduate studies when I became involved in the Tapestry Tool project. Over time, this project has evolved into an open-source online learning platform that has captured my fascination. What intrigued me most was its innovative graph-based structure, which simplifies the concept of content relevance through edges and nodes, transforming traditional teaching into a multidimensional and engaging experience. This unique approach harnesses the human memory's capacity for making correlations, thus enhancing the overall quality of Education. This experience has fueled my desire to explore the relationship between human cognitive-level instincts and learning and the possible role of Educational Technology within this sector.

As a part of the CS6460 Research track, I am undertaking a research endeavour to investigate the pivotal role of educational software in fostering student

engagement within online learning. My research investigates how technology can effectively stimulate students' curiosity and intrinsic motivation, ultimately leading to increased engagement and improved learning outcomes to what extent. To achieve this, my primary focus is graduate students, chosen due to their practical feasibility for surveys and data collection and leveraging Duolingo as the tool to conduct my research.

This report primarily covers the progress I have made in my research this semester and how I have achieved these results. Also, I discuss the limitations of this research and look at interesting future research directions at the end of the report. I also want to thank the researchers in the Education Field whose previous work inspired me and made my research feasible.

2 METHODOLOGY

2.1 Research Design

I conducted controlled experiments for my research. I have 18 participants, randomly divided into two groups - one study the new language traditionally, and the other will be leveraging Duolingo. As highlighted by Luis von Ahn in his interview with Lindsay (Luis & Lindsay, 2023), the traditional approach to language learning, which they initially attempted with the first version of Duolingo, was perceived as dull and challenging to sustain motivation. Consequently, they transformed Duolingo into a levels-based gamified platform rather than a robotic learning tool. This aligns seamlessly with my educational philosophy and research interests. Furthermore, the fact that Duolingo is a free application eliminates the need for a budget to conduct my research.

Using quantitative and qualitative analysis methods, I aim to examine the effects of interest and self-motivation on Student Engagement and Learning Outcomes in a remote learning context. The following are what my participants need to do to engage with my research:

- Commit 30 minutes of learning daily.
- Send me weekly feedback about their learning state, answering the questions of "Did you spend half an hour a day studying last week?", "Did you find it interesting to learn a new language?", "Did you feel confident about what you will learn after two months?" and so on.

- By December 1st, 2023, let me know their overall feeling about the learning experience.

During this process, I tracked each week whether participants sent me weekly feedback and tabulated the data to send to the TAs to calculate these participants' participation scores for CS6460. The following two sections describe how the data are collected and analyzed.

2.2 Data Collection

There are two types of data that I mainly collect: Participants' weekly feedback and their final thoughts on the whole learning experience. I emailed participants a daily study reminder for the first two weeks, starting in week 8, to help them get into the swing of remote learning, and then on Fridays, I reminded them to reflect on the previous week and send me the feedback.

Due to the infeasibility of online language testing and internet/software access limitations, I do not directly use the participants' final Duolingo Levels and Badges to assess their learning outcomes. Instead, I derive the participants' learning outcomes from their level of engagement from the available research outcomes. This method of assessment applies to both the experimental and control groups.

2.3 Data Analysis

I aim to comprehensively analyze the data by combining quantitative and qualitative approaches, offering a multifaceted understanding of student engagement and participants' learning experiences in remote learning, followed by the learning outcomes inferred from extant research.

2.3.1 Quantitative Analysis

I quantitatively evaluated the engagement levels of both the experimental and control groups by examining the participants' weekly feedback. This encompasses tracking how many participants in each group reduced the frequency of their learning or discontinued learning in any given week.

I first organized the raw data and analyzed the learning engagement using the combined Double Line and Area Chart (*Figure 1*) and a Double Line Chart (*Figure 2*), which effectively displays two sets of data trends simultaneously, allowing

direct comparison of contrast, patterns, or correlations between the control and experimental groups over the shared axis.

Two Word Clouds (*Figure 3*) are applied to analyze participants' reflections on the learning experience, presenting a visual summary by emphasizing frequently used words and revealing dominant themes or sentiments in participants' feedback. Comparing control and experimental groups' clouds helps identify commonalities or differences in perceptions, enabling quick identification of critical focus aspects within their respective experiences.

2.3.2 Qualitative Analysis

I leveraged Qualitative Analysis to assess the participants' learning outcomes over the months, which included reading the feedback submitted by the participants in the final week and making logical inferences from the existing research on the relationship between learning engagement and learning outcomes.

3 RESULTS

Three types of charts are selected to present the data and deliver analysis results:

- Line Chart - great for showing trends over time. It can be helpful to illustrate the change in student engagement over weeks.
- Area Chart - emphasizes not only the trend but also the magnitude or volume of change in data, making it useful for showcasing cumulative totals.
- Word Cloud - visual representations where words are sized based on frequency. They can be used to analyze qualitative data, such as student feedback or responses related to the learning process.

3.1 Student Engagement

As shown in *Figure 1*, the solid Area represents the cumulative weekly feedback received on time by the experimental and control groups; the lines represent the trend of the total feedback received per week. It is clear from this that the experimental group sent more feedback each week than the control group, which means that the participants who used Duolingo to learn the new language were more active and engaged in the learning process than those who used the traditional method. *Figure 2* also corroborates this finding from a different perspective

- the total number of words fed back by participants in the experimental group was also higher than in the control group each week.

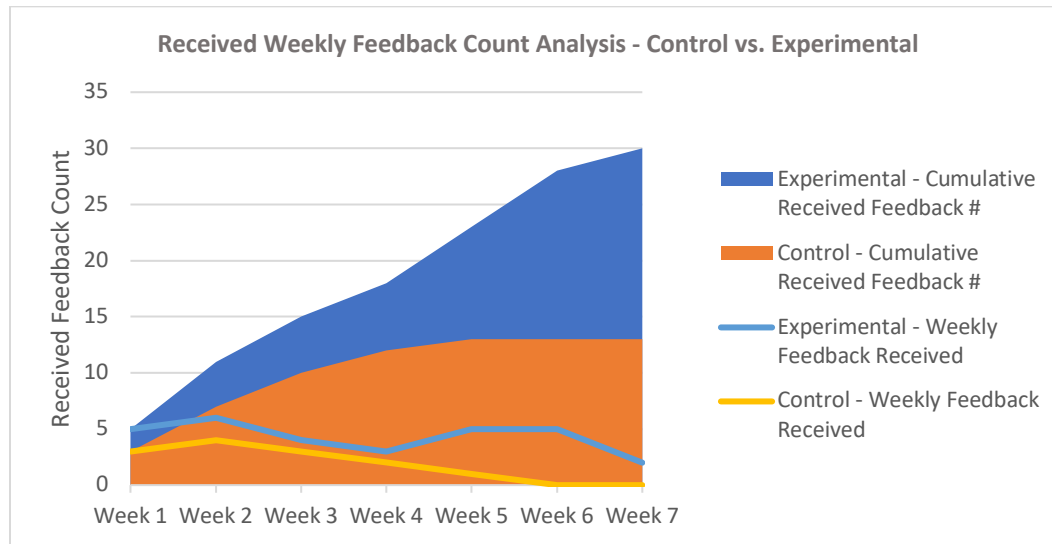


Figure 1— Received Weekly Feedback Count Analysis - Control vs. Experimental

In addition, the region in *Figure 1* that tracks the cumulative amount of feedback from the experimental group has a steeper and more steady growth curve than the control one, which means that participants who used Duolingo had more stable learning behaviours and engagement levels throughout the learning activity, while the control group's engagement was decreasing.

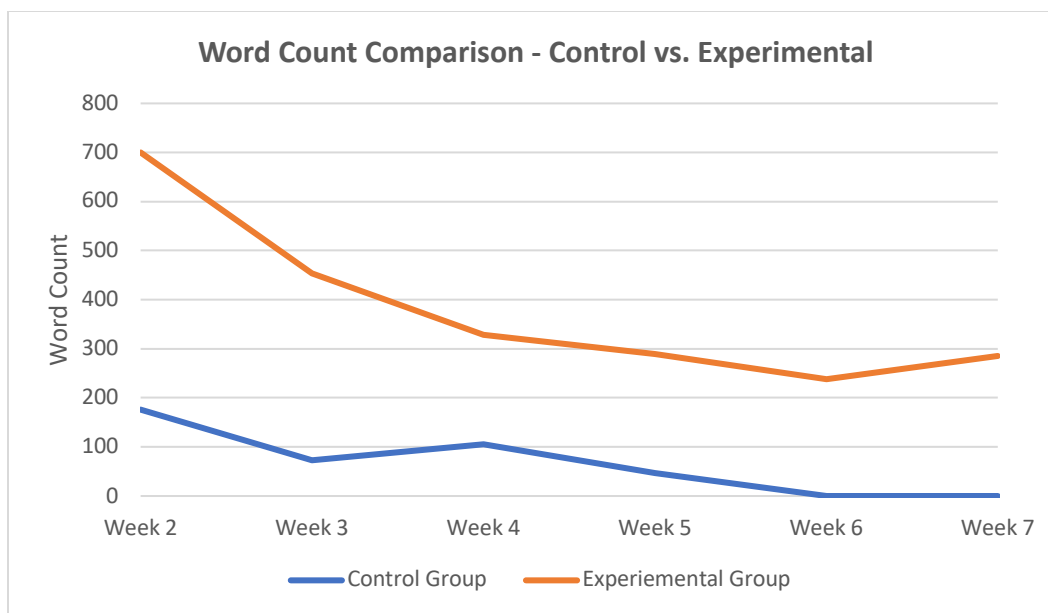


Figure 2 — Word Count Comparison - Control vs. Experimental

3.2 Learning Process

Figure 3 shows the word clouds I generated through all collected feedback except for the last week, and by comparing them, I noticed the following:

- The keyword "learning" appears more often in the feedback from the control group - this suggests that although learning behaviours were occurring in both groups of participants, the learners using Duolingo seemed to be consciously weakening the learning word or the experience of using Duolingo is not compatible with what most people perceive as learning behaviour.
- Descriptive words such as "challenged" and "struggled" appeared more often in the control group than in the experimental group - this indicated that learning a new language using traditional methods seems to encounter more considerable resistance than with Duolingo. The greater this resistance is for the learner, the less likely the learner is to continue learning.
- From the participants' feedback, it seems that the two groups of participants have a different focus on learning languages through different learning approaches - the experimental group mentioned sensory words such as "see," "feel," and "sense" multiple times during the learning

process. In contrast, the control group mentioned "vocabulary," "grammar," and "practice" more often.

- Participants in both groups frequently mentioned the words "update" and "progress" in their weekly feedback, which shows that as long as learning behaviours were engaged, the participants more or less believed that they were progressing in learning the new language.

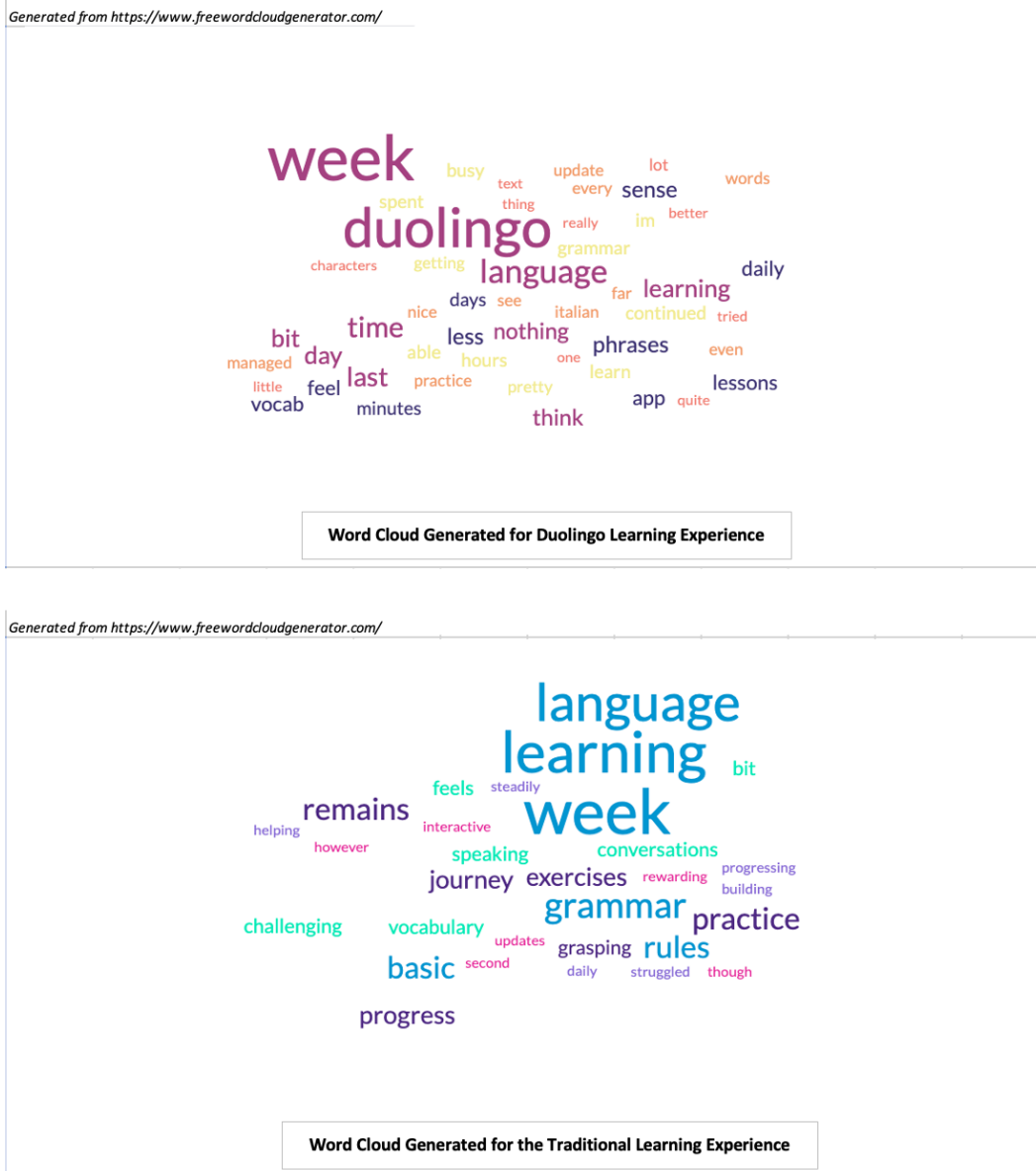


Figure 3 — Word Count Comparison - Control vs. Experimental

3.3 Learning Outcomes

Since I have not received any feedback from the control group participants in the final week and because of the limitations mentioned in subsection 2.2, I will explain the participants' learning outcomes from the existing research findings and the quantitative analysis results in this section.

Research conducted by Richardson and Swan (2003) emphasized that higher social presence within online courses directly links to increased perceived learning and satisfaction among students, which echoes Garrison and Vaughan's (2008) findings, highlighting how effective communication and interaction among participants significantly contribute to student engagement and overall learning experiences. It is evident that fostering engagement, as supported by these scholars, is pivotal in positively correlating with learning outcomes in Education.

Since previous quantitative analysis has shown that the experimental group using Duolingo had a higher engagement level compared to the control group, I predict that participants in the experimental group will have better outcomes in learning the new language than the control group.

4 CONCLUSION

The study employed *Figures 1, 2, and 3* to compare participants' engagement levels using Duolingo versus traditional methods for language learning. *Figures 1 and 2* demonstrated a consistent trend: the experimental group, using Duolingo, consistently provided more feedback and words, highlighting their higher engagement throughout the learning process. Moreover, the experimental group exhibited a steadier growth curve in their cumulative feedback, suggesting sustained and stable engagement compared to the declining engagement in the control group. *Figure 3's* analysis revealed nuanced differences in learning approaches, with the experimental group focusing on sensory experiences while the control group emphasized vocabulary and grammar.

Despite limitations such as the absence of feedback from the control group in the final week, the comprehensive quantitative data analysis reinforced the association between engagement levels and learning outcomes. The study's conclusions highlighted the instrumental role of heightened engagement, facilitated by innovative tools like Duolingo, in enhancing learning experiences and influencing

positive learning outcomes. Studies by Garrison and Vaughan (2008) and Richardson and Swan (2003) also support this assertion.

5 DISCUSSIONS

5.1 Critique

Different languages have different learning curves, and the difficulty level of listening, reading, and writing varies from language to language. For example, Chinese pronunciation is relatively easy for native English speakers because there are fewer vowel combinations. Still, Kanji can be challenging for native English speakers because the hieroglyphics radically differ from the alphabet. McLaughlin and Nayak's research (1989) focused on, analyzed, and discussed this phenomenon in depth. Even though in my research I have suggested that native English speakers learn Japanese and native Chinese speakers learn Arabic, some participants learned the language they were interested in anyway. This makes it difficult for me to consider the final learning outcomes of the participants based on the "same starting point," as Schiefele's (1991) studies have pointed out a relevance between the motivation and the interest in learning.

The data collected is limited regarding the time frame and number of participants, resulting in a limited scope for Generalization and Sampling Bias. With a small number of participants, the diversity within the population may need to be adequately represented. If a study only includes a small group from a particular age group or socioeconomic background, the conclusions drawn might not accurately reflect the perspectives or behaviours of the wider population.

Mechanical intervention by the researcher and the external makes the data hardly reflect reality - To obtain a certain amount of data to complete the study; I sent participants a daily study reminder for the first two weeks and a weekly feedback reminder every Friday. Although I informed participants before they started to conduct the learning action that the weekly feedback did not imply that I encouraged them to learn more and that feedback was to tell me they were engaged participants, many participants still interpreted my weekly feedback reminder as a predisposing behaviour and thus bring potential confirmation biases to the data I collected.

5.2 Future Work

Based on the previous section's critique, two exciting directions in the field of online learning will fuel the development and research in Educational Technology.

1. The impact and extent of the researcher's intervention on the study participants, regarding the data collection and the analysis results - Findings in this track help researchers revisit their experimentation flow and data-gathering design in educational technology to obtain more reliable experimental results. In practical terms, these findings may inspire educators on how to guide their students.
2. Implicit assessment of online learning outcomes - Research on implicit assessment of online learning outcomes can benefit educational technology by providing nuanced insights into learners' subconscious knowledge retention and skill acquisition. This data enables tailored adaptive systems, improving personalized learning experiences and refining pedagogical approaches for more effective online Education (Ha & Im, 2020).
3. Effects of repetitive behaviours on learning - As shown in *Figure 2*, the word count in the participant's feedback decreased week by week; this phenomenon seems to implicitly indicate that the participants were tired of the act of sending me weekly feedback, even though repetition is essential for reinforcing memory and skill acquisition (Wogan & Waters, 1959). Therefore, I think it is meaningful to discuss how the form and frequency of repetitive behaviours affect learners' learning status and outcomes.

6 REFERENCES

- Garrison, D. R., & Vaughan, N. D. (2008). Blended learning in higher Education: Framework, principles, and guidelines.
- McLaughlin, B., & Nayak, N. (1989). Processing a new language: Does knowing other languages make a difference. *Interlingual processes*, 1(5), 179–187.
- Richardson, J. C., & Swan, K. (2003). Examining Social Presence in Online Courses in Relation to Students' Perceived Learning and Satisfaction.

Schiefele, U. (1991). Interest, learning, and motivation. *Educational psychologist*, 26(3-4), 299-323.

Wogan, M., & Waters, R. H. (1959). The role of repetition in learning. *The American journal of psychology*.

Ha, Y., & Im, H. (2020). The Role of an Interactive Visual Learning Tool and Its Personalizability in Online Learning: Flow Experience. *Online Learning* (Newburyport, Mass.), 24(1), 205-. <https://doi.org/10.24059/olj.v24i1.1620>.

Zou, L., & von Ahn, L. (2023). Interview with the founder of Duolingo. YouTube. https://www.youtube.com/watch?v=ejlixcznawg&ab_channel=%Ed%BP%jFLin%Ej%AF%BO.