Is Note-Taking More Effective with a Keyboard or a Pen?

Stacee M. Horwitz

Department of Psychology and Neuroscience at University of Colorado Boulder

Thesis Defense: April 4, 2017

Committee Members:

Psychology Department: Eliana Colunga and Tim Curran (Thesis Advisor)

Chemistry Department: Nabilah Carlon

Author Note

Stacee M. Horwitz, Undergraduate Student in the Department of Psychology and Neuroscience, University of Colorado, Boulder.

Tim Curran, Alice Healy, and Lakshmi Lalchandani provided guidance on this project, including advice about the experimental design and the writing of the thesis. In addition, Lakshmi Lalchandani provided her program for computing verbatim overlap. Alice Healy encouraged me to conduct this thesis through her course on research methods, in which I conducted an original mini-experiment on this topic. Tim Curran helped with the statistical analyses and served as my primary advisor on this project. I am deeply grateful for their full dedication in helping me complete this thesis.

Correspondence concerning this manuscript should be addressed to Stacee M. Horwitz, Undergraduate Psychology Major, 909 14th St, Boulder, CO 80302. E-mail: Stacee_horwitz@colorado.edu
Abstract

As the educational system attempts to incorporate technology into the daily regimen, controversy over technology’s impact has escalated amongst students and teachers. This study researched the impact of longhand versus typed note-taking medium on test performance. Past research has discovered that longhand note taking of a video lecture outperforms typing notes for conceptual test questions (Mueller & Oppenhime, 2014). In Experiment 1, the purpose was to examine how longhand note taking compares to typing notes from a textbook passage, and also the effects of taking notes versus only receiving notes to review. The participants read a textbook passage and either took longhand, laptop, or no notes of the passage. Before taking the posttest, participants reviewed notes they created, or received from past participants. A surprising result occurred in which all subjects had similar factual and conceptual test performance, regardless of the note-taking medium. Experiment 2 was later created to examine if taking or receiving notes had any effect on test improvement, by having a no-note control group only read the passage. Even though this group did not create or receive notes, there was no difference in test scores compared to subjects who created notes, implying that test improvement was mainly due to reading the passage. However, subjects who received typed notes had a higher factual test performance than the control group, which signifies the value of receiving verbatim notes for learning factual information.

*Keywords*: technology, note-taking medium, word count, verbatim overlap, comprehension, memory, generative versus shallow processing
As technology changes, the types of note-taking methods used in the educational system are changing as well. Typing notes on a laptop is increasingly popular; however, educational psychologists have suggested that longhand notes are more beneficial for learning (Aragón-Mendizábal, Delgado-Casas, Navarro-Guzmán, Menacho-Jiménez, & Romero-Oliva, 2016; Kiewra, 1985; Mangen & Velay, 2010; Mueller & Oppenheimer, 2014; Smoker, Murphy, & Rockwell, 2009; Yamamoto, 2007). Past studies pertaining to this topic have focused on laptops being distracting since Internet browsing is so easy (Hembrooke & Gay, 2003; Kraushaar & Novak, 2010). Research has also suggested that laptops result in shallower processing of learning when subjects listen to a lecture (Mueller & Oppenheimer, 2014). There is a gap in the educational field for research on learning from textbook passages, and this study explores the most optimal note-taking methods to improve academic performance; the current study researches the effects of longhand notes versus laptop notes of a textbook passage, along with the impact of not creating notes, but receiving notes for review. This study will hopefully bring clarity to the current controversy in education with regards to technology. The millennial generation prefers to learn with technology; however, many educators view technology as controversial (Gipson et al., 2017).

Members of the millennial generation grew up with technology, such as cable TV, Internet, cellphones, and laptops. For this generation, technology is a suitable medium for entertainment, education, and socializing (Palekanda & Venkataraman, 2012). This generation challenges the educational system with its
preference for laptops in the classroom; however, teachers commonly find that technology is counterproductive to the learning regimen (Gipson et al., 2017; Yamamoto, 2007). The distractive roles of computers serve as a major disadvantage for education; it is easy for students to chat with friends or search the web while taking notes, instead of focusing on the teacher (Hembrooke & Gay, 2003; Mueller & Oppenheimer, 2014). Class performance tends to decline when students pay more attention to their device than the lecture (Hembrooke & Gay 2003).

Typing is a quicker medium than longhand note taking; this gives students an opportunity to browse the Internet as they wait for their peers to finish writing (Brown 1988; Hembrooke & Gay 2003; Yamamoto 2007). Research shows that 42% of the time students have non-coursework applications active on their laptops, which exhibits how often multitasking occurs (Kraushaar & Novak, 2010). When multitasking in the classroom, students are not using their full cognitive ability to comprehend or retain the information (Linnell, Caparos, de Fockert, & Davidoff, 2013). Instead, they are passively taking notes while they check their email or social media. On the other hand, since longhand note taking is slower, it allows a deeper engagement with the material, and gives time for the brain to store the incoming information (Alonso, 2015; Longcamp et al., 2008). Longhand note taking is more advantageous because technology not only decreases task performance in recall and recognition skills, but also shortens attention spans (Hembrooke & Gay, 2003; Linnell et al., 2013).
Laptops give students the opportunity to become further distracted from the cognitive task at hand (Linnell et al., 2013). The Pew Research Center discovered that 87% of teachers find that technology causes shorter attention spans and that 67% of teachers believe that the distraction of technology hinders education more than it benefits (Purcell et al., 2012). These studies on laptop distraction represent how technology gives students an extra opportunity to become distracted, which results in a decline of class performance. Incorporating technology into education has disadvantages that extend from class distraction to impaired memory.

Laptops have the potential to negatively impact memory. In research testing memory, a group of adults typed or hand wrote a list of words, and afterwards participated in word recall and recognition tests. The results exhibit that longhand words have higher accuracy scores on recall and recognition, compared to typed notes (Smoker et al., 2009). However, longhand notes do not only benefit adults. Research on preschool children found similar results that support the longhand method over typing. Learning letters longhand resulted in significant improvement on recognition performance, compared to the children (3-5 years old) who typed letters (Longcamp et al., 2005). This memory retention expands beyond just letters to learning new geometrics (Longcamp et al., 2008); the sole act of taking longhand notes improves memory in a way that typing does not offer. The memory performance linked to longhand notes is due to the mere efforts of taking the extra time to use fine motor skills in creating generative notes (Alonso, 2015).
Handwriting involves motor memory through the high intake of kinesthetic information that is encoded while learning. The process of learning kinesthetically is through the physical movement of a pen on paper; these movements are activating regions and connections in the brain that do not occur through typing (Longcamp et al., 2008). The muscle movement used to handwrite letters activates the sensorimotor region of the brain, which causes reinforcement of the learned information through motor memory (Longcamp et al., 2008; Alonso, 2015). Handwriting involves strong connections between the left fusiform gyrus and the middle frontal gyrus, which are regions associated with working memory and fine motor skills (James, Jao, & Berninger, 2016). These brain connections are not found in typing; it is easier to passively type on a keyboard compared to the motor demands of handwriting. Further, longhand notes not only affect the brain due to its motor movements, but from cognitive and visual aspects as well (Longcamp et al., 2008).

Writing is a complicated mental process that heavily relies on a combination of sensory and motor techniques; when a person writes by hand, the brain plays a more active role in forming each letter, compared to passively pushing letters on a keyboard (Mangen & Velay, 2010). Handwriting involves a deeper cognitive process that leads to better comprehension, ideas, and memory (Bui & McDaniels, 2015; Longcamp et al., 2008; Lotze, Erhard, Neumann, Eickhoff, & Langner, 2014). Many more brain regions in the frontal and parietal lobe play a role in longhand note taking compared to typing; also, there are more brain connections (i.e., left fusiform with left pre/postcentral gyri) between regions
that only occur from handwriting (James et al., 2016). Since handwriting is a more complex thought process, it relates to creativity in paraphrasing notes, which is more generative than verbatim repetition from a laptop (Mueller & Oppenheimer, 2014). Further, research tested how the brain functions differently when it uses deep cognitive processes, like creativity, compared to only copying words using shallow cognition (Berninger, Abbott, Augsburger, Garcia, 2009).

Research involving a link between handwriting and creativity found that students who hand wrote their essays had more ideas and complete sentences, in comparison to their typed essays (Berninger et al., 2009). Handwriting involves a deeper thought process with more planning and creativity, than typing essays (Berninger et al., 2009). Furthermore, functional magnetic resonance imaging (fMRI) discovered that some areas of the brain (in the occipital, frontal, and temporal lobe) make unique connections during the creative process (Lotze et al., 2014; Howard-Jones, Blakemore, Samuel, Summers, & Claxton, 2005; Seger et al., 2000). In the Howard-Jones et al., (2005) study, subjects received random words from which they had to make a story, while an fMRI assessed their brain activity. When subjects were creative, the fMRI showed the main increase in bilateral activity between prefrontal brain regions (Howard-Jones et al., 2005). This study explains another benefit of longhand notes, because they involve more creativity to write informational summaries, compared to typing notes verbatim (Lotze et al., 2014). Furthermore, when a student listens to a lecture, longhand notes lead to lower word count and less verbatim overlap with the lecture, because of the enhanced complexity of the longhand generative brain
process (Mueller & Oppenheimer, 2014). Longhand note taking is a more complicated process mentally and physically; higher brain functions are necessary for optimal memory and comprehension, which is why current research studies the effects of technology on learning.

The educational system cannot hide from technology; therefore, it is important to look at sufficiently integrating modern technology with the traditional learning of longhand note taking. Transitioning from longhand to typing notes could have a major impact on memory because the note-taking medium affects the cognitive processes (Alonso, 2015). Students’ preference for technology is increasing; they enjoy the speed and efficiency that laptops give. Laptops allow students to type faster, and for a longer time (Bui, Myerson, & Hale, 2013). The classroom needs to keep the positive effects of handwriting, while finding effective methods to incorporate technology into learning (Caruso & Kvavik 2005; Gipson et al., 2017). Current research explores incorporating handwriting and technology, so that benefits from each medium will optimize education. If technology incorporates generative note-taking media, then it will have the potential to enhance academic performance (Alonso, 2015).

Current research explores touch-screen devices that allow students to use a stylus to create longhand notes, because the kinesthetic movements of typing are ineffective with helping memory (Smoker et al., 2009). Touch screen technology will only be effective if it can incorporate the generative brain processing that is derived from longhand (Alonso, 2015). A stylus provides notes that are written on the screen, and can still carry the advantages of the traditional
pen and paper style notes (Longcamp et al., 2008). Furthermore, a popular touch screen device is the iPad, which is commonly found within classrooms. This technology has a positive acceptance from students who believe it enhances their education and note taking (Alyahya & Gall, 2012); however, academics commonly find iPads a distraction. Current research is inconclusive for the impact of iPads on learning outcomes (Kinash, Brand, & Mathew, 2012; Nguyen, Barton, & Nguyen, 2015). Future research is necessary for determining effective technology methods in classrooms, and the effects of current touch-screen devices. Educational technologies continue to evolve at a fast pace; the exploration of technologies is an active and growing field (Nguyen et al., 2015). Further, research looks into the effects of technology on memory and comprehension.

In a three-part study, Pam Mueller and Daniel Oppenheimer (2014) conducted a research project on this controversial topic of laptops in the classroom. In the first study, subjects watched TED talks and created typed or handwritten notes on the lectures. After they watched the lectures, they had 30 min of distractor tasks. Afterwards, they were asked factual recall and conceptual application questions. The participants in the two note-taking conditions performed the same for factual questions. However, students who used laptops performed significantly worse on conceptual questions. They found that lecture typed notes have a larger word count compared to longhand notes, because subjects can type significantly faster than they can write. Additionally, they discovered that the amount of verbatim overlap grew as word count increased.
Subjects who hand wrote notes had a verbatim overlap score of 8.8% with the lecture, compared to the subjects who typed notes with a verbatim score of 14.6% (Mueller & Oppenheimer, 2014).

Mueller and Oppenheimer (2014) wanted to see how instructions played a role in preventing verbatim overlap. In the second experiment, the subjects were instructed to take notes just like they would outside of a lab and not to repeat verbatim what they heard; however, these instructions did not prevent verbatim overlap. The longhand subjects performed the best on conceptual questions, and subjects with a higher word count also performed better overall. There was no difference between note-taking media for factual questions (Mueller & Oppenheimer, 2014). These results were all consistent with their first study. In the third study, subjects watched recordings of a “lecturer” who read from various educational passages. A week later, Mueller and Oppenheimer (2014) had the participants review their notes, in hopes of an impact on test performance from notes with a higher word count. The results were consistent that longhand subjects still had higher test performance than laptop subjects for conceptual questions, with the opportunity to review their notes. However, when subjects did not study their notes, there was no difference between the note-taking media. This three-part study revealed the benefits for longhand notes on test performance. Furthermore, Mueller and Oppenheimer solely focused on note-taking media for video lectures; they did not research any other methods for encoding information: for example, from an educational reading passage or
PowerPoint presentation. It is important to have a complete understanding of how technology affects different types of learning.

A follow-up study to Mueller and Oppenheimer (2014) tested for an interaction between question type and note-taking medium (Lalchandani & Healy, 2016). In Lalchandani and Healy’s (2016) first experiment, subjects listened to lectures and either took longhand, computer, or no-notes of the lecture. Further, some subjects saw a PowerPoint presentation that coincided with the lectures. After subjects listened to the lectures, they performed 30-min distractor tasks, took an immediate posttest, and a retention test one-week later. Lalchandani and Healy (2016) had similar results to Mueller and Oppenheimer (2014) in that longhand subjects had higher conceptual test scores on the immediate posttest, compared to computer and no-note takers. However, when subjects took the retention test, computer note takers surprisingly outperformed longhand note takers on conceptual questions. In the second experiment, Lalchandani and Healy (2016) tested the effects of reviewing notes, and not creating them. The subjects followed a similar procedure to the first experiment, but instead of creating notes they received notes from their yoked counterpart in Experiment 1; this resulted in no difference between groups on the immediate posttest. However, participants who received notes from past subjects who viewed lecture slides performed better on the factual portion of the retention test, which is most likely because the lecture slide notes had more specific information (Lalchandani & Healy, 2016). To better understand the influence of receiving notes and not
creating them, the current experiment has a “review only” condition (Lalchandani & Healy, 2016).

In the current experiment, the hypothesis that longhand note taking leads to higher conceptual test scores was formed from studies indicating that typed notes are done with shallow cognitive processing. This type of cognition leads to verbatim overlap, which impairs comprehension and conceptual test performance (Kiewra, 1985; Mueller & Oppenheimer, 2014; Yamamoto, 2007). When students hand write notes in a lecture they do not have enough time to write down every word, so they condense the material and write key details. This technique of summarizing information is a higher level of cognition compared to typing verbatim notes (Bui, et al., 2013). To handwrite notes at an appropriate pace, the students must decide what is important, and they must summarize and plan how they want to produce efficient notes. This is more cognitively demanding and complex than typing (Bui, et al., 2013). On the other hand, when students type notes, they have more time to type almost exactly what they hear, without summarizing their notes.

The primary goal of this study was to research the effect of note-taking medium on academic performance for information delivered through a textbook passage. For the educational system to benefit academia beyond a lecture format, it is essential to compare how different encoding methods play a role on test performance. This study is a follow-up to Mueller and Oppenheimer (2014); it carries similarities in procedures and test measurements. However, Mueller and Oppenheimer never looked into the benefits of only reviewing notes, not creating
them. Lalchandani and Healy (2016) tested for this by having a control group not take notes, but receive notes from a yoked counterpart in their previous experiment; this method was used in the current experiment as well. The purpose of this group was to examine the effects of review on test performance, and also how the quality of notes influenced test scores. Note quality was determined by the word count and verbatim overlap with the textbook passage. It is important to research note-taking media for reading passages, so that students know proper techniques that are efficient and beneficial to learning.

**Experiment 1**

There were four experimental groups that either created or received notes to test the effects of note-taking medium (see Table 1). The first group, *longhand note takers*, created handwritten notes of the passage. The second group, *laptop note takers*, created typed notes of the passage. The third and fourth groups did not create any notes; they received notes from a yoked counterpart in the note creating groups. The third group, *longhand note receivers*, reviewed notes from the longhand note takers. The fourth group, *laptop note receivers*, reviewed notes created by the laptop note takers. Mueller and Oppenheimer (2014) did not have groups who only received notes; the receive groups were added as a baseline in examining the sole effects of taking notes, regardless of the medium (Lalchandani & Healy, 2016). Subjects that received notes had a yoked counterpart (from one of the create conditions) who had the same test order, because test types were fully counterbalanced. Furthermore, these four groups all reviewed notes before the posttest.
Table 1. The study procedure used for testing the effects of note-taking medium on test performance, and the different experimental conditions.

During the experiment, all experimental groups took a pretest to measure prior knowledge regarding the marketing passage. Following the pretest, subjects read the passage in which some groups created notes (longhand or typed), depending on experimental group assignment. After reading the passage, all participants took a math distractor task. Then, the groups reviewed the notes that they either created or received. After participants studied their assigned notes they took a crossword puzzle distractor task, followed by a posttest.

It was hypothesized that longhand note takers would perform better on conceptual questions than laptop note takers, because research supports that longhand note taking can enhance the encoding process (Kiewra et al., 1991, 1991). Also, this hypothesis supports the results found by Mueller and Oppenheimer (2014) when testing note-taking medium for lectures. Word count and verbatim overlap with the passage were hypothesized to be higher for laptop note takers than longhand note takers, which resemble the note quality found in the Mueller and Oppenheimer (2014) and Lalchandani and Healy (2016) studies. Furthermore, the note taking groups were hypothesized to have higher test
improvement (between pre and post tests) than subjects who did not create
notes, because any act of note taking carries benefits (Dunlosky, Rawson,
Marsh, Nathan, Willingham, 2013; Kiewra, 1985; Luo, Kiewra, & Samuelson,
2016).

This experiment follows the Mueller and Oppenheimer (2014) study by
applying similar methods for testing the effects of note-taking medium on text,
rather than lecture. This study also applies the experimental design from
Lalchandani and Healy (2016) by having a review-only condition. Subjects were
tested on a textbook passage to uncover optimal study practices for coursework
that requires informational learning from reading. This experiment’s goal was to
clarify the best learning techniques of reading text, so that students are
performing to their highest academic potential.

Method

Participants. The participants were 48 undergraduate college students
from the University of Colorado Boulder enrolled in General Psychology. They all
received partial course credit for their participation. The majority of subjects were
college freshmen; all participants were over the age of 18.

Ethics, consent, and permissions. The University of Colorado
Institutional Review Board, Protocol Number 16-0664, approved the use of
human subjects.

Design. A 2 x 2 x 2 x 2 mixed factorial design was used. There were two
between-subjects independent variables; participants either took notes by
longhand or by laptop, and they either wrote their own notes or received them
from another subject \((n=12/\text{group})\). There were two within-subjects independent variables; participants were tested on conceptual and factual questions during a pretest and posttest that were given before/after reading the passage. The dependent variables are the percent correct on the factual and conceptual questions, the word count, and verbatim overlap of notes. Participants were randomly assigned to conditions by the time and date they signed up for the study.

**Materials.** Subjects received all materials needed for the experiment in a testing pamphlet. The pamphlet had general directions, and verbal instructions were stated to subjects when there were differences among the between-variable groups. Participants read a passage from the first chapter of an introductory course textbook on marketing, *Fundamentals of Marketing* (Kerin, Hartley, & Rudelius, 2015).

There were 12 participants who took notes on white, blank, printer paper that was provided. Another 12 participants took notes on a personal laptop. They typed on a blank document of Microsoft Word, with no distractions on the laptop. The other 24 participants were not allowed to take notes when reading the passage. Instead, 12 of the participants received notes from past subjects in the longhand group; the experimenter typed these notes to prevent handwriting issues. The last 12 subjects received notes from participants who used laptops.

All subjects received a posttest at the end of the study. The posttest was fully counterbalanced with the pretest, meaning that half of the participants in each experimental condition took “Test A” as the pretest and the other half took
“Test B” as the pretest, which was later followed by the opposite test to measure learning. Subjects who received notes had the same test order as their yoked counterpart who created the notes. The paper-based tests contained six factual and six conceptual test questions, in a multiple-choice format with four answer choices. An example of a factual question from Test A was, “From what country is the CEO of Chobani yogurt?” From Test B, a factual question example was, “What were the three products used as examples about satisfying consumer needs?” An example of a conceptual question from Test A was, “What is a major essence of success to marketing?” An example of a conceptual question from Test B was, “What type of product traits will make consumers more likely to purchase the MacBook Air laptop?” Test questions were chosen from all portions of the reading passage and are shown in the Appendix (“Marketing kerin 12th edition test bank,” n.d.).

**Procedure.** The experiment took place in a classroom, and all subjects were tested in a similar room on the University of Colorado Boulder campus. The subjects were randomly assigned to each group and were tested with other subjects (ranging from one to eight subjects per testing group) of the same between-subjects condition. Following the informed consent process, the subjects took a baseline pretest for 6 minutes. The participants who took notes were told that other research participants would read their anonymous notes. In the laptop condition, subjects brought their own laptop, had a blank Word document open, and were instructed that they were not allowed to browse the Web or have any other distracting elements on the screen. In the laptop and
longhand note taking conditions, participants were instructed to create notes as if they were studying for an exam (Lalchandani & Healy, 2016). Further, participants in the receive conditions were instructed to read the passage the same way they would for a class that would test them on the textbook material. Participants were not allowed to reread the passage.

There were 12 subjects who hand wrote notes as they read the passage. Another 12 subjects typed notes as they read the passage. There were 24 subjects who read the passage and did not take any notes. However, 12 of these subjects reviewed the notes from a yoked participant with the same test order. The researcher typed the longhand notes to avoid informational confusion from illegible handwriting. Twelve participants reviewed the notes from the participants who typed notes on a laptop, which were also edited by the experimenter to make sure spelling was correct and that format was consistent. Word count and verbatim overlap was not affected by the edits.

As shown in Table 1, after the participants completed the reading passage, they completed one distractor task for 7 min, and then had 6 min to review the notes they created or received, followed by 7 min for the second distractor task, and then they completed the posttest for 6 min. One distractor task was about math and the other task was a crossword puzzle about the states and capitals (Lalchandani & Healy, 2016).

To measure learning, the proportion of correct answers on each test was analyzed. All notes were analyzed for word count and verbatim overlap with the reading passage. Verbatim overlap was measured using three-word sequences
from the participant’s notes, and comparing the three-word sequences with the marketing passage (Lalchandani & Healy, 2016). The amount of identical three-word sequences found in the notes and reading passage was converted to the proportion of verbatim overlap found in the subject’s notes (the number of identical sequences divided by total number of sequences). These three measurements were used as the dependent variables to analyze the data.

**Results**

**Test performance.** A 2 longhand/laptop x 2 create/receive x 2 conceptual/factual x 2 pre/post mixed-design Analysis of Variance (ANOVA) on test accuracy revealed the main effect of pretest and posttest scores as significant, $F (1,44) = 295$, $p < .001$. As shown in Figure 1, the posttest scores were consistently more accurate than the pretest scores. Further, there was a significant interaction between the test types (pre and post) and the question types (conceptual versus factual). Figure 2 illustrates that test score improvement (posttest - pretest accuracy) was higher when the questions were factual, compared to conceptual, $F (1,44) = 27.18$, $p < .001$. However, the overall ANOVA did not show significant interactions between the note-taking medium and test scores.
Figure 1. Mean proportion of correct test answers based on test type, note-taking medium and question type.

Welch 2-sample t-tests, which correct for group differences in variability, were used to individually determine the effect of note-taking medium (longhand versus laptop) on test improvement scores. As shown in Figure 2, subjects who received laptop notes over longhand notes had marginally significantly more improvement over those who received handwritten notes, $t(21.93) = -1.92, p = 0.068$. No other conditions differed between the handwritten and laptop groups (all $p > .10$).
**Figure 2.** Mean proportion of test improvement based on note-taking medium and question type.

**Note content analyses.** When compared to those who typed notes, longhand notes had less verbatim overlap. Verbatim overlap was scored by the proportion of three-word sequences in the notes that duplicated the same three-word sequences from the reading passage. Figure 3 illustrates that subjects who typed notes had a significantly higher percent of verbatim overlap (35.47%), compared to subjects who used the longhand medium (19.98%), $t(21.96) = -2.79, p = .011$. 
IS NOTE-TAKING MORE EFFECTIVE WITH A KEYBOARD OR PEN?

**Figure 3.** Mean proportion of verbatim overlap with the textbook passage based on note-taking medium.

Note-taking medium also impacted word count, or the total number of words produced. As shown in Figure 4, a $t$-test was used to determine that those who typed notes had a significantly higher word count, $t (15.93) = -3.98$, $p = .001$.

**Figure 4.** Mean word count based on note-taking medium.

Pearson correlation tests were used to look at the relationship between quality of notes (word count and verbatim overlap) and overall test improvement.
(post-pre accuracy). The tests combined the laptop and longhand groups so that the data analyzed the relationship between test performance and note quality (word count and verbatim overlap), regardless of the note-taking medium. Separate correlations were computed for the create (Figure 5) and receive (Figure 6) groups, resulting in a critical correlation value of $r = +/- .404$, for $p<.05$ ($n = 24$). As seen in Figures 5 and 6, none of the correlations between test performance and note quality reached $r = +/- .404$, so the correlations were not statistically significant.

**Figure 5.** $R$ scores based on correlations between creating notes (longhand or laptop), question type, and note quality (word count and verbatim overlap) on test performance.
**Figure 6.** R scores based on correlations between receiving notes (longhand or laptop), question type, and note quality (word count and verbatim overlap) on test performance.

**Discussion**

Past research has discovered that taking longhand notes of a lecture improves test performance (Mueller & Oppenheimer, 2014). However, the current experiment is the first to test the effects of note-taking medium on reading an educational text. The study focused on the different effects for longhand notes versus typed notes of a textbook passage on test performance (conceptual and factual questions). In this study, four experimental groups read an educational passage after taking the pretest. The first experimental group, *longhand note takers*, created handwritten notes of the passage and had an opportunity to study their own notes during the allocated review time. The second experimental group, *laptop note takers*, typed notes that were later reviewed. In the third group, *longhand note receivers*, participants did not take notes, but received
notes from a yoked counterpart in the longhand note taking condition. The fourth experimental group, *laptop note receivers*, received notes from a yoked counterpart in the laptop note taking condition. After the participants read the passage, they participated in a math distractor task, reviewed notes, and then took a crossword puzzle distractor task, followed by the posttest with six conceptual and six factual questions.

An interaction was hypothesized between note-taking medium and question type for those who created notes, which is also what occurred in Mueller & Oppenheimer (2014). Longhand note takers were believed to have higher conceptual test performance, compared to subjects who typed or did not take any notes. However, the only difference between groups was the effect in which the laptop note receivers outperformed the creating notes conditions in factual questions (Figure 1). These results were inconsistent with previous studies by Mueller & Oppenheimer (2014) that suggested longhand note takers have better memory and understanding of newly learned information compared to laptop note takers. A second experiment was performed to possibly explain why there was no significant interaction between note-taking medium and test performance for learning from a textbook passage, since past research shows this interaction by learning from a video lecture. In the second study, there was a control group that read the same passage from the first study, but they did not create or review notes. This study tested for possible benefits of creating or receiving notes from the textbook passage. By comparing each of the 4 groups from Experiment 1 to the control group of Experiment 2, we can ascertain which Experiment 1 groups
benefitted from creating and/or reviewing notes, over and above the effects of reading the passage alone.

**Experiment 2**

The control group was used to test the benefits of creating or reviewing notes. This group followed a similar procedure to Experiment 1, but subjects did not create or receive notes, they were the *no-note* group. During the experiment, they took a pretest, read the textbook passage, spent 10 min on each distractor task, and then took a posttest (Table 2).

<table>
<thead>
<tr>
<th>Pretest</th>
<th>Reading Passage</th>
<th>Math Distractor Task</th>
<th>Crossword Distractor Task</th>
<th>Posttest</th>
</tr>
</thead>
</table>

**Table 2.** The study procedure used for testing the effects of not creating or receiving notes on test performance.

Past research discusses that creating and reviewing notes both carry benefits for test performance (Kiewra, 1985; Kiewra et al., 1991), and this no-note group did neither. It was hypothesized that this no-note group should perform worse than all other experimental groups that created or received notes. If this group has the lowest test performance, then it would mean that note taking and/or note receiving affects test performance.

**Method**

**Participants.** The participants were 12 undergraduate college students from the University of Colorado Boulder enrolled in General Psychology. They all
received partial course credit for their participation. The majority of subjects were college freshmen; all participants were over the age of 18.

**Ethics, consent, and permissions.** The University of Colorado Institutional Review Board, Protocol Number 16-0664, approved the use of human subjects.

**Design, materials, procedure.** Participants were in a control group that was tested on conceptual and factual questions during a pretest and posttest that were given before/after reading the passage. The dependent variable is the percent correct on the factual and conceptual questions.

Participants received the same testing pamphlet from Experiment 1, and followed a similar procedure. As shown in Diagram 2, the 12 subjects took a pretest, read the textbook passage, spent 10 min on the math distractor task, spent 10 min on the crossword puzzle distractor task, and then took a posttest. The subjects spent an extra 3 min on each distractor task, so that subjects had the same amount of time between reading the passage and taking the posttest as Experiment 1 (subjects had 6 min to review notes). The participants did not take any notes of the passage or review notes before the posttest.

**Results**

**Test performance.** Figure 7 shows test accuracy for Experiment 1 (solid bars) along with Experiment 2 (hatched bars). ANOVA on Experiment 2 test accuracy alone revealed the main effect of pretest and posttest scores as significant, $F(1,11) = 12.08$, $p = .005$. As shown in Figure 7, the posttest scores were consistently more accurate than the conceptual and factual pretest scores.
Further, there was a significant interaction between the test types (pre and post) and the question types (conceptual versus factual). Figure 8 (green bars) illustrates that test score improvement (posttest - pretest accuracy) was higher when the questions were factual, compared to conceptual, $F(1,11) = 5.66$, $p = .037$.

![Test Accuracy Diagram]

**Figure 7.** Mean proportion of correct test answers based on test type, note-taking medium (create or receive) and question type. Solid bars are from Experiment 1, and hatched bars are from Experiment 2.

Welch 2-sample $t$-tests, which correct for group differences in variability, were used to individually determine the effects of taking or receiving notes, or doing neither. Test improvement from each condition of Experiment 1 (each red or blue bar in Figure 8) was compared with the corresponding condition of
Experiment 2 (green bars in Figure 8), this was a between experimental comparison. As shown in Figure 8, laptop note receivers had significant test improvement on factual questions compared to subjects who did not create or receive notes, $t(21.76) = 3.64, p < 0.01$. All other $t$-tests had $p > .10$, and were not significant.

![Test Improvement](image)

**Figure 8.** Mean proportion of test improvement based on note-taking medium and question type. Blue and red bars are from Experiment 1, and green bars are from Experiment 2.

**Discussion**

In this study, there was a control group that took a pretest and then read an educational passage, without creating or reviewing any notes. After the participants read the passage, they participated in a math distractor task, a crossword puzzle distractor task, followed by the posttest with six conceptual and six factual questions. The purpose of this experiment was to compare the results
to the first experiment, in order to analyze the effects of note taking or note reviewing on test performance.

It was hypothesized that this control group would have the lowest test improvement, because they did not create or review notes. However, this outcome did not consistently occur. The no-note control subjects had a similar test improvement for all experimental conditions, except for a greater improvement on factual questions shown by subjects who received laptop notes. The direct implication of these findings is that creating notes (longhand or typed) of a reading passage did not affect test improvement. The majority of what subjects learned was from the reading passage, not from reviewing notes. This explains why the control group performed similarly to the other experimental groups, reading the passage was enough to improve test performance. Further, this experiment shared many consistencies with the first experiment, in regards to the differences between pre/post scores and the interaction of conceptual/factual questions with pre/post scores. These results were replicated, exemplifying a consistency within the data across experiments.

**General Discussion**

The overall purpose of these two experiments were to follow-up the Mueller and Oppenheimer (2014) studies to support that longhand notes are more beneficial to typed notes for learning (from a lecture). The current experiments examine if longhand note taking has the same benefits for learning information from a textbook passage, not a lecture (Mueller & Oppenheimer, 2014). The current experiment studies the effects of note-taking media and
reviewing notes for learning information from a textbook passage. The results are often compared to Mueller and Oppenheimer (2014) and Lalchandani and Healy (2016) studies, because the testing conditions and hypotheses are similar. Several other studies have found benefits for longhand note taking, except they commonly test subjects on random letters and information (Aragón-Mendizábal et al., 2016; James et al., 2016; Longcamp et al., 2008; Naka, 1998).

In Experiment 1, it was hypothesized that longhand note creators would have a higher test performance than subjects who typed or only received notes for review. There was no significant interaction involving note-taking medium on test scores (Figure 1). A possible explanation is that longhand participants did not have as much factual content to study compared to laptop note takers. Longhand note takers used more generative thinking, resulting in a smaller word count (less information) and verbatim overlap. Since there was a lower word count, subjects had less information to study in the review process. These results for no factual differences across longhand and laptop note takers were also seen in the first two Mueller and Oppenheimer (2014) studies and in Lalchandani and Healy’s (2016) study.

In the first study of Mueller and Oppenheimer (2014) the subjects took longhand or typed notes of TED talks, and they did not review their notes before taking a posttest. The second study was very similar, except subjects were instructed to not transcribe notes directly from the lecture, in hopes of decreasing verbatim overlap. Results of these two studies showed no difference between factual questions; however, the longhand subjects had higher test scores for
conceptual questions, compared to the laptop note takers. Lalchandani and Healy (2016) had the same results in higher conceptual question test scores for longhand subjects. However, in the current experiment there were no differences for conceptual questions across groups. A possible explanation is that longhand subjects might have been tired from creating generative longhand notes, and were too exhausted or unmotivated for additional complex thinking on the posttest. Another possible reason for different experiment results than Mueller and Oppenheimer (2014) is that the effects of learning and the encoding process differ between reading and listening to information (Vidal, 2011). Further, besides the difference in delivering the information, subjects learned different material (and had different posttests). These reasons might explain the differences between the current experiment and Mueller & Oppenheimer’s (2014) first two studies in test performance, along with Lalchandani and Healy’s (2016) study.

This study supported previous findings that laptop note takers have a higher verbatim overlap score than longhand note takers. As shown in Figure 3, laptop note takers had an average verbatim overlap score of 35% with the textbook, compared to longhand note takers who averaged 20%. Similar results were found with the three studies of Mueller and Oppenheimer (2014) and Lalchandani and Healy (2016) study. However, in regards to Mueller and Oppenheimer’s (2014) three studies, the average verbatim overlap score for laptop note takers was 12.6%, and for longhand note takers it was 6.63%. These averages are much lower than the current study, which is most likely because it is easier to copy verbatim when looking at the words in a passage, compared to
copying verbatim from hearing the words in a lecture. In the current experiment, subjects who received typed notes with a 35% verbatim overlap score, showed significantly better factual test improvement than those who did not receive or take notes. A possible reason is because they had the opportunity to study factual information that was very similar to the passage.

Further research found benefits for subjects creating typed notes, not receiving them. Bui et al., (2013) discovered that laptop note takers who typed verbatim notes from a lecture had a higher recall and short answer test performance than longhand note takers. However, these subjects were instructed to ignore their general study habits, and to transcribe notes; this instruction could explain why their results were inconsistent with Mueller and Oppenheimer's (2014) findings that longhand notes are more beneficial to test performance. Further, the higher test performance for laptops in Bui et al., (2013) study, might be explained because laptops result in more notes than longhand note taking, meaning that more information is recorded, even if it is processed shallowly. Further, this idea supports why the laptop note receivers had the highest factual test performance. They received notes with the highest verbatim overlap with the textbook passage, which resulted in them studying a lot of factual information before taking the posttest.

The results support the hypothesis that laptop note takers will have a higher word count of notes, compared to longhand note takers. As shown in Figure 4, laptop note takers averaged 436 words, compared to longhand note takers that averaged only 219 words. These results are consistent with Mueller
and Oppenheimer (2014) and Lalchandani and Healy (2016) in that laptop note takers had a higher word count.

It was hypothesized that subjects who received notes would perform worse than subjects who took notes, because creating notes involves deeper mental processing than passively reading a passage (Kiewra, 1985). However, this did not occur; all groups had higher test improvement for factual questions over conceptual questions, but the longhand note takers did not perform better than those in the receiving groups (Figure 2). Many factors could have influenced these results.

First, subjects who created notes (by hand or laptop) were told that another participant might read their anonymous notes. This could cause subjects to create notes that were not semantically personal, but more general for an outside reader to understand. Research suggests that memory improves for specific and detailed information when the learner relates it to themselves; this is called the Self Reference Effect (Serbun, Shih, & Gutchess, 2011). When subjects created notes, they might not have applied the information to themselves. This could have prevented them from remembering the essential information, which might explain their test performance not being higher than subjects who received notes.

Secondly, subjects who produced notes might not have chosen to review them seriously, because they wrote the notes shortly before. This could cause a decline in their posttest performance compared to subjects who received unfamiliar notes and possibly studied them in more depth. On the other hand,
subjects who did not create notes probably spent more time and energy studying the newly received notes, which could have improved their test scores. Thirdly, participants who created notes spent more time on the passage, which might have decreased their motivation to reexamine their notes or place a maximum effort in the posttest.

An unexpected finding occurred when laptop note receivers performed better on factual questions than longhand note receivers. As shown in Figure 2, participants who received notes from a laptop note taker had a higher test improvement than subjects who received notes from a longhand note taker. This is most likely due to the high levels of verbatim overlap that played as an advantage for the laptop note receivers. They had received notes with more information that could include factual content straight from the passage. This applies to real-world scenarios in which students receive notes from classmates. Students possibly benefit more from receiving verbatim notes with the learned material, compared to longhand notes that the note taker believed to be the key points.

In the second experiment, the primary goal was to research if subjects who did not create or review notes would have the lowest test improvement. As shown in Figure 8, numerically these subjects had the lowest improvement for conceptual questions, but less so for factual questions. The laptop note receivers performed significantly better on factual questions compared to the control group. There are a few possible explanations for no difference in factual improvement across the other experimental conditions (not including laptop note receivers).
First, the factual questions are like trivia statements, subjects typically need previous knowledge from the textbook passage to answer them correctly. This makes it more difficult for subjects to receive high factual scores on the pretest. On the other hand, conceptual questions use more common sense that subjects can apply for finding the right answer, which results in higher conceptual scores for the pretest. Since there were lower factual pretest scores, it leaves room for more factual improvement, which may explain the results.

Lalchandani and Healy (2016) had a surprising result in which subjects that received longhand notes performed worse on conceptual questions in comparison to subjects who received laptop notes. This result is similar to the findings in the current experiment. However, the subjects who received longhand notes performed worse on factual questions, not conceptual questions. These similar experimental results show an advantage in test performance for receiving laptop notes. Further, this might explain the results in the current experiment in which subjects who received laptop notes had the highest test improvement for factual questions.

The results from this study can benefit not only the educational field, but also society in general by explaining the practical applications of learning with technology. This study is unique in that it is the first to look at the effects of note-taking medium for learning by reading a textbook passage. Through comparing this study’s results with Mueller and Oppenheimer (2014), it appears that laptops are more harmful to test performance in note taking for a lecture, compared to reading a passage. It is less clear how technology impacts test performance from
reading passages because the results did not support past studies stating that longhand notes help improve conceptual understanding.

Past research clearly shows that longhand notes are more beneficial than typing for learning in lectures (Longcamp et al., 2008; Mueller & Oppenheimer 2014). To keep this traditional note taking method alive, it is important that teachers speak at an appropriate pace that gives students enough time to handwrite, so that they do not find the need for laptops to increase their note taking speeds (Dezure, Kaplan, & Deerman, 2001). Further, teachers can promote longhand note taking for lecture and reading passages by giving note outlines. These note handouts influence students to format newly learned information into their own words, which increases generative thinking (Stacy & Cain, 2015).

The present study had limitations that might have caused the note-taking medium to limit the effects on test performance. First, the sample size of 12 subjects per condition was small, which could have resulted in not seeing significant differences among groups; in Mueller and Oppenheimer (2014) study a sample size of 45 participants in each experimental condition were suggested by a power analysis of 95% (Lalchandani & Healy, 2016). Secondly, there was a very short time between subjects reading the passage (and possibly taking notes) and reviewing their notes. This could have resulted in subjects who created notes not to study them closely, because they just wrote them. To resolve this issue, subjects could have reviewed their notes and taken the posttest a couple of days later, so that benefits of note review might be clearer.
Thirdly, some of the conceptual questions on the pre/posttest could be answered through common knowledge, this might have caused less conceptual test improvement compared to factual improvement. Having conceptual questions is important, but future studies would improve if questions could not be answered by applying common sense.

Further research needs to compare how taking notes from a reading passage versus a lecture are affected by time allotted for the note taking process. Finding the proper techniques in blending generative longhand note taking, with the speed and efficiency of modern technology is essential in finding an appropriate medium between the controversies. It is important to properly integrate technology within classroom learning and studying process, so that students have the maximum potential to improve their memory and understanding of information.
References


Palekanda, V., & Venkataraman, R. (2012). Technology innovations in teaching and learning - amalgamation of technology and learning sciences has improved the
value creation for the learner and teacher. *5th International Conference of Education, Research and Innovation (Iceri 2012)*, 3416–3423.


Appendix

Test A:
1) A student would like to buy a Hyundai Elantra from a local dealer, but she thinks the payments will be too high. Marketing does not occur in this situation because:
   a) Two or more parties have unsatisfied needs
   b) There is no desire for the dealership to lower the car price
   c) There has been no assessment of consumer wants and needs
   d) One of the involved parties does not have the ability to satisfy its needs
   Answer: D/Conceptual.

2) From what country is the CEO of Chobani yogurt?
   a) United States of America
   b) Turkey
   c) Italy
   d) Greece
   Answer: B/Factual.

3) Steve just invented a new type of athletic shoe that has installed air-conditioning, to keep feet cool during exercise. However, he has been unsuccessful in selling his product. What is the best type of advice you give him?
   a) Advise him to make sure his prospective consumers have a want or need for this product.
   b) Advise him to keep trying, and eventually the idea will pick up, or he should try changing physical characteristics of the product.
   c) Explain to him the idea of profit, and how he can only make money off of the product if the shoes are expensive, because consumers believe that the quality always matches the cost of the product.
   d) Advise him that waiting is part of the process, as long as the product satisfies the consumer, then the product will gain popularity with time, word of mouth, and advertisement.
   Answer: A/Conceptual.

4) What is the source of the majority of new U.S.A. jobs?
   a) Renewable Energy
   b) Large businesses
   c) Small businesses
   d) Non-profit organizations
   Answer: C/Factual.
5) What is a major essence of success to marketing?
   a) Making enough money to pay off product debt.
   b) Having customers who consistently purchase a product, due to the unique value it provides.
   c) Word of mouth, where customers advertise product
   d) Having advertisement on TV, radio, and print that constantly remind consumers of the target product
Answer: B/Conceptual.

6) What is the definition of market?
   a) People with both desire and ability to buy a specific offering
   b) Having the best selling tools, product and design for customer convenience
   c) Having a cost and inventory that satisfies customer requests
   d) Customer and producer relationship, focusing on customer during sale
Answer: A/Factual.

7) A local candidate running for office would very much like to have your vote. She promises that she will “make the country better.” Because all candidates for public office say this, you decide not to vote for her. Marketing will NOT occur in this situation because:
   a) Marketing occurs in an economical sense, not political sense
   b) There is only one party involved in the decision
   c) There is no direct way for the parties to easily communicate
   d) Something to exchange is missing
Answer: D/Conceptual.

8) All of the following are marketing mix strategies that Chobani used to sell its yogurt EXCEPT:
   a) Develop new yogurt line product for kids
   b) Price Chobani yogurt slightly below price line of major competitors
   c) Use Facebook, YouTube, and other social media to promote Chobani yogurt
   d) Employ a spokesperson to open a yogurt bar
Answer: B/Factual.

9) A television advertisement shows several teenagers searching through the refrigerator for something to satisfy their hunger. The refrigerator offers the teenagers many alternatives- celery, bologna, cottage cheese, and box of Hot Pockets Bacon Cheddar Cheese Melt sandwiches. The ad, which shows teens happily eating the hot pockets, appeals to the consumers’ _____ for sustenance to satisfy their hunger and attempts to shape their ______ for the advertised product.
a) needs; wants
b) wants; needs
c) wants; preferences
d) preferences; wants
Answer: A/Conceptual.

10) Who was the leading person to launch Chobani?

a) Halis Unken
b) Hadith Uilliam
c) Halid Uduak
d) Hamdi Ulukaya
Answer: D/Factual.

11) The American Red Cross created TV advertisements to increase the likelihood of people to donate blood. Holly saw the ad, and went to her local Red Cross office to donate her blood. When Holly got home she was happy that she performed a good deed. Did an exchange occur in a marketing sense?

a) No, money must be exchanged for it to be considered a marketing activity.
b) No, the Red Cross did not provide Holly with a product, so no exchange occurred.
c) Yes, the blood was donated due to advertisement, which is an aspect of marketing.
d) Yes, Holly donated blood and through her good deed she felt happy.
Answer: D/Conceptual.

12) What company toothpaste was described in the passage?

a) Colgate
b) Dr. Care
c) Close-up
d) Arm and Hammer
Answer: B/Factual.

Test B:

1) What is the definition of customer value?

a) The unique combination of benefits received by targeted buyers
b) The customer satisfaction of feeling valued when culture and personality are thought about
c) The art and science of choosing valued customer groups and building relationships with them
d) The idea that consumers will not buy enough of the firm's products unless it undertakes a large-scale selling and promotion effort of value
Answer: A/Factual.

2) A church placed advertisements in its weekly bulletins to encourage its members to participate in the services by telling a positive short story about fellow members helping each other in times of need. Jack volunteered to share his story during the service, and afterwards he felt joyous. Was this an exchange in a marketing sense?

a) Yes, because the church ran an advertisement, a marketing activity.
b) No, the church did not provide Jack with a tangible product or service.
c) Yes, because sharing his story at a service was exchanged for a feeling of joy.
d) No, the church is a nonprofit organization, and these actions are expected without the expectation of “exchange”
Answer: C/Conceptual.

3) What was the major breakthrough for Chobani gaining public awareness?

a) Creating specific yogurt for children
b) Gaining popularity through social medial
c) Sponsorship in U.S. Olympic and Paralympic teams
d) Sponsorship in the Super Bowl
Answer: C/Factual.

4) What type of product traits will make consumers more likely to purchase the MacBook Air laptop?

a) The laptop costing $3,000, and .008 pounds lighter than standard version.
b) The laptop costing $1,200, and .008 pounds heavier than standard version
c) The laptop costing $500, but very short battery life.
d) The laptop costing $700, but is heavy.
Answer: B/Conceptual.

5) What were the three products used as examples about satisfying consumer needs?

a) Domino’s, Conoco slushy, IPhone 6
b) Cat collars, Dr. Watches, Coffee Maker
c) Books2go, IScam, Water blaster
d) Dr. Care, Terrafugia Transition, Pepsi Next
Answer: D/Factual.
6) Which is an important and reasonable aspect of satisfying a consumer's needs?
   a) Have product satisfy majority of consumers needs
   b) Have product satisfy people from different socioeconomic groups
   c) Have product concentrate on specific groups of people
   d) Have product make people happy
   Answer: C/Conceptual.

7) Mark Zuckerberg's launch of “TheFacebook.com” website became a huge success. Yet, more than half of all new businesses fail within _____ years of their launch.
   a) Two
   b) Three
   c) Four
   d) Five
   Answer: D/Factual.

8) Which option is an element of the marketing mix, and has a controllable factor?
   a) Walmart forcing toothpaste to be sold at a cheaper price, if the company wants their product in the Walmart stock.
   b) A company deciding to place their toothpaste in every Walmart throughout America.
   c) The CEO of a company telling the marketing department that there will be no sales on the toothpaste in Walmart
   d) The toothpaste not being on commercial TV, because Walmart’s marketing department cannot afford it.
   Answer: B/Conceptual.

9) All of the following are departments in a typical organization, except:
   a) Suppliers
   b) Finance
   c) Human Resources
   d) Information Systems
   Answer: A/Factual.

10) What is a potential legal or ethical issue that is central to marketing?
    a) Writing up contracts missing vital information
    b) Having producers falsely advertise product
    c) Not considering race, sex, or religion when industrializing product
    d) Marketing influencing what consumers buy to large extremes
Answer: D/Conceptual.

11) A soda company created a new brand that is described in the reading, what is it called?
   
   a) Mist Twist
   b) Pepsi Next
   c) Mountain Dew Baja Blast
   d) AMP energy

Answer: B/Factual.

12) Chose the incorrect statement:

   a) A person wants shelter
   b) A person needs food
   c) A person needs clothes
   d) A person wants the new Nike rebound shoes.

Answer: A/Conceptual.