The Effect of Note-Taking Medium on Memory Recall as Guided by the External Storage Hypothesis

20 March 2025

Word Count: 2195

Introduction

Currently the prevalence of higher education in the United States is significant, with a 116% increase in college enrollees over the past decade. This high number of students inspires the question which this study will investigate: the effect of note-taking mediums on memory recall. The idea of recall and precursory actions such as note-taking cannot be investigated without mentioning the external-storage hypothesis, the idea that studying from notes allows for easier retention of information by providing a medium to review the material (Mueller & Oppenheimer, 2014, 2). This study exemplifies the external-storage hypothesis by placing a requirement upon all participants to take notes, creating an external-storage bank, and providing participants a brief opportunity to review their notes before undergoing a memory recall task, which will demonstrate whether the participants are able to retain information better due to the creation of a medium from which they can review.

This study is based on Pam A. Mueller and Daniel M. Oppenheimer's 2014 study titled "The Pen is Mightier than the Keyboard: Advantages of Longhand Over Laptop Note Taking", published in the *SAGE journal* of psychological science. The study aimed to investigate whether taking notes on a laptop versus writing longhand affects academic performance(Mueller & Oppenheimer, 2014, 4). The study was based on the external-storage hypothesis as well as the encoding hypothesis, the idea that "the processing that occurs during the act of note taking improves learning and retention" (Mueller & Oppenheimer, 2014, 1). The study tested this on 67 Princeton university students, nearly evenly split between male and female, and excluded only two subjects from the reported data set (Mueller & Oppenheimer, 2014, 2), finding that "participants who took longhand notes and were able to study them performed significantly better than participants in any of the other conditions" (Mueller and Oppenheimer 6). This suggests that hand-written notes are a more effective external information store as compared to typed notes.

The aim of this study is to investigate which medium of note-taking is more effective in memory retention for high school students: hand-written or digital notes. The population used in the study involves high school freshmen aged around 14, due to their unique academic circumstance. Transitioning into high

schoolers, their workload has drastically increased, requiring all students to present a greater effort in their classes, one example being taking notes. Additionally, the current assimilation of technology into the classroom poses an interesting situation; on one hand, technology empowers students to take extensive notes that are easy to access, while on the other hand, technology poses a distraction from class. This study aims to address the benefits of technology in terms of memory retention to ultimately suggest the preferred memory storage method which students should adopt for optimal memory retention.

The study is based on the external storage hypothesis, the idea that learning from a bank of information can improve memory recall ability in a manner which scales with the size of the bank of information.

The independent variable of this study is the medium used for notetaking, either handwritten notes or typed notes, with an established control group of no medium for notes. The dependent variable of this study is the participant's score on a memory recall exam which requires participants to complete short-answer questions related to a lecture over which the participants will take notes.

The null hypothesis of this study is that the notetaking method used by the participant will have no effect on their score on the memory recall task.

The research hypothesis of the study is that participants' scores on the memory recall tasks will be higher while utilizing a computer to take notes because the lower dexterity requirement for typing allows for more information to be stored in the same time frame, allowing for a more extensive bank of information to study from.

Exploration

The experiment was a repeated measure design. This design was selected to account for different abilities in concept retention. The individuals first undergo the experiment with the handwritten note condition, then the typed note condition. This allows a participant to be compared to a self-established benchmark, thus eliminating a potential confounding variable: differences in students' abilities to retain information. Following the design, participants are allocated into both conditions.

In order to ensure the ethics of the experiment, participants were briefed and debriefed; participants were told the research area beforehand and the aim after the fact and reminded of their right to withdraw. Participants were required to provide consent before the experiment was carried out; the consent form can be found in appendix A, the briefing script can be found in appendix B, and the debriefing script can be found in appendix C.

The sample was acquired through convenience sampling. The participants were organized into two rooms made up of two sets of participants. The participants were split due to the nature of our agreement to procure subjects, which allotted us limited time slots with each group. The rooms started with opposite initial conditions, and then swapped. There was no concern of participants in different rooms interacting with each other and potentially introducing demand characteristics into the data set. The participants were provided by our educational institution.

The sample consisted of 29 high school freshmen with an average age of 15.. The participants were all students at a prestigious institution known for academic rigour, and as such may consume information at an accelerated pace as compared to others of their age group.

The lectures presented to students were selected to be specific and niche, and were played in 5 minute increments. Additionally, the lectures were sourced from college lectures to ensure that participants would be able to understand the material.

In order to avoid confounding variables, participants were asked to use the restroom prior to the experiment, and were asked not to leave to prevent experimental mortality and history bias, a request which all participants honored. Additionally, participants were not permitted to speak to one another, nor were they permitted to share notes in order to prove the effect of note taking on a student's success rather than the effect of studying new material. On the tests, questions were made extremely specific to ensure no instrumentation bias, and participants were required to answer with short phrases to ensure that participants' scores were not inflated by guessing. Participants were asked prior to each experimental condition whether they had any familiarity with presented topics, and no participants admitted familiarity.

Materials

- HDMI cord (To connect the laptop to the projector screen)
    - Sound system in projector
- Laptops (school provided)
- Test papers (To be able to get our results in)
    - Quiz I (see Appendix II)
    - Quiz II (see Appendix II)
- Extra paper (for notes)
- Extra pencils (To write notes)
- 1st lecture:
    - https://www.youtube.com/watch?v=mcrGC0KTfnY&t=1085s
        - 14:35 to 19:35
- 2nd lecture:
    - https://www.youtube.com/watch?v=EDRa-ESxmJY
        - 2:00 to 7:00

Procedure

1. Participants are briefed (See Appendix B)
2. Participants receive a 5-minute lecture about a given topic (this lecture will include objective details that remove ambiguity in the experiment) (ex colors of objects or quantities of objects) during which they will take notes on the given topic on paper.
    a. This is the first lecture:
        i. https://www.youtube.com/watch?v=mcrGC0KTfnY&t=1085s
        ii. 14:35 to 19:35
3. Participants take 2 minutes independently in the room to study the topic based on their notes.
4. Participants take a 10-question post-lecture quiz to determine their new familiarity with a topic (see Appendix D)
5. The process will be repeated for steps 1-4 on a different, yet similar in genre, lecture except the notes will be taken on a computer.
    a. This is the second lecture: https://www.youtube.com/watch?v=EDRa-ESxmJY
    b. See Appendix D for Quiz II
6. Participants will be debriefed (See Appendix C)
7. The change in score for each participant will be measured and averaged for both the control (handwritten) and the experimental (computer-taken notes) experiment and will be compared using a paired-sample t-test.

Analysis

	Handwritten Notes	**Typed Notes**
Mean	3.759	3.897
Standard Deviation	2.016	1.520

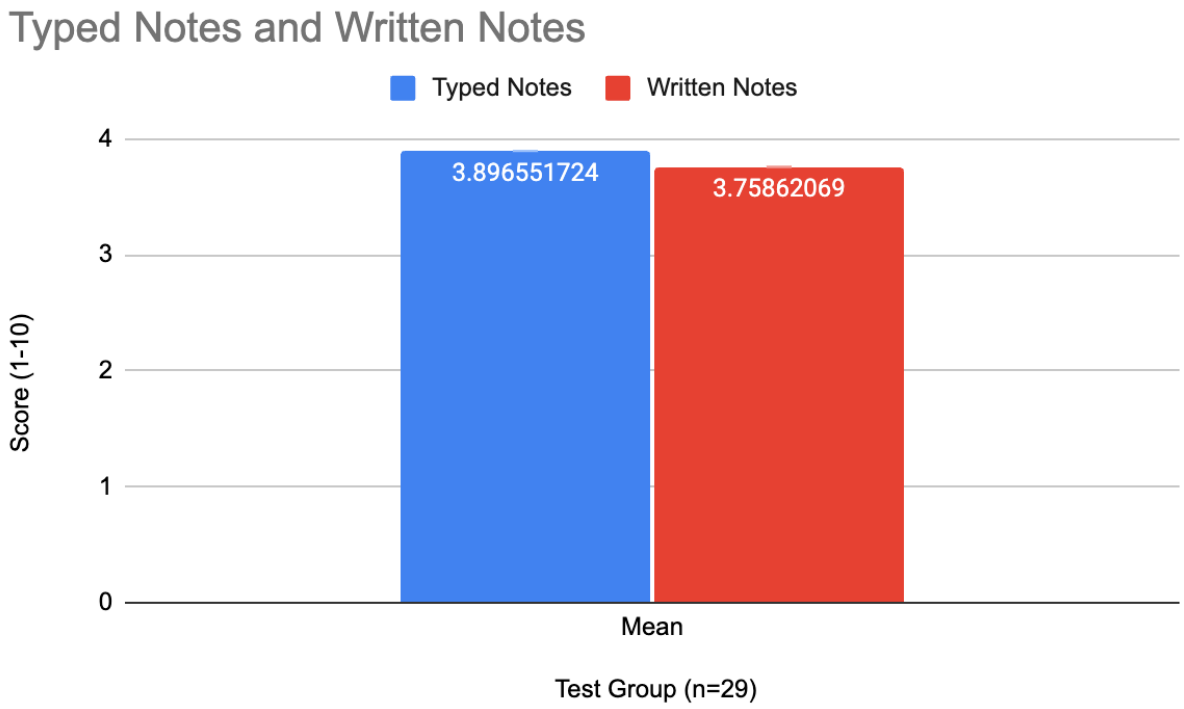
The mean was higher for the condition in which participants took typed notes as compared to the condition in which participants took handwritten notes. The mean portrays that participants were able to very quickly consume information, specifically retaining at least 37.6% of presented information for about five minutes when hand writing notes compared to at least 38.9% of presented information for about five minutes when typing notes.

The standard deviation puts all scores for handwritten notes within 2.016 points of the mean and within 1.520 points of the mean for typed notes. Note that the standard deviation for the handwritten condition is significantly higher, in part due to an extreme data point which was included. This extreme, a score of 0, was the result of a subject's limited command of the English language, which impaired their ability to understand the material. The data point was included in the set due to the presence of other extreme values. Including the low extremes allows the data set to be more balanced with the high extremes, allowing for conclusions to be drawn while retaining the sample size.

Aside from these extremes, there were no outliers. The measure of variance signifies insignificant data The data is positively skewed for the handwritten condition and is negatively

skewed for the typed condition. The data overlaps almost exactly across conditions, making the results of the study insignificant.

The mean and standard deviation were selected due to their accurate representation of groups as a whole while minimizing extreme data. Additionally the standard deviation depicts the deviation of the data allowing outliers to be identified.



To determine the ability of the data in making conclusions, inferential statistics were gathered. The data collected was ratio, with equal intervals of one from zero to ten, with zero representing a lack of knowledge. The data was collected in a repeated measures design, in which participants underwent each condition, with the conditions being reversed for opposing rooms to retain environmental integrity. The inferential statistic measured was standard deviation. The standard deviation for the Handwritten condition was higher than the mean, while it was lower than the mean in the Typed condition. Due to the standard deviation and the data overlap, the data collected is insignificant. Thus, the null hypothesis is accepted. There is no concern about type II error, due to the small sample size and low population validity of the experiment. The p value of the data is .3.

Evaluation

The original study conducted by Mueller and Oppenheimer found “ a significant interaction between” note-taking “medium [and] opportunity to study” (Mueller and Oppenheimer, 6). Additionally, the study found that “participants who took longhand notes and were able to study them performed significantly better than participants in any of the other conditions” (Mueller and Oppenheimer 6), suggesting that hand-written notes are a more effective external information store as compared to typed notes. However, our study was unable to replicate these results, instead finding insignificant data relating the conditions, meaning no relationship can be established.

The results may be attributed towards the low population validity of our experiment. The sample of participants was composed of only 29 high-school aged participants, who had no incentive to cooperate. Additionally, errors occurred during testing which could have introduced confounding variables. During one of the test conditions, a researcher failed to properly display the qr code linked to the quiz. This failure prolonged the time during which respondents could study their notes. The researchers attempted to address this by immediately collecting participant notes, but the confounding variable cannot be ignored. Additionally, during one of the test conditions, participants were incentivized to leave quiz answers blank. One of the test conditions was conducted before the student participants of the study would be released for lunch, and the experiment was run past the time at which students would have been released. Researchers did not draw attention to the time during the experiment, yet some students noticed and may have left quiz questions blank in order to leave quickly.

The external storage hypothesis does not support the insignificant results of the study, nor does it support the results of Mueller and Oppenheimer’s original study. The external storage

hypothesis suggests that learning from a bank of information can improve memory recall ability in a manner which scales with the size of the bank of information. Although the typed condition was able to amass a significantly larger bank of information due to the rapid nature of typing, the condition still had lower average scores than that of the handwritten condition.

The use of a repeated measures design in the study was effective in eliminating the confounding variable of an intelligence difference between students across conditions. However, the design was unable to factor in the interdependence of students upon each other with questions.

The sample was gathered in volunteer sampling of high school freshmen from Southwest Indiana. The population was homogeneous with an even split of males to females, and a variety of racial identities. The sample was halved to accommodate two different testing rooms. In Room A, the sample moved very smoothly through their conditions, while Room B had difficulty. One participant in Room B reported a low aptitude for English which skewed their understanding of the lecture material and thus lowered their score. Additionally, participants attempted to converse during the experiment, requiring researchers to intervene.

The procedure used had consistent sampling, and high internal validity. The controlled experimental environment allows for identification of a causal relationship between the measured and manipulated variables. However, the environment reduces the ecological validity of the results, suggesting that participants' results from a classroom environment differ. The procedure managed to eliminate instrumentation bias with a standardized response grading key, and eliminated history bias by encouraging participants to stay in the room until the experiment was completed, which all participants adhered to.

The materials were standardized and taken from introductory college psychology classes, and were selected to be unknown to participants. This may have reduced participant scores by presenting them with topics that they had difficulty understanding.

In order to address the low population validity, the population could be sampled from a university psychology department, representing more age groups, races, and socioeconomic backgrounds. Additionally, the materials can be improved by presenting custom-made lectures on an imaginary topic, presented in a way that isn't whimsical, requiring participants to focus during lecture to answer questions.

To conclude, the p value of the experiment was .3, leading to the acceptance of the null hypothesis, the idea that note-taking medium has no effect on recall. This means that no relationship can be established between the independent variable and the dependent variable.

References

MIT OpenCourseWare. (2012a, August 9). *5. Are Infants Little Scientists?* YouTube. https://www.youtube.com/watch?v=mcrGC0KTfnY

MIT OpenCourseWare. (2012b, August 9). *8. Addressing Molyneux's Query*. YouTube. https://www.youtube.com/watch?v=EDRa-ESxmJY

Mueller, P. A., & Oppenheimer, D. M. (2014, May 22). The Pen is Mightier Than the Keyboard: Advantages of Longhand Over Laptop Note Taking. *SAGE psychological science*, 1-10. 10.1177/0956797614524581

Appendix A. Consent Form

Hello, we are experimenting as part of our internal assessment of our IB Psychology SL class. We are investigating the effect of different note-taking methods on memory recall. We would like to ask you to be part of our experiment. If you agree to take part in this experiment, you should know that:

- All data that we obtain will be kept confidential and anonymous.
- You may stop participating in this experiment at any time.
- You shall receive information about the nature of the experiment and our results after we have obtained results.

The experiment will take about twenty minutes to complete.

If you agree, we ask that you sign the form below and fill in the following information relevant to our experiment. If you are under the age of 16, a parental guardian must also sign the letter of consent.

Age _____

I, _____understand the nature of this experiment and I agree to participate voluntarily. I permit the researchers to use my data as part of their experimental study.

Participant Signature: _____

Parental signature: _____

Date:

Appendix B. Briefing

Thank you for participating in this experiment.

First I will give you a paper to sign. This is called an informed consent letter. You should know that you are entitled to withdraw from the experiment in case you change your mind and don't want your data used in the study, please contact an experimenter after the study in case you want your data to be omitted. Furthermore, your data is confidential.

Don't hesitate to ask questions if you have any. After you have signed the paper, I will read some instructions about the experiment to you.

Please listen carefully and do not talk to the other participants. The other experimenter will now distribute a piece of paper to you. Please take out a pencil to write on the paper when I instruct you to do so.

The experiment consists of two stages and in each stage, you will be given an audio of a short lecture about the same topic. Take notes on the paper given to you. After the lecture, you have two minutes to independently study your notes. Finally, you will be handed a 10-question quiz on the subject, you are not permitted to use your notes during these three minutes. We will collect this form at the end of stage one. Please do not write your name on any materials.

I will then repeat the same procedure in the second stage of this experiment and your note-taking method will be altered.

Appendix C. Debriefing

Thank you for participating. Please raise your hand if you still have any of the following: a computer with notes on it, a piece of paper with notes on it, a question, or a test.

This experiment was based on a psychological study conducted by scientists Mueller and Oppenheimer titled "The Pen is Mightier than the Keyboard", which aimed to understand how different note-taking methods can affect memory recall. In the experiment, participants who took notes on paper performed 8% better overall when given time to review after the lecture compared to participants who typed out their notes.

We gave you a post-test to gauge your understanding of the topic after the lecture. We will compare the difference in your scores between the two different note-taking methods using a paired-sample t-test, to see which note-taking method is more conducive to learning on average for individuals.

As participants, you are well within your rights to ask about the results of the experiment, and if you'd like you can read our paper once it is complete. Thank you for your participation and feel free to ask any final questions!

Appendix D. Questionnaires

Appendix D.1

1. What year was the Baillargeon study released?
2. What paradigm did the Baillargeon study use?
3. What is evidence for the violation of an expectation?

4. Aside from triangles, what was the other shape presented in the example used to explain how the infant-looking time paradigm works?
5. How many degrees does the board in the study rotate throughout the experiment?
6. What is the name for the phase of the experiment where infants are getting used to seeing a certain kind of display? (not familiarization)
7. Which test condition would be considered physically impossible by an adult?
8. What concept does Baillargeon pit against the physical possibility?
9. What is the name of the second phase of the experiment?
10. In the first test frame, an adult identifying that the board stopped on the hidden rubber duck proves what concept?

Appendix D.1A

1. What year was the Baillargeon study released?
    a. 1987
2. What paradigm did the Baillargeon study use?

    a. Infant-looking time paradigm

3. What is evidence for the violation of an expectation?
    a. Surprise
4. Aside from triangles, what was the other shape presented in the example used to explain how the infant-looking time paradigm works?
    a. squares
5. How many degrees does the board in the study rotate throughout the experiment?
    a. 180
6. What is the name for the phase of the experiment where infants are getting used to seeing a certain kind of display? (not familiarization)
    a. habituation
7. Which test condition would be considered physically impossible by an adult?
    a. Condition 2
8. What concept does Baillargeon pit against physical possibility
    a. Perceptual similarity
9. What is the name of the second phase of the experiment?
    a. test
10. In the first test frame, an adult identifying that the board stopped on the hidden rubber duck proves what concept?
    a. solidity

Appendix D.2

1. What condition does the man in Molyneux's Query Possess?
2. What century did Molyneux send his query to its recipient?
3. How many applicable cases have there been over the past 1000 years?
4. Name one of the conditions that can cause congenital blindness.
5. Name one of the regions where these cases are virtually nonexistent.
6. Light cannot be placed onto which part of the eye when one receives congenital blindness?
7. What is the country mentioned where many people grow up blind?
8. What is the most likely cause of congenital eye blindness
9. What was the project mentioned in the video?
10. Who was the Molyneux Query sent to?

Appendix D.2A

1. What condition does the man in Molyneux's Query Possess?
    a. Blind(ness)
2. What century did Molyneux send his query to its recipient?
    a. 17th century
3. How many applicable cases have there been over the past 1000 years
    a. Less than 20
4. Name one of the conditions that can cause congenital blindness
    a. Extreme Refractive Error/Corneal Opacities/Occlusion of lights by cataracts
5. Name one of the regions where these cases are virtually nonexistent

    a. The United States/Europe

6. Light cannot be placed onto which part of the eye when one receives congenital blindness
    a. The retina
7. What is the country mentioned where many people grow up blind
    a. India
8. What is the most likely cause of congenital eye blindness?
    a. Cataracts
9. What was the project mentioned in the video
    a. Project Prakash
10. What is the speaker's job (other than professor)
    a. Neurologist/neuroscientist

Appendix E. Raw Data

Written Score	Typed Score
3	5
0	1
2	5
4	4
3	6
2	3
5	5
4	5
4	6
2	5
1	5
3	4
0	5
3	5
5	4
7	2
8	4
6	3
3	3
7	4
6	2
3	6
5	1
4	4
6	3
3	5
5	2
2	1
3	5

Appendix F. Descriptive Statistics

To find the mean, all of the scores were added and divided by 29.

$$\sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{N}}$$

Where

N (population size) = 29

μ (population mean) = 3.759 for Handwritten and 3.897 for Typed

x_i (score of a respondent)

Appendix G. Inferential Statistics

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$s^2 = \frac{\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2 + \sum_{j=1}^{n_2}(x_j - \bar{x}_2)^2}{n_1 + n_2 - 2}$$

Where

x = sample mean

s^2 = pooled sample variance

n = sample size

t = Student t quantile with $n_1 + n_2 - 2$ + degrees of freedom