Which modality results in superior recall for students: Handwriting, typing, or drawing?

Lindsay Richardson & Guy Lacroix

Carleton University, Ottawa | Canada

Abstract: One of the most common interests among cognitive psychologists is establishing ways to enhance human learning. An additional layer of complexity has been brought on by the rapid evolution of technology. Specifically, examining if the mechanisms involved in typing differ from those involved in handwriting. The literature concerning the implications of encoding modality on memory have been inconclusive. This present research examined whether encoding modality resulted in performance differences for word recall. Wammes et al.'s (2016) drawing versus handwriting methodology was utilized with the addition of a typing condition. The results replicated the *drawing effect*, whereby drawn words were better recalled than handwritten ones. Overall, the evidence did not suggest that the mechanisms involved in handwriting led to better free recall than those involved in typing. However, if the pen is indeed mightier than the keyboard (Mueller & Oppenheimer, 2014), then the effect is not explained by visual attention or sensorimotor action differences between modalities. Implications for education are discussed.

Keywords: drawing, handwriting, typing, recall, memory

journal of WRITING RESEARCH

Richardson, L., & Lacroix, G. (2024). Which modality results in superior recall for students: Handwriting, typing, or drawing?. *Journal of Writing Research, 15*(3), 519-540. DOI: 10.17239/jowr-2024.15.03.04

Contact: Lindsay Richardson, Carleton University, 1125 Colonel By Dr, Ottawa, ON K1S 5B6 Ottawa, Canada – lindsayrichardson@cunet.carleton.ca.

For decades, cognitive psychologists have sought to identify ways to promote more effective recall and these efforts have yielded many well-known memory enhancing phenomena. For instance, expanding on a concept (i.e., the elaboration effect; Craik & Tulving, 1975), reading aloud (i.e., the production effect; MacLeod et al., 2010), and creating something (i.e., the generation effect; Karpicke & Roediger, 2008) have all been shown to increase one's likelihood of remembering encoded information. More recent research has focused on the influence of motor movement on memory. This field, known as embodied cognition, examines the influence that the mind-body connection has on memory and learning (Stolz, 2015). It has produced many fascinating findings. For example, participants who perform actions or observe others doing them remember the actions more accurately than those who simply hear a verbal description (i.e., enactment effect; Madan & Singhal, 2012). Even movements that are unrelated to words can improve memory. In fact, Sullivan et al. (2018) found that the act of pulling words written on paper toward oneself rather than pushing them away increased the likelihood of recall. They dubbed this phenomenon the *mere ownership effect*, whereby information related to one's possessions is well-remembered (Cunningham et al., 2008).

Research on embodied cognition is directly relevant to determining whether the memory benefit offered by digital encoding strategies differ from those associated with analogue methods. This present research aims to explore the effect that encoding modality has on one's recall ability. Handwriting has been used as a means of encoding for centuries, but relatively recent technological advances have allowed people to use a keyboard instead. Ostensibly, the mechanisms involved in typing on a keyboard are not the same as those involved in handwriting. If typing is used as a method for encoding information, it makes sense that psychologists would be interested in whether it offers similar memorial benefits. Longcamp et al. (2005) explored this idea with a sample of pre-literate children aged 3 to 5. They were asked to either handwrite or type letters with feedback from the experimenter. Then, they completed a letter recognition task. Longcamp et al. found that the older children, but not the younger ones, recognized more letters if they had handwritten them instead of having typed them. This result supports the argument that typing and handwriting involve distinct sensorimotor processes. With handwriting, the relationship between movement and the shape of the letter (or word) is learned over time. Conversely, with typing, one merely needs to locate a key on a keyboard and tap it.

Even though handwriting might yield better memory for letters or strings of letters, the same might not be true for words. When semantics are applied to strings of letters, as they are with real lexicon words, the mechanism of encoding differs because the encoder can group the letters of the string into one meaningful unit (i.e., or chunk) of information. Thus, it is likely that literate individuals, as opposed to Longcamp et al.'s (2005) pre-literate children, encode real words as opposed to a series of letters from the string. So, when the modality for retrieval is not either handwriting or typing (e.g., verbal), the encoder should be able to rely on the meaning of the word to recall it. Still, Longcamp et al.'s research provided evidence to support the notion that the peripheral kinaesthetic signals that accompany handwriting and typing movements differ. This finding raises the question of how these modality differences evolve developmentally.

Investigating the development of each modality can elucidate the underlying cognitive and motor processes to help explain the mechanistic differences between the two. For instance, Kiefer et al. (2015) studied the effects of handwriting and typing training on preschool children's reading and writing performance. They found that handwriting training led to better outcomes, suggesting a strong link between motor action and perception in the context of literacy training (see also Bouriga & Olive, 2021). Studies by Grabowski (2010) and Bourdin & Fayol (1994) found that writing is more cognitively demanding than speaking, and that written language production requires more working memory than oral language production. These findings contribute to our understanding of the cognitive processes involved in handwriting and typing as well as their effects on learning and memory.

Mangen et al. (2015) also argued that there are physiological, cognitive, and ergonomic differences that make handwriting and typing distinctly different processes. Specifically, handwriting is a more kinaesthetic process that is governed by the spatial consistency involved in the motor action of shaping letters. This theoretical framework was the basis for their hypothesis that free recall and recognition of words would be superior when written by hand compared to both a screen and a mechanical keyboard. To this end, 36 participants completed a within-subjects encoding task, where they were asked to use a pen and paper, a mechanical keyboard, or a screen keyboard to record words played through headphones. After the encoding task, half of the participants were asked to take as much time as they needed to orally produce the words they remembered. The other half were presented words orally and asked to indicate whether the word was present in the encoding task. The results demonstrated no effect in the recognition group. However, free recall was better for words that had been handwritten than those that were typed (regardless of keyboard type).

Mangen et al. (2015) claimed that handwriting was associated with better free recall compared to typing because of the cognitive and sensorimotor differences between the modalities. They also argued that keyboarding can be considered less generative than handwriting because keystrokes do not need to be created by a typist; that is, the keys remain in the same position and the word dictates where the fingers must go to produce it on a page. Comparatively, when handwriting, one must generate each letter from scratch to form the words. Ultimately, Mangen et al. provided empirical evidence to suggest that humans' sensory and perceptual experience of logging information differs depending on whether the information is handwritten or typed and that this is partly explained by the kinaesthetic differences between modalities.

However, the literature suggests that it is quite difficult to produce a reliable modality effect, as seen with the number of mixed results found throughout the literature. This is especially true when they have been examined under more ecologically valid conditions (e.g., Bui et al., 2013; Morehead et al., 2019; Mueller & Oppenheimer, 2014). While researchers agree that there are mechanistic differences between handwriting and typing that could lead to performative differences when used for encoding (e.g., Aragón-Mendizábal et al., 2016; Bouriga & Olive, 2021; Longcamp et al., 2005; Mangen & Balsvik, 2016; Mangen et al., 2015; Smoker et al., 2009), it is not clear which modality leads to superior memory for the information encoded.

A salient example of these mixed results can be found in the literature that pertains to modality differences in university notetaking. It is unequivocally true that the act of taking notes during a lecture enhances university students' academic performance (Dunkel & Davy, 1989). Until recently, though, most of the notetaking literature pertained to the act of handwriting. Because it is now very common for students to bring their laptops into the classroom (Morehead et al., 2019), many of them choose to take lecture notes by typing them. This provides yet another implication for understanding how the mechanistic differences that exist between these two modalities affects memory. Students' in-class notetaking serves two main purposes: encoding the information and storing it for later review (see Kiewra ,1989 for a thorough review). If the mechanisms involved in one of the notetaking modalities enhances memory for the information, students would benefit from adopting that modality when taking notes in class. Unfortunately, answering this question empirically has been proven difficult because many variables come into play (e.g., review period, instructor differences, handwriting and typing proficiency, working memory capacity, academic support, organization, study habits, class attendance). Still, many researchers have investigated how lecture notetaking modality affects students' academic success (e.g., Blankenship, 2016; Bui et al., 2013; Fried, 2008; Gaudreau et al., 2014; Grace-Martin & Gay, 2001; Kay & Lauricella, 2011; Kodaira, 2017; Kraushaar & Novak, 2010; Manzi, et al., 2017; Morehead et al., 2019; Muller & Oppenheimer, 2014; Skolnick & Puzo, 2008; Sovern, 2013; Urry et al., 2021; Yamamoto, 2007).

Alas, this line of research has produced inconsistent results. For instance, Bui et al. (2013) found that students who took notes with the assistance of a keyboard outperformed their handwriting counterparts on a comprehension test. Mueller and Oppenheimer (2014) obtained results in the opposite direction. They found that participants retained more conceptual information when they took notes by hand compared to typing them, as would be predicted by the research suggesting that handwriting would lead to superior memory for information encoded compared to typing (e.g., Longcamp et al., 2005; Mangen et al., 2015). Nonetheless, Morehead et al. (2019) and Urry et al.'s (2021) exact replications of Mueller and Oppenheimer's experiment failed to generate a difference between the handwriting and typing conditions.

When investigating how students' typed and handwritten notes differed, Fiorella and Mayer (2017) found that typists tended to use more verbal strategies. They also found that some students opted to use maps and drawings when taking notes by hand and that this ability had an impact on their learning. Interestingly, Alesandrini's (1981) results showed that concepts were better recalled when drawn compared to paraphrased. More recently, Wammes et al. (2016) investigated the difference between handwritten transcription of words and the drawing of what they represent. Because the processes associated with visual and tactile perception are reciprocally connected (Mangen et al., 2015), it makes sense that the extended integration of visual attention and sensorimotor action involved in drawing would result in memory improvements over handwriting. In fact, Wammes et al. sought to demonstrate this empirically. As such, participants drew and handwrote words presented on a screen during an incidental learning task. After a three-minute distractor task, they were asked to orally recall as many words as they could in one minute in whatever order they chose. Wammes et al. found that the words that were encoded by drawing were better remembered than those that were transcribed with a pen and paper. Importantly, these results established that the integration of the multiple memory trace components involved in drawing (i.e., visual, semantic, and motor) improves memory.

To support this interpretation of *the drawing effect*, Wammes et al. (2016) empirically and systematically ruled out alternative accounts. For instance, a *levels of processing* account for the results could have been argued (Craik & Lockhart, 1972). Simply put, information that is encoded more deeply (e.g., creating cue cards) is better remembered than information encoded in a shallow way (e.g., re-watching lecture videos). To rule this potential explanation out, Wammes et al. added a feature-listing condition to their basic paradigm. This condition required participants to list the features of words. For instance, when given the word *dog*, participants might have listed features such as *four legs*, *barks*, *fur*. If a levels-of-processing account could explain the drawing effect, then listing should have been shown to be just as beneficial as drawing in terms of memory improvements over handwriting. The experiment did not generate this result, however. Drawing improved memory more than listing or handwriting. Fernandes et al. (2018) replicated this finding, whereby drawing was shown to promote memory improvements compared to listing. Ultimately, Wammes et al. similarly excluded a series of other plausible alternate explanations for *the drawing effect* (e.g., encoding time, visual imagery, picture superiority).

If the memory enhancement for drawn concepts can be attributed to the integration of multiple well-known encoding techniques (i.e., elaborative, motoric, and pictorial), then recall for drawn words should be superior to any one technique alone. This was exactly what Wammes et al. (2019) sought to demonstrate. To this end, they replicated Wammes et al.'s (2016) methodology and demonstrated that drawing was superior to four novel conditions: drawing without visual feedback, tracing, imagining, and viewing. This provided strong evidence to support Wammes et al.'s (2016) argument that drawing recruits multiple memory-enhancing encoding techniques, which results in superior performance to any single encoding technique. Ultimately, the literature has provided ample empirical support for the notion that drawing concepts promotes improved memory for information when compared to handwriting because of the integration of multiple memory trace components (e.g., Fernandes et al., 2018; Wammes et al., 2016; Wammes et al., 2019).

When comparing handwriting to typing, though, the act of handwriting promotes the integration of more memory trace components. So, taking Wammes et al.'s (2016) memory trace component explanation of *the drawing effect* combined with Mangen et al.'s (2015) sensorimotor and visual attention explanation of *the handwriting effect*, one might conclude that the act of handwriting, albeit inferior to drawing, is superior to typing when attempting to encode to-be-remembered information. In other words, even though handwriting does not involve as many encoding benefits as drawing, it might involve more than typing.

The goal of this present research was to build upon Mangen et al.'s (2015) research to examine whether the mechanistic differences between handwriting and typing impact memory for information encoded. As such, Wammes et al.'s (2016) paradigm was replicated with the addition of a typing condition. Moreover, the present research used some of Wammes et al.'s words in addition to others frequently found in academic texts (also see Roberts & Wammes, 2021). Across three experiments, 191 participants completed an encoding task, whereby they were asked to draw what the words represented or to transcribe them by either typing or writing. Experiments 1 and 3 employed a within-subjects design, whereas participants in Experiment 2 used a between-subject design. After a 3-minute distractor task, participants were surprised with a free recall task, where they were asked to recall as many words from the encoding task as they could in one minute. If Wammes et al.'s results are replicated, then words that are drawn should be the most likely to be recalled by participants. Moreover, if support for Mangen et al.'s *handwriting superiority effect* is found, then handwritten words should be recalled more than those that are typed.

1. Experiment 1

1.1 Method

Participants

Participants were 64 undergraduate students recruited from Carleton University's SONA system. They earned 1% bonus credit in their psychology courses as compensation for participation.

Materials

Target items. The 80-item word list is shown in Table 1. Forty words were selected from the Academic Word List, provided by the EAP Foundation (publicly available online) to be representative of words most used in academic texts. The other 40 words comprised a selection of the words used by Wammes et al. (2016). The words ranged in length from 3 to 12 (M = 6.66); in frequency from 379 to 326 758 (M = 28 233.10); and in number of syllables from 1 to 4 (M = 2.11). The English Lexicon Project's website was used to gain the lexical information on the words within the list (Balota et al., 2007).

Filler task. Following Wammes et al.'s (2016) methodology, the filler task comprised a continuous reaction time task (CRT). Sound files were constructed to represent low-, medium-, and high-pitched tones using Audacity 3.2 software (Crowder, 2015). Each tone was presented to participants for 500ms at frequencies of 350, 500, and 650 Hz for low-, medium-, and high-pitched tones, respectively.

Procedure

The recruitment announcement indicated that participants would be required to use Zoom along with their own writing utensil and notepad. Participants completed the experiment online and in a location of their choosing over a recorded Zoom call with one of two different experimenters. Participants were instructed to keep their webcams on for the duration of the experiment.

After providing informed consent, participants completed the encoding task. PowerPoint slides were created to display all instructions and stimuli. The instructions prompted participants to either handwrite, type, or draw the words shown on the screen. They were informed that the prompt "handwrite" meant to continuously write down the word in their notebook. Similarly, the prompt "type" meant to continuously type the word into the Zoom chat window. Additionally, they were informed that the prompt "draw" meant to take the allotted time to draw out what the word represented in their notebook.

Table 1. This 80-item word list was used in the encoding task of all three experiments and is presented in alphabetical order.

analyze	duck	indicative	shoe
assess	ear	individual	significant
assumption	elephant	interpret	similar
available	environment	involved	skirt
benefit	establish	issue	specific
concept	estimate	jacket	spider
consist	evidence	kettle	spoon
context	factors	method	stool
couch	flute	occur	stove
cow	fork	percentage	strawberry
create	formula	proceed	sweater
data	frog	process	theory
define	function	require	toaster
derive	giraffe	research	trumpet
desk	glove	response	turtle
distribution	grapes	ruler	variable
divided	guitar	sailboat	violin
doll	hammer	scissors	wagon
door	harp	screwdriver	whistle
drum	identity	sheep	wrench

Note. Words in this list are sorted alphabetically.

An Excel spreadsheet was used to randomize the display order of prompts and record the randomized order of target words. Participants were not informed of the upcoming memory task.

Participants completed two practice trials for each prompt to familiarize themselves with each of them. From the list of 80 words, 30 of them were randomly selected per participant such that 10 were randomly selected to belong to each of the three prompts. This was done so that each participant received a unique set of target words. For each participant, each of the 30 words was shown on the screen, one at a time. A cue card creating site was used to randomize and display target words and a mask "&". Participants were shown the target word for approximately 1s. Then, the mask replaced the word on the screen and participants had 40s to follow the prompt for that word. Next, a tone sounded to indicate that a new word would be displayed on the screen in 3s. Immediately following the encoding task (i.e., 30 randomized trials), participants were instructed to take a self-paced break before continuing to the next task. They were also instructed to put their notebooks out of sight and reach.

Subsequently, participants completed the CRT as a filler task. The experimenter's computer played a shuffled playlist of tones that included 1500ms of silence between tones. This was played through the participants' speakers via Zoom screenshare (audio only). Hence, there was no way to control for audio quality and volume. Participants orally classified each of the tones, in turn, as either "low", "medium", or "high". Each tone played for 500ms and participants had 1500ms to respond. Participants continued to classify tones for three minutes. Following the filler task, they were then be asked to freely recall as many words as they could in 60s in whatever order they choose. For this free recall task, participants spoke the words into their microphone and the Zoom recording was used to record their responses. In total, the experiment lasted approximately 35 minutes.

All methods were approved by the Carleton University Research Ethics Board-B (CUREB-B) at Carleton University, which is constituted and operates in compliance with the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS2).

1.2 Results

Participants' CRT accuracy was calculated (M = 0.64, SD = 0.26). However, the data for 27 participants were missing because they chose to take a self-paced break that lasted the given maximum of three minutes, which left them no time to attempt the CRT. Consequently, this metric was not used to determine participant compliance.

To ensure that results could not be attributed to experimenter differences, an independent samples t-test was conducted on words recalled. The results found no significant differences between participants run by Experimenter 1 (n = 33, M = 6.94, SD = 0.52) and Experimenter 2 (n = 31, M = 6.42, SD = 0.45), $t(62) = 0.75$, $p = .46$. Thus, participants' ability to recall words cannot be explained by the difference in experimenter.

The proportion of words recalled (M = .22, SD = .09) was calculated for each of the three prompts by dividing the number of words participants recalled for that prompt by the number of words they were shown for that same prompt. To investigate whether modality influenced the proportion of words recalled, a 1 x 3 Within-Subjects ANOVA was conducted on proportion of words recalled. The independent variable was Prompt, which comprised three levels: Draw (M = .34, SD = .18), Write (M = .16, SD = .14), and Type (M = .17, SD = .13). The results are presented in Figure 1 and they demonstrated a significant omnibus test, $F(2, 126) = 32.27$, $p < .001$, $\eta^2 = .34$. It was expected that this study would replicate the effects found by Wammes et al. (2016), whereby drawn words would be better recalled than handwritten ones. The results did replicate Wammes et al.'s *drawing effect*, $t(63) = 6.71$, $p < .001$, $d = 0.84$.

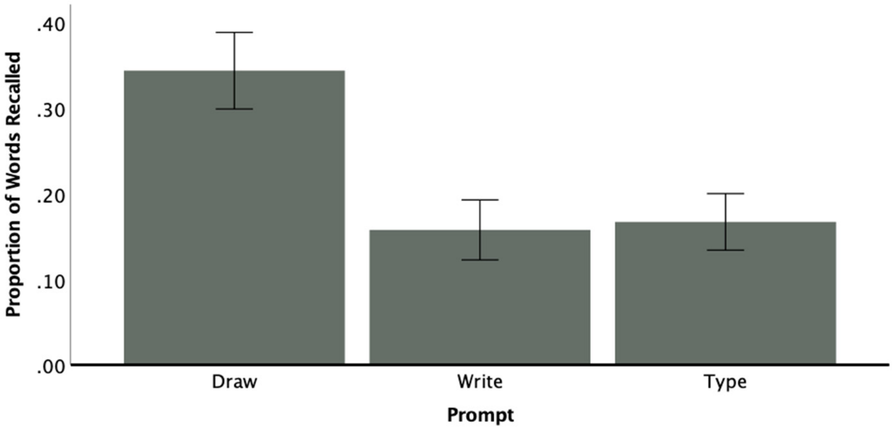

Figure 1: The proportion of words freely recalled is depicted and differentiates between whether they were drawn, written, or typed during the encoding task. Error bars represent 95% confidence intervals.

Expanding upon Wammes et al.'s results, this experiment demonstrated that drawn words were also better recalled than typed ones, $t(63) = 6.76$, $p < .001$, $d = 0.84$. Finally, it was also hypothesized that handwritten words would be better recalled than typed ones. Unfortunately, this hypothesis was not supported, $t(63) = 0.39$, $p = .35$, $d = 0.05$. To follow up on this null effect, a Bayesian one-way Repeated Measures ANOVA was conducted, which found strong evidence to support the null, BF = 0.10.

1.3 Discussion

The results from this present study have provided an independent replication of Wammes et al.'s (2016) very robust *drawing effect*. Free recall for drawn words was very clearly superior to that of written or typed words. It failed, however, to find a memory benefit of handwriting over typing. Nonetheless, it should be noted this inability to replicate Mangen et al.'s (2015) result might have stemmed from the drawn words distinctiveness as originally argued by Wammes et al. (2016, Experiment 7).

A distinctiveness effect is an enhancement of memory that is experienced for stimuli that are unique in some aspect (Restorff, 1933). Hunt (1995) argued that perceptual salience is not required for the effect to emerge. Rather, it is the differential attention elicited by the distinctive stimuli that yields it whether it is due to perceptual salience or contextual incongruity. With regard to the present research, it could be argued that the drawing condition enhanced memory for these words not because of any encoding benefit inherent to the task, but because drawing made them more distinctive relative to the words in the other conditions. In fact, Wammes et al. (2016) noted the possibility for the drawing effect to be accounted for, at least in part, by the distinctiveness effect. Hence, this possibility will be explored in Experiment 2 using a between-subjects design.

2. Experiment 2

If drawing certain words yielded a distinctiveness effect so powerful that it nearly precluded the retrieval of handwritten or typed words, then it might still be possible to generate a handwriting effect using a between-subject design. Experiment 2 was designed to examine this notion. More specifically, participants were randomly assigned to one of three encoding conditions: Draw, Write, or Type. If the distinctiveness of drawing prevented an effect between handwriting and typing conditions to emerge in Experiment 1, then participants who handwrote during Experiment 2's encoding task should outperform those who typed on the surprise free recall task.

2.1 Method

Participants

Ninety undergraduate students from Carleton University's SONA system participated in the Draw (n = 30), Write (n = 30), or Type (n = 30) condition. They each earned 1% bonus credit in their psychology courses as compensation for participation. Exclusion criteria ensured that no participants had previously completed Experiment 1.

Materials and Procedure

The materials and procedure used in this experiment replicated those of Experiment 1. The only difference was that this present experiment utilized a between-subjects design. Therefore, participants only completed one of the three prompt conditions: Draw, Write, or Type to which 30 randomly-selected words from Table 1 were assigned. Once more, this encoding phase was followed by the CRT used as a filler task and a surprise test of free recall.

2.2 Results

As was done in Experiment 1, participants' CRT accuracy was calculated (M = 0.66, SD = 0.26). This time, the data for five participants were missing due to time constraints during the retention interval. Still, this metric was excluded from determining compliance.

Again, an independent samples t-test was conducted on words recalled ensuring that results could not be attributed to experimenter differences. The results found no significant differences between participants run by Experimenter 1 (n = 45, M = 7.98, SD = 0.53) and Experimenter 2 (n = 45, M = 3.29, SD = 0.49), t(88) = 0.71, p = .48. Thus, participants' ability to recall words cannot be explained by the difference in experimenter.

The proportion of words recalled (M = .27, SD = .11) was calculated for all participants. To investigate whether modality influenced the proportion of words recalled, a 1 x 3 Between-Subjects ANOVA was conducted on proportion of words recalled. The independent variable was Prompt, which comprised three levels: Draw (n = 30, M = .35, SD = .11), Write (n = 30, M = .26, SD = .11), and Type (n = 30, M = .21, SD = .08). The results are presented in Figure 2 and they yielded a significant omnibus test, F(2, 87) = 14.94, p < .001, η^2 = .27. Just like in Experiment 1, drawn words were better recalled than handwritten ones, t(58) = 3.17, p = .001, d = 0.81. This replicated Wammes et al.'s *drawing effect*. Also, replicating the results found in Experiment 1, drawn words were better recalled than typed ones, t(58) = 5.57, p < .001, d = 1.44. While the results of the within-subjects Experiment 1 were unable to support the hypothesis that handwritten words are more likely to be recalled than typed ones, the results of this between-subjects Experiment 2 did. Specifically, participants in the Write condition remembered more words than those in the Type condition, t(58) = 2.15, p = .02, d = 0.55.

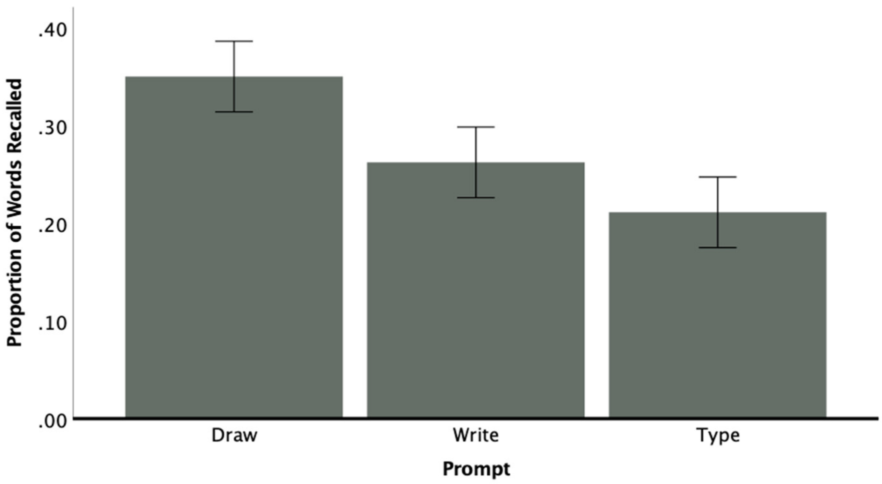

Figure 2: The proportion of words freely recalled is depicted and differentiates between participants who completed the drawing, writing, or typing encoding task. Error bars represent 95% confidence intervals.

However, a Bayesian independent sample test—conducted between the Write and Type conditions—found anecdotal evidence to support a null effect, BF = 0.67.

2.3 Discussion

The contrast between the results of Experiments 1 and 2 was informative. While Experiment 1 was unable to generate evidence in support of a *handwriting superiority effect*, the change from a within-subject to a between-subjects design in Experiment 2 did replicate Mangen et al.'s (2015) results. However, while the Bayesian test provided merely anecdotal evidence in support of a null effect, it was unable to provide evidence of a difference between conditions. Therefore, one last more stringent experiment was conducted to establish the existence of a handwriting superiority effect. A within-subject design including only writing and typing conditions was conducted.

3. Experiment 3

3.1 Method

Participants

Participants were 37 undergraduate students recruited from Carleton University's SONA system. They earned 1% bonus credit in their psychology courses as compensation for participation. Exclusion criteria ensured that no participants had previously completed Experiment 1 or 2.

Materials and Procedure

The materials and procedure used in this experiment replicated those of Experiment 1. However, participants did not experience a drawing condition. As such, to ensure all participants received the same number of words in total, 15 randomly selected words were written and 15 were typed. Moreover, the research assistants from Experiments 1 and 2 had since graduated and 3 new research assistants were hired to be experimenters for this experiment.

3.2 Results

As was done in Experiment 1 and 2, participants' CRT accuracy was calculated (M = 0.76, SD = 0.09), and eight participants' data were missing due to the retention interval time constraints. This metric was not used to determine participant compliance.

Once more, potential differences in words recalled between experimenters was examined. Since there were three experimenters, a 1 x 3 Between-Subjects ANOVA was conducted to determine if they had influenced the results. Fortunately, no significant differences between participants run by Experimenter 1 ($n = 12$, $M = 7.42$, $SD = 2.91$), Experimenter 2 ($n = 12$, $M = 7.25$, $SD = 2.00$), and Experimenter 3 ($n = 13$, $M = 7.92$, $SD = 3.75$), $F(2, 34) = 0.17$, $p = .84$. Thus, participants' ability to recall words cannot be explained by the difference in experimenter.

The proportion of words recalled ($M = .25$, $SD = .10$) was calculated for all participants. To investigate whether modality influenced the proportion of words recalled, a 1 x 2 Within-Subjects ANOVA was conducted on proportion of words recalled. The independent variable was Prompt, which comprised two levels: Write ($M = .25$, $SD = .17$) and Type ($M = .25$, $SD = .15$). As seen in Figure 3, the conditions generated almost identical recall. Participants were just as likely to remember words they had handwritten compared to those they had typed during encoding, $F(1, 36) = 0.02$, $p = .90$, $\eta^2 < .001$. To examine an absence of effect, a Bayesian one-way Repeated Measures ANOVA was conducted with the dependent variables proportion of written words recalled and proportion of typed words recalled. It found strong evidence to support the null effect, BF = 0.12.

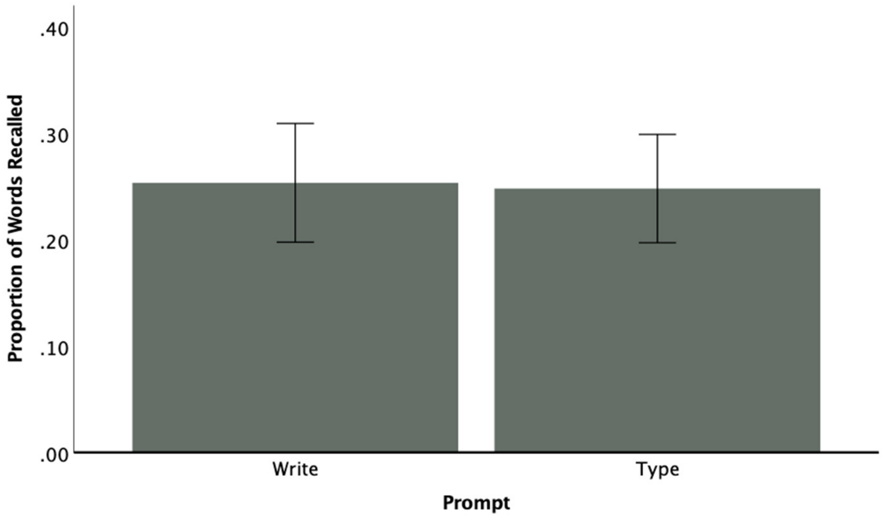

Figure 3: The proportion of words freely recalled is depicted and differentiates between whether they were written or typed during the encoding task. Error bars represent 95% confidence intervals.

3.3 Discussion

Experiment 3 demonstrated no effect of handwriting or typing on the memory of incidentally learned information. Moreover, it is unlikely that the null effect is explained by lack of power. The only reasonable conclusion that can be drawn from this experiment is that there is no difference in performance between the act of moving a pen on paper and making keystrokes when transcribing words before a surprise free recall task.

4. Differences Between Experiments

Across the three Experiments, 191 participants completed both the encoding and the free recall task. To investigate potential differences between experiments, a 1 x 3 Between-Subjects ANOVA was conducted on Words Recalled (M = 7.58, SD = 3.18). The results were found to be significant, $F(2, 188) = 4.58$, $p = .01$. Post hoc comparisons demonstrated that the difference existed between Experiment 1 and 2, $t(152) = 2.98$, $p = .003$. Participants in Experiment 1 recalled fewer words (M = 6.69, SD = 2.77) compared to those in Experiment 2 (M = 8.23, SD = 3.42). A Bayesian independent sample test found strong evidence for a null effect, however, BF = 0.12.

5. The Effect of Word Concreteness

One of the key differences between this present research and that of Wammes et al. (2016) was the use of abstract or unimageable words. It is possible that the words' concreteness might have impacted participants' ability to recall them differentially depending on condition. Thus, word concreteness was examined to determine whether it impacted word recall. As such, Brysbaert et al.'s (2014) word concreteness ratings database was used to collect concreteness ratings for the words used in this present research. These ratings matched our expectation based on the source of the words. When a median split was conducted, words with concreteness ratings below the median were those from academic texts, whereas words with ratings above the median were those from Wammes et al.'s materials selected to facilitate drawing. The likelihood of recall (i.e., the number of times a word was recalled as a function of how many times it was shown) was found to be correlated with concreteness ratings, $r = .38$, $p > .001$. Moreover, abstract words (M = .20, SD = .09) were less likely to be recalled than concrete ones (M = .28, SD = .11), $t(78) = 3.53$, $p < .001$.

To investigate this issue further, the number of concrete words seen relative to abstract words was calculated for each participant. Numbers greater than 1 indicated that participants received more abstract words and those less than 1 indicated that they received more concrete words. The correlation between this Concreteness Ratio (M = 1.13, SD = 0.36) and proportion of words recalled was not significant, $r = .01$, $p = .77$. Moreover, an independent samples t-test found that the proportion of words recalled did not differ between participants who saw more concrete words (M = .24, SD = .10) and those who saw more abstract words (M = .26, SD = .11), $t(189) = 1.41$, $p = .16$. Therefore, even though the concreteness of a word can influence the likelihood of recall, this did not impact participants' overall performance on the word recall task.

6. General Discussion

The goal of this research was to determine whether the visual attention and sensorimotor processes that distinguish writing and typing were sufficient to generate a handwriting superiority effect, which could be explained by embodied cognition. Theories of embodied cognition posit that cognitive processes, such as memory, are closely tied to the physical actions and experiences, such as the kinaesthetic and spatial aspects involved in handwriting (e.g., Longcamp et al., 2005; Mangen et al., 2015). It was hypothesized that if the mechanisms involved in handwriting—a more kinaesthetic process governed by the spatial consistency of shaping letters—are more conducive to memory compared to those involved in typing, then participants would recall more words that were handwritten compared to those that they had typed during an encoding task. Following Wammes et al.'s (2016) general methodology, participants in Experiment 1 drew, wrote, and typed words that were presented on a screen, during an encoding task. After a three-minute filler task, they were surprised with a one-minute free recall task. The results demonstrated that participants who typed words performed just as well on the free recall task as those who wrote them by hand. Interestingly, a between-subjects version of the exact same experiment (i.e., Experiment 2) yielded a significant difference between writing and typing groups. It provided evidence to suggest that handwriting might be more conducive to memory. Given the mixed results from the first two experiments, Experiment 3 was conducted to examine whether a distinctiveness effect could explain the inability for Experiment 1 to demonstrate a *handwriting effect*. More specifically, it was argued that the distinctiveness of drawn words might have reduced the experiment's potential to reveal a difference between the handwriting and typing conditions. To this end, participants in Experiment 3 only encoded words by writing or typing them.

The results of Experiment 3 failed to demonstrate an effect of modality on memory. In other words, participants were just as likely to recall words they had handwritten compared to those they had typed during incidental learning. Collectively, this series of experiments failed to replicate the results of Mangen et al. (2015). However, the one that most closely resembled that of Mangen et al. (i.e., Experiment 2) did replicate their findings. In Mangen et al.'s experiment, the presentation of condition was blocked, and thus, participants completed one encoding type at a time. In our within-subjects experiments (i.e., Experiment 1 & 3), we did not use a blocked presentation of the different encoding conditions. However, the results of our Experiment 2, which used a between-subjects design, successfully replicated Mangen et al.'s results. It might be that a handwriting effect is only achievable when participants are not switching encoding strategies within an experiment.

Thus, these considerations beg the question: Why would a between-subjects design result in a handwriting effect, whereas a within-subjects design would not?

Since participants were generally able to recall fewer words in Experiment 1 compared to Experiment 2, one possible explanation is that the inclusion of three different encoding prompts generated a cognitive load that was difficult to manage for the participants. At least one other example of this exists within the literature, albeit the procedure and task are not directly comparable. That is, Bouriga and Olive (2021) found that typing takes up more cognitive resources than handwriting, but they did not investigate those involved in drawing. Moreover, drawing is certainly a cognitively demanding elaborative process (Jonker et al., 2019), so it might have added to task's complexity too. This does not explain the lack of effect found in Experiment 3, however, since there was no drawing condition and the words recalled were not significantly different from the other experiments. Future research might help clarify the issue, but in comparison, the drawing effect is certainly not vulnerable to changes in experimental design (see also Wammes et al., 2016).

Still, this research provided more evidence to support the notion that drawn words are better recalled than handwritten ones. Expanding upon this result, the present research also demonstrated that drawn words are better recalled than typed ones. Moreover, it added evidence to support the notion that the drawing effect is not sensitive to word concreteness (Roberts & Wammes, 2021). Our experiments used a wider variety of words than those included in most work on the drawing effect (e.g., Fernandes et al., 2018; Wammes et al., 2016; Wammes et al., 2019). It included more abstract expressions found in academic texts in addition to words that were easily imageable. The results demonstrated that the concreteness of words was related to the likelihood of a word being recalled. Because participants' word lists were randomized, it was possible for them to have received a different number of concrete or abstract words. This was found to not have impacted participants' ability to recall the words, however.

6.1 Limitations and Future Directions

While these experiments offer meaningful insights into the effect of different encoding modalities on word recall, this present research is not without its limitations. Notably COVID-19 restrictions necessitated online testing, which reduced our ability to control the participants' compliance with experimental instructions and prevented us from using the CRT filler task to measure it. Moreover, we did not anticipate individual differences based on participants' demographics (e.g., age, gender, ethnicity, first language, learning differences) and, as such, this information was not collected. This resulted in our inability to examine whether demographic differences could explain the results and reduced its external validity in this regard. Finally, no stratification of the words was conducted prior to randomization to ensure that participants received an equal number of concrete or abstract words. Thus, it is impossible to say with certainty that word concreteness did not affect the results. Future research should aim to intentionally manipulate word concreteness to examine its impact on the relationship between modality and word recall.

6.2 Educational Implications

Across three experiments, we explored participants' ability to recall lists of unrelated words that they had encoded using drawing, handwriting, and typing. The drawing effect proved to be extremely robust, but unfortunately, the evidence to support a handwriting effect was not particularly convincing. Thus, if there is any benefit to taking handwritten notes over typed notes in the classroom, then we would argue that it probably does not stem from the cognitive and sensorimotor differences between these modalities (see Mangen et al., 2015).

Nonetheless, our experiments do not allow us to exclude other reasons why handwriting notes might enhance students' encoding more than typing. When students take notes under more ecologically valid conditions, they are recording lectures; that is, rich narratives designed to impart expertise in a given academic domain. Thus, in that context, other factors might lead handwritten notes to yield better encoding than typed ones. For instance, Muller and Oppenheimer (2014) argued that handwriting forces students to use "deeper processing" when taking notes more than typing because the speed at which lectures are given require them to paraphrase the information if their handwriting is to keep up. Typing unfolds more quickly and allows for notes to be taken verbatim, which was argued to yield "shallow processing." Unfortunately, while Muller and Oppenheimer's results initially supported their claim, more recent research has failed to replicate them (Morehead et al., 2019; Urry et al., 2021). In fact, it can be argued that this line of research has led educators to give more attention to the role of external storage (Kiewra, 1989). If students take notes to build studying materials in preparation for examinations, then the more complete notes afforded by the fastest notetaking method (i.e., typing) might be preferred. Yet, this possible advantage must be weighed in relation to the finding that laptops in the classroom tend to generate many distractions (e.g., access to social media and disruptive notifications) that compete with the lecture for students' attention (Fried, 2008; Gaudreau et al., 2014). Thus, the issue is complex, and students ought to consider the advantages and disadvantages to both notetaking modalities. Nevertheless, if handwriting can be shown to generate superior encoding under certain conditions, the explanation will be at some higher level of cognitive processing like Muller and Oppenheimer's proposed depth of processing proposal.

Finally, this research and the growing literature supporting the drawing effect (Fernandes et al., 2018; Jonker et al., 2019; Roberts & Wammes, 2021; Wammes et al, 2016; Wammes et al., 2019) do lead to several suggestions. In the context of lecture notetaking, it might encourage students and educators to explore if a modern version of shorthand-like notetaking: "visual notetaking", whereby symbols are used to replace words and phrases, could increase students' academic success (Tutt, 2021). The multiple memory traces involved in visual notetaking could enhance memory and ultimately lead to better learning. Minimally, when the content allows, drawing certain concepts presented in class—colloquially referred to as "Sketchnoting"—might yield better encoding than writing them. Importantly, like concept mapping, this practice would not thwart the external storage benefit of academic notetaking.

References

Alesandrini, K. L. (1981). Pictorial–verbal and analytic–holistic learning strategies in science learning. *Journal of Educational Psychology, 73*(3), 358–368. https://doi.org/10.1037/0022-0663.73.3.358

Aragón-Mendizábal, E., Delgado-Casas, C., Navarro-Guzmán, J.-I., Menacho-Jiménez, I., & Romero-Oliva, M.-F. (2016). A comparative study of handwriting and computer typing in note-taking by university students. *Comunicar (Huelva, Spain)*, *24*(48), 101–107. https://doi.org/10.3916/C48-2016-10

Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*, 445-459.
https://doi.org/10.3758/BF03193014

Blankenship, M. W. (2016). *Comparing note taking and test performance in methods and modes of note taking conditions*. ProQuest Dissertations Publishing. Retrieved from proxy.library.carleton.ca/login?url=https://search-proquest-com.proxy.library.carleton.ca/docview/1905877115?accountid=9894

Bouriga, S. & Olive T. (2021). Is typewriting more resources-demanding than handwriting in undergraduate students? *Reading and Writing, 34*, 2227–2255. https://doi.org/10.1007/s11145-021-10137-6

Bourdin, B., & Fayol, M. (1994). Is Written Language Production More Difficult Than Oral Language Production? A Working Memory Approach. *International Journal of Psychology, 29*, 591-620. DOI: 10.1080/00207599408248175.

Brysbaert, M., Warriner, A.B. & Kuperman, V. Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46*, 904–911 (2014). https://doi.org/10.3758/s13428-013-0403-5

Bui, D. C., Myerson, J., & Hale, S. (2013). Notetaking with computers: Exploring alternative strategies for improved recall. *Journal of Educational Psychology, 105*(2), 299-309. https://doi.org/10.1037/a0030367

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior, 11*(6), 671-684. https://doi.org/10.1016/S0022-5371(72)80001-X

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268

Crowder, M. (2015). Audacity (Version 3.2). [Computer software]. https://www.audacityteam.org/.

Cunningham, S. J., Turk, D. J., Macdonald, L. M., & Neil Macrae, C. (2008). Yours or mine? Ownership and memory. *Consciousness and Cognition, 17*(1), 312–318. https://doi.org/10.1016/j.concog.2007.04.003

Dunkel, P., & Davy, S. (1989). The heuristic of lecture notetaking: Perceptions of American & international students regarding the value & practice of notetaking. *English for Specific Purposes, 8*, 33-50. https://doi.org/10.1016/0889-4906(89)90005-7

Fernandes, M. A., Wammes, J. D., & Meade, M. E. (2018). The Surprisingly Powerful Influence of Drawing on Memory. *Current Directions in Psychological Science*, *27*(5), 302–308. https://doi.org/10.1177/0963721418755385

Fiorella, L., & Mayer, R. E. (2017). Spontaneous spatial strategy use in learning from scientific text. *Contemporary Educational Psychology, 49,* 66–79. https://doi.org/10.1016/j.cedpsych.2017.01.002.

Fried, C. B. (2008). In-class laptop use and its effects on student learning. *Computers & Education, 50*, 906-914.
https://doi.org/10.1016/j.compedu.2006.09.006

Gaudreau, P., Miranda, D., & Gareau, A. (2014). Canadian university students in wireless classrooms: What do they do on their laptops and does it really matter? *Computers & Education, 70*, 245-255. https://doi.org/10.1016/j.compedu.2013.08.019

Grabowski, J. (2010). Speaking, writing, and memory span in children: Output modality affects cognitive performance. *International journal of psychology, 45,* 28-39. DOI:10.1080/00207590902914051

Grace-Martin, M., & Gay, G. (2001). Web Browsing, Mobile Computing and Academic Performance. *Educational Technology & Society*, *4*(3), 95–107.

Hunt, R. R. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review, 2*, 105–112. https://doi.org/10.3758/BF03214414

Jonker, T. R., Wammes, J. D., & MacLeod, C. M. (2019). Drawing Enhances Item Information but Undermines Sequence Information in Memory. *Journal of Experimental Psychology, 45*(4), 689-699. https://doi.org/10.1037/xlm0000610

Karpicke, J. D., & Roediger, H. L., 3rd (2008). The Critical Importance of Retrieval for Learning. *Science, 319*(5865), 966–968. https://doi.org/10.1126/science.1152408

Kay, R., & Lauricella, S. (2011). Exploring the Benefits and Challenges of Using Laptop Computers in Higher Education Classrooms: A Formative Analysis. Canadian *Journal of Learning and Technology, 37*(1).
https://doi.org/10.21432/T2S598

Kiefer, M., Schuler, S., Mayer, C., Trumpp, N. M., Hille, K., & Sachse, S. (2015). Handwriting or Typewriting? The Influence of Pen- or Keyboard-Based Writing Training on Reading and Writing Performance in Preschool Children. *Advances in cognitive psychology*, *11*(4), 136–146. https://doi.org/10.5709/acp-0178-7

Kiewra, K. A. (1989). A review of note-taking: The encoding-storage paradigm and beyond. *Educational Psychology Review, 1*, 147-172. https://doi.org/10.1007/BF01326640

Kodaira, Y. (2017). *An Investigation of Cognitive Processes Associated with Notetaking and Notes-Review*. ProQuest Dissertations Publishing. Retrieved from http://proxy.library.carleton.ca/login?url=https://search-proquest-com.proxy.library.carleton.ca/docview/1896530052?accountid=9894

Kraushaar, J. M., & Novak, D. C. (2010). Examining the Affects of Student Multitasking with Laptops during the Lecture. *Journal of Information Systems Education*, *21*(2), 241-251. https://proxy.library.carleton.ca/login?qurl=https%3A%2F%2Fwww.proquest.com%2Fscholarly-journals%2Fexamining-affects-student-multitasking-with%2Fdocview%2F745597375%2Fse-2%3Faccountid%3D9894

Longcamp, M., Zerbato-Poudou, M.T., & Velay, J.L. (2005). The influence of writing practice on letter recognition in preschool children: A comparison between handwriting and typing. *Acta Psychologica, 119*(1) 67-79. https://doi.org/10.1016/j.actpsy.2004.10.019

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of experimental psychology: Learning, Memory, and Cognition*, *36*(3), 671–685. https://doi.org/10.1037/a0018785

Madan, C. R., & Singhal, A. (2012). Using actions to enhance memory: effects of enactment, gestures, and exercise on human memory. *Frontiers in Psychology*, *19*(3), 507. https://doi.org/10.3389/fpsyg.2012.00507

Mangen, A., Anda, L. G., Oxborough, G. H., & Brønnick, K. (2015). Handwriting versus keyboard writing: Effect on word recall. *Journal of Writing Research, 7*(2), 227-247. https://doi.org/10.17239/jowr-2015.07.02.1

Mangen, A., & Balsvik, L. (2016). Pen or keyboard in beginning writing instruction? Some perspectives from embodied cognition. *Trends in Neuroscience and Education, 5*(3), 99–106. https://doi.org/10.1016/j.tine.2016.06.003

Manzi, A., Martinez, S., & Durmysheva, Y. (2017). Cognitive correlates of lecture notetaking: Handwriting speed and attention. *North American Journal of Psychology, 19*(1), 195-217. Retrieved from http://proxy.library.carleton.ca/login?url=https://search-proquest-com.proxy.library.carleton.ca /docview/1878904220?accountid=9894

Mueller, P. A., & Oppenheimer, D. M. (2014). The Pen Is Mightier Than the Keyboard: Advantages of Longhand Over Laptop Note Taking. *Psychological Science*, *25*(6), 1159–1168. https://doi-org.proxy.library.carleton.ca/10.1177/0956797614524581

Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from https://support.pstnet.com/.

Roberts, B.R., Wammes, J.D. (2021). Drawing and memory: Using visual production to alleviate concreteness effects. *Psychonomic Bulletin & Review, 28*, 259–267 (2021). doi.org/10.3758/s13423-020-01804-w

Restorff, H. V. (1933). Ueber die Wirkung von Bereichsbildungen im Spurenfeld. Analyse von Vorgängen im Spurenfeld. I. Von W. Köhler und H. v. Restorff [On the effect of field formations in the trace field. Analysis of processes in the trace field. I. By W. Kohler and H. v. Restorff]. *Psychologische Forschung, 18*, 299-342. https://doi.org/10.1007/BF02409636

Skolnick, R., & Puzo, M. (2008). Utilization of laptop computers in the school of business classroom. *Academy of Educational Leadership Journal, 12*, 1-10.

Smoker, T. J., Murphy, C. E., & Rockwell, A. K. (2009). Comparing Memory for Handwriting versus Typing. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 53*(22), 1744–1747. https://doi.org/10.1177/154193120905302218

Sovern, J. (2013). Law Student Laptop Use During Class for Non- Class Purposes: Temptation v. Incentives. *University of Louisville Law Review, 51*, 483-517. DOI:10.2139/ssrn.1805107

Stolz, S. A. (2015). Embodied Learning. *Educational Philosophy and Theory*, *47*(5), 474–487. https://doi.org/10.1080/00131857.2013.879694

Sullivan, J. V., Potvin, J. M., & Christmas, S. D. (2018). Mere ownership of memory: motor manipulation during encoding affects memory for words. *Cognitive Processing, 19*(3), 429–434. https://doi.org/10.1007/s10339-018-0860-y

Tutt, P. (2021, April 9). *How—and Why—to Introduce Visual Note-Taking to Your Students. Edutopia*. edutopia.org/article/how-and-why-introduce-visual-note-taking-your-students

Urry, H. L., Crittle, C. S., Floerke, V. A., Leonard, M. Z., Perry, C. S., Akdilek, N., Albert, E. R., Block, A. J., Bollinger, C. A., Bowers, E. M., Brody, R. S., Burk, K. C., Burnstein, A., Chan, A. K., Chan, P. C., Chang, L. J., Chen, E., Chiarawongse, C. P., Chin, G., … Zarrow, J. E. (2021). Don't Ditch the Laptop Just Yet: A Direct Replication of Mueller and Oppenheimer's (2014) Study 1 Plus Mini Meta-Analyses Across Similar Studies. *Psychological Science, 32*(3), 326-339.

http://dx.doi.org/10.1177/0956797620965541

Wammes, J. D., Jonker, T. A., & Fernandes, M. A. (2019). Drawing improves memory: The importance of multimodal encoding context. *Cognition, 191*(2019), 1–9. https://doi.org/10.1016/j.cognition.2019.04.024

Wammes, J. D., Meade, M. E., & Fernandes, M. A. (2016). The drawing effect: Evidence for reliable and robust memory benefits in free recall. *The Quarterly Journal of Experimental Psychology, 69*(9), 1752-1776. http://dx.doi.org.proxy.library.carleton.ca/10.1080/17470218.2015.1094494

Yamamoto, K. (2007). Banning Laptops in the Classroom: Is it Worth the Hassles? *Journal of Legal Education*, *57*(4), 477–520.