Elucidating the Cognitive Processes Involved in the Note-Taking Effect

by

Lakshmi Arjan Lalchandani

B.A., University of Colorado Boulder, 2011

M.A., University of Colorado Boulder, 2016

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Psychology and Neuroscience

and the Institute of Cognitive Science

2018

This thesis entitled:
Elucidating the Cognitive Processes Involved in the Note-Taking Effect
written by Lakshmi Arjan Lalchandani
has been approved for the Department of Psychology and Neuroscience
and the Institute of Cognitive Science

____
Dr. Alice F. Healy

____
Dr. Eliana Colunga

____
Dr. Michael C. Mozer

____
Dr. Michael C. Stallings

____
Dr. Tamara Sumner

Date_____________

The final copy of this thesis has been examined by the signatories, and we
Find that both the content and the form meet acceptable presentation standards
Of scholarly work in the above mentioned discipline

HRC IRB protocol # 18 - 0308

Lalchandani, Lakshmi Arjan (Ph.D. Psychology and Neuroscience; Institute of Cognitive Science)

Elucidating the Cognitive Processes Involved in the Note-Taking Effect

Thesis directed by College Professor of Distinction Emeritus Dr. Alice F. Healy

The fact that research on note taking extends as far back as the early 1900's is unsurprising as it is a practice that is ubiquitous throughout our lives. Note taking helps us to retain information even when there is no opportunity to review the notes. This phenomenon is known as the *encoding effect* or the *note-taking effect*. Much of the research investigating the note-taking effect focuses on the impact of note-taking media or note-taking strategy on the size of the effect. However, there is no consensus on the cognitive mechanisms underlying the note-taking effect. When discussed, there are three primary hypotheses of cognitive processing during note-taking: *generative processing*, *cognitive effort*, and *sustained attention*. After thoroughly comparing these hypotheses, there were only three unique cognitive mechanisms that required further investigation: generative processing, summarization, and sustained attention. Therefore, the purpose of this investigation was to compare the separate effects of the three cognitive mechanisms in relation to the note-taking effect.

Two experiments were designed to compare generative processing to (a) sustained attention and (b) summarization. Generative processing is the active construction of associations between novel information and prior knowledge. Generative processing is most cited as the explanation for the note-taking effect. Summarization forces the learner to identify the most pertinent information to create a coherent synopsis. This process thereby facilitates retention and comprehension. Sustained attention, in this context, is selectively concentrating on novel information while ignoring all irrelevant distraction. Experiment 1 in the present investigation, through the measurement of task-relevant and task-irrelevant distraction, found that sustained attention is positively related to retention and generative processing is negatively associated with retention. Analyses of the content of the notes revealed that students are more likely to take summary style notes when not given specific instruction and moreover, the rating of summary style is positively correlated with performance on the retention test. Experiment 2 found that generative processing impeded and summarization facilitated retention. Therefore, although most cited for the benefits of note taking, generative processing appears to be detrimental to retention. This conclusion was deduced from both internal (distraction reports) and external (test performance) measures.

ACKNOWLEDGMENTS

For the past several years, I received support and encouragement from many people. First and foremost, Dr. Alice F. Healy has been (and continues to be) more than just a mentor and colleague to me. Since the first day I stepped into her Experimental Methods class, Alice gave me endless support and guidance. She not only taught me fundamentals of conducting rigorous, scientific research, but she also inspired me with her work ethic and unwavering passion. I would like to express my deepest appreciation to Alice for her positivity, guidance, and genuine concern for my success.

In addition to my advisor, I would like to thank the rest of my dissertation committee, Drs. Eliana Colunga, Michael Mozer, Michael Stallings, and Tamara Sumner for their insights and feedback that helped shape my research plan. I also wish to thank the many people from the University of Colorado Boulder within the Department of Psychology and Neuroscience, the Institute of Cognitive Science, and the Center for Research on Training. In particular, Jon Roberts and Ernie Mross provided an abundance of programming and technical expertise that made this research possible. Moreover, I am grateful to Shannon McKnight, Heejung Jung, Jude DeLaubenfels, and Timothy Greeson for their unconditional support, words of wisdom, and encouragement. I would also like to thank the undergraduate research assistants: Rebekah Benjamin-Pollak, Timothy Ghan, Calista Prananta, and Majd Salim Al Suqri, without whom I might still be processing the data.

Thanks also for the generous funding that helped support my research, including the National Science Foundation and the University of Colorado Institute of Cognitive Science.

Last but not least, I am grateful for my family: Kyle, for being there for me day after day; Kubs and Geeta for continually believing in me; and to my parents, Arjan and Sharmilla, for their love and support. This dissertation would not have been possible without their continued patience and wisdom. Thank you!

CONTENTS

FIGURES

Experiment 1

Figure

Experiment 2

Figure

TABLES

Experiment 1

Table

Experiment 2

Table

INTRODUCTION

There is substantial research in the fields of education and psychology investigating the role of note taking in learning. The widespread interest in note taking is not surprising given its pervasiveness in education beginning as early as primary school. Interest in note taking extends back to the early 1900's (Seward, 1910), and experimental inquiries have been conducted since 1925 (Crawford, 1925). The bulk of the research has found benefits in retention and comprehension for students taking notes from both lecture and text compared to those who just listened or read (e.g., Bretzing & Kulhavy, 1981). However, a few studies have found either no benefit for note taking or in some cases a detriment for those in the note-taking conditions in terms of retention (e.g., Kiewra, 1985; Kobayashi, 2005). The large variation in methodology might be why there is not a consensus on whether there is an advantage to note taking.

When found, the retention advantage for note taking has been attributed to two different functions of note taking, the process and the product (Di Vesta & Gray, 1972). The process function refers to the retention benefit of note taking when there is no opportunity to review the notes. This benefit of note taking has been referred to as the *encoding effect* or the *note-taking effect*. Conversely, *external storage*, also known as the product function, increases test performance due to studying the notes produced compared to only taking notes. In my previous research (Lalchandani, 2016; Lalchandani & Healy, 2017), I had investigated whether the retention benefits for handwritten notes were due to the act of taking notes or to the content of the notes. The findings suggested that the act of taking notes but not the content of the notes facilitated the retention benefits.

A review investigating the process function of note taking found that nearly half of the experiments (n = 23 of 56) reported either a null effect (n = 21) or a detriment (n = 2) for the note-taking condition (Kiewra, 1985). There are many possible explanations for why the encoding effect is not consistent across the experiments. One possibility is that there is a physical component to note taking that might impact the students' ability to process the presented information by causing the note takers to experience fatigue. It is also possible that the manner in which information is presented might hinder note taking, thus negating the encoding benefits. For example, extremely rapid presentation of the material has been attributed to some of the negative findings. Such manipulations do not yield ecologically valid learning environments. In a similar vein, lecture material that involves visual presentations create a multiple task situation involving note taking, audio processing, and visual processing, forcing the students to divide their attention, thus lowering the level of processing of the lecture material. A more recent review found that the presentation mode significantly moderated the effect sizes of 57 studies (306 *Cohen's d* effect sizes) that compared a note-taking condition to a non-note-taking condition (Kobayashi, 2005). The experiments where note taking was completed for text material resulted in larger effect sizes (d = .27) than for audio-visual presentations (d = -.02). Because the methodology in these experiments varied greatly, it is possible that the specific cognitive processes involved in the encoding effect were not utilized in the note-taking conditions or were masked by ecologically invalid manipulations.

Much of the more recent literature regarding the encoding effect has focused on different note-taking techniques and strategies to increase the benefit of note taking. Some of these studies of note taking have moved on to investigate the effect of technology used for note taking. It has been found that the note-taking strategy can be tied to the media used for note taking. For example, students who take notes on laptops and computers tend to transcribe the lecture material, whereas longhand note takers tend to summarize the material (Lalchandani, 2016; Lalchandani & Healy, 2017; Mueller & Oppenheimer, 2014). Other recent investigations have considered the effect of providing instructional aids for note taking, such as skeletal outlines or illustrated diagrams (e.g., Bui & McDaniel, 2015; Stacy & Cain, 2015). Further research has also investigated which note-taking technologies or strategies are most beneficial in non-educational settings such as legal, business, and medical contexts (e.g., Kiewra, 2016; Mueller & Oppenheimer, 2016).

The findings of my Master's research (Lalchandani, 2016; Lalchandani & Healy, 2017) have helped me to define the scope of and motivation for this current line of research. In the Master's research, I conducted two experiments about note taking. The goal of the first experiment was to replicate and extend the results of the Mueller and Oppenheimer (2014) investigation, specifically their first experiment. Their investigation used a 2 x 2 mixed factorial design. The between-subjects variable was note-taking medium (longhand vs. laptop). Question type (specific vs. conceptual) varied within subjects. Learning was assessed through a test one half hour after the lecture material was presented (notes were not available for study). Participants taking longhand notes performed better on conceptual test questions than did those taking notes on a computer but as well as those in the computer condition for the specific test questions. They also successfully linked note-taking strategy to the note-taking medium. They found that longhand note takers wrote fewer words and transcribed less of the spoken lecture than did computer note takers. In my replication of this study, I also added a no-note control condition to investigate the more general note-taking effect and I added a second retention test 1 week after the first test to investigate the durability of any effects of note-taking. Overall test performance was significantly better for students in the longhand and computer note-taking conditions than for those in the no-note control condition, showing that note taking facilitates retention and comprehension of the lecture material for computer note takers as well as for longhand note takers, even though students did not have access to their notes and thus were not permitted to review them prior to taking either retention test. I also found that although students in all three note-taking conditions performed equally well on the specific questions, longhand note takers performed better on the conceptual questions, in accordance with Mueller and Oppenheimer, but the interaction of note-taking medium and question type was not significant. However, this trend disappeared at the time of the 1-week test, such that no advantage for performance on conceptual test questions was observed for the longhand condition relative to the computer condition after the 1-week delay.

The second experiment of my Master's research examined whether the effects of note-taking medium were due to the content of the notes, as opposed to the act of taking notes, by eliminating the possibility of participants' producing their own notes. Instead participants were given notes produced by other students (the students in the first experiment), so that any remaining differences between longhand and computer notes would be due strictly to the content of the notes rather than to the act of taking notes. The participants in the second experiment were yoked to those in the first. Each student in the second experiment viewed the same lectures in the same order as did his or her yoked counterpart in the first experiment, but was not permitted to take notes. The students were then provided the notes of their counterpart to study. The better performance for the longhand note takers than for the computer note takers on conceptual questions at the time of the 30-min test in the first experiment seems attributable to the act of taking notes because there was somewhat better performance for computer note takers than for longhand note takers on conceptual questions on the critical test in the second experiment, when taking notes was not possible.

The two experiments from my Master's research have motivated the present line of research in a few ways. The note-taking effect is present regardless of the note-taking strategy used and the medium used to take notes; however it is strongest for those taking notes longhand. It is also most clearly found a short time after the presentation of the to-be-learned material. For those reasons, in the present line of research, I limit (a) the note-taking media to longhand notes and (b) the retention test to a one-half hour delay. Furthermore, based on the findings from the second experiment, I (c) focus specifically on the process function of note taking in the present line of research. Although the process function has been the subject of the majority of the studies of note taking, there has been little agreement regarding which cognitive processes are involved in note-taking benefits. In order to address the questions of how and why the note-taking effect is sometimes present and sometimes not, it is important to identify which are the pertinent cognitive processes elicited by the act of note taking. Therefore, the purpose of the present experiments is to elucidate those cognitive processes, specifically. After the roles of the relevant cognitive processes are more clearly defined, we will be able to leverage them to increase the retention benefits and to better develop effective pedagogical practices or educational technologies.

I have found that of the research that focused on cognitive processing during note taking, the findings can be categorized into three hypotheses: the *generative processing* hypothesis, the *attention* hypothesis, and the *cognitive effort* hypothesis. The note-taking effect has been explained by each of these views individually, but the hypotheses have not been formally compared. I analyzed these hypotheses for similar and differentiating features and proposed the present two experiments to help identify the different cognitive mechanisms underlying the note-taking effect and, thus, help fill this gap in the literature.

The Hypotheses

The generative processing hypothesis, the cognitive effort hypothesis, and the attention hypothesis, as they have been defined in the existing note-taking literature, each have many overlapping features. By ascertaining the differences between them, it becomes possible to more narrowly identify which of the differentiating features contributes to the encoding effect. However, these views are not mutually exclusive. It is possible that the underlying mechanisms of these three hypotheses work together to elicit the note-taking benefit. First, I will describe each of the three hypotheses as they are described in the current literature. Then I will identify the overlapping and discriminating cognitive mechanisms in the hypotheses. Finally, I describe a research plan including the present two experiments to test each mechanism in order to develop a unified hypothesis to explain the note-taking effect.

Generative Processing Hypothesis

Wittrock and Carter's (1975) model of generative processing largely relies on a levels-of-processing framework abstracted from Craik and Lockhart (1972). Although the original levels-of-processing framework applies to word lists, generative processing presumes that information is more deeply encoded when the learner actively constructs associations between the to-be-learned material and his or her prior knowledge than when simply listening to or reading the material. Generative processing is described as continuous in the level of encoding in that superficial processing leads to a lower level of encoding and deep processing leads to a higher level of encoding, with intermediate levels of processing leading to intermediate levels of encoding. Higher levels of encoding translate to increased degrees of retention. Reproductive processing, on the contrary, refers to repetition or rehearsal of information without elaboration, resulting in a shallower level of encoding.

The degree of generative processing is influenced by the style of note taking used. Transcription style notes (i.e., verbatim or near verbatim copying such as nongenerative notes) would be considered reproductive processing resulting in shallow levels of encoding, as the students are not connecting the novel material to their prior knowledge. Generative processing has been attributed to the note-taking styles that require summarizing, paraphrasing, outlining, or concept mapping (e.g., generative, organized, matrix, or Cornell style notes) (e.g., Bui, Myerson, & Hale, 2013; Lalchandani, 2016; Lalchandani & Healy, 2017; Mueller & Oppenheimer, 2014). One study examined the differences between two note-taking styles, transcription and organized notes (Bui et al., 2013). The study revealed that students who transcribed notes outperformed students who took organized notes on a test immediately following the presentation of the material. However, of the students who took organized notes, the students who completed a retention test immediately after the lecture period performed equivalently to those who completed a 1-day delayed retention test. In contrast, of the students who transcribed notes, those who took the delayed retention test performed worse than those who took the immediate retention test.

Under the generative processing hypothesis, this finding is due to organized notes requiring the students to construct associations between their prior knowledge and the lecture material. However, it appears that generative processing has been confounded with summarization. Summarization is believed to focus attention on the most relevant information, thereby improving comprehension and retention (Anderson & Armbruster, 1984). Although summarization does not require the students to actively link the novel information to their prior knowledge, Wittrock and Alesandrini (1990) consider summaries in which the students use their own words and connect across concepts as generative processing. They propose that memory mechanisms allow students to naturally associate their prior knowledge with the novel material through their choice of words (King, 1992; Wittrock & Alesandrini, 1990). Nevertheless, studies that have examined the effect of summarization on learning have most often used summaries that “simply select, delete, and modify existing sentences to produce a summary” (King, 1992, p.306). However, because the content of the notes in the Bui et al. (2013) experiment was only analyzed for the number of idea units noted, there is no way to determine if the effect found by them was due to students’ using their own words in their summaries (generative processing) or to students’ selecting and modifying the material to be written down (selection summarization).

Although it is often conflated with generative processing, the *generation effect,* as found by Slameka and Graf (1978), is a separate and distinct psychological phenomenon. The generation effect is the finding that people recall information (most often shown using word lists) significantly better when they generate the information as opposed to when they are provided with it (e.g., reading or listening). Experimental investigations have found that even completing word fragments can elicit the generation effect. Furthermore, the size of the generation effect is often larger when the generation constraints are fewer for both item and context memory (McCurdy, Leach, & Leshikar, 2017). The differences due to the number of constraints maps on to various degrees of depth of processing (Craik & Lockhart, 1972), more constraints leading to shallower processing of both the content and the context. Word completion is considered highly constrained because the number of correct responses is very limited.

However, some styles of note taking might elicit the generation effect because notes are self-generated without a correct response. Therefore, there are relatively few constraints related to creating notes. For example, generative processing notes that involve the students' generating connections across concepts might benefit from the generation effect because the students are producing the information. However, selection summaries or notes that are restatements of the material might not elicit the generation effect as the students are simply selecting which material to record from the provided information. Nevertheless, the generation effect would not account for the note-taking effect and is distinctly different from generative processing.

Because note taking is a complex, cognitive task, individual differences in attentional resources such as working memory ability could also limit the degree of generative processing. For example, working memory is necessary to manipulate information in order to construct associations with prior knowledge (Bui et al., 2013). Individuals with lower working memory ability might struggle to maintain the novel information in their working memory, thereby reducing the degree of generative processing. It has been found that working memory ability correlates with both note quantity (for all note-taking strategies) and free recall performance (for generative note-taking strategies). However, even with the increased cognitive demands, note taking could be facilitating generative processing that, in turn, leads to an increase in the level of processing for note takers compared to those who simply listen or read, resulting in better retention.

In sum, the generative processing hypothesis confounds two different cognitive processes. First is the construction of associations between new information and prior knowledge. This aspect of the generative processing hypothesis is the most cited in reference to the benefit of note taking. Second, summarization could underlie the note-taking effect due to the students' choices in selecting which material to include in their notes and what type of modifications to make. These two processes are not incompatible. It is possible that both processes could facilitate the note-taking effect. However, the content of the notes must be examined more thoroughly than has been done in previous studies in order to determine the separate effects of generative processing and summarization.

Attention Hypothesis

Also embedded in the levels-of-processing framework (Craik & Lockhart, 1972) is the attention hypothesis. This hypothesis asserts that note taking forces the learner to pay attention to the material and thereby process the information more deeply (Peper & Mayer, 1986). Under this hypothesis, attention can be defined as selectively concentrating on a discrete stimulus (the novel material) while ignoring irrelevant information (distractors). The novel material specifically refers to the lecture or text, whereas distractors can be either task-related or task-unrelated. For example, rumination over a misconception of the novel material is task-related. In contrast, rumination over general class performance is task-unrelated. Both types of distractors are thought to remove attention from the lecture or text material. Evidence for this hypothesis comes as the result of comparing retention test performance between note takers and those who simply listen to or read the material (Kiewra, 1987), without emphasis on the note-taking strategy employed. Another study found sustained attention ability to be a predictor for note style, where style was related to the note-taking strategy used (Peverly, Garner, & Vekaria, 2014). However, this experiment did not find that sustained attention ability was a significant predictor of performance, although note-taking style was. Nevertheless, with respect to the note-taking effect, the increase in attention during the presentation of the material, compared to just listening or reading, could be attributed to various phenomena such as the *cognitive antidote principle*, *decreased mind wandering*, or increased *engagement*.

The cognitive antidote principle is the finding that interrupting a task with cognitive complications can enhance task performance (Kole, Healy, & Bourne, 2008). This phenomenon has been seen with particularly tedious tasks (e.g. data entry) as an antidote to boredom (e.g., Chapman, Healy, & Kole, 2016; Kole et al., 2008). The addition of a mental math component to a data entry task reduced the speed-accuracy tradeoff in the second half of the experimental session, such that while speed increased between the first half and the second half of the session, so did overall accuracy (Kole et al., 2008). A subsequent study (Chapman et al., 2016) forced participants to remember the target and respond while the following cue was on the screen. The authors concluded that increased memory load also reduced the speed-accuracy tradeoff. More recently, the cognitive antidote principle was applied in a simulated classroom setting in the form of interleaved quiz questions in a fact-learning task (Healy, Jones, Lalchandani, & Tack, 2017). Students who answered quiz questions interspersed throughout the training period performed better on the retention tests than those who answered quiz questions at the end of the training period.

As applied in the Healy et al. (2017) study, the cognitive antidote to boredom was thought to be the result of higher cognitive engagement in the interspersed condition. Specifically, boredom and task disengagement increased in the later task blocks for the condition that lacked the additional cognitive complications. In contrast, there was little to no change in a self-reported engagement and boredom metric in the cognitive antidote condition.

Note taking, itself, can be considered an additional cognitive complication. As such, reduced mind wandering and increased task engagement could be the underlying mechanism of the retention advantage for note taking. In this context of the note-taking effect, the concepts of mind wandering and the levels of distraction are equivalent. Moreover, task-related distraction is synonymous with task-related mind wandering and task-unrelated distraction with task-unrelated mind wandering. Another interpolated testing paradigm used the quantity of notes taken (measured as the number of slides that contained any notes) during the lecture period as an assessment of task engagement and mind wandering. The authors used the measurement such that more notes recorded indicated less mind wandering and increased task engagement (Szpunar, Jing, & Schacter, 2014; Szpunar, Khan, & Schacter, 2013). The quantity of notes is also a positive predictor of test performance (Mueller & Oppenheimer, 2014). The benefit in retention due to note taking could be the result of such reduced mind wandering and increased task engagement. However, there has yet to be an empirical study that has examined the relationship between mind wandering or task engagement and note taking.

Overall, the attention hypothesis asserts that note-taking benefits are due to focused attention to the new material. Both increased task engagement and decreased mind wandering serve to increase attention to the presented material and the act of note taking could be serving as a cognitive antidote. Furthermore, the attention hypothesis predicts note-taking benefits to retention independent of the note-taking strategy used.

Cognitive Effort Hypothesis

Often confounded with the levels-of-processing framework is the distinct concept of cognitive effort. Cognitive effort can be described as the amount of a limited-capacity central processor that is engaged with a specific task (Kahneman, 1973; Tyler, Hertel, McCallum, & Ellis, 1979). More explicitly, cognitive effort is the proportion of available cognitive resources, including working memory, used to complete the task at hand. High effort tasks have been found to increase both incidental and intentional learning regardless of the level of processing (e.g., Tyler et al., 1979). In the context of note taking, a larger proportion of the cognitive resources are employed in the act of note taking compared to simply listening or reading, regardless of the note-taking strategy used.

The depth-of-processing framework predicts better retention of a semantic task (e.g., sentence completion) than a non-semantic task (e.g., anagrams) irrespective of the level of cognitive effort. However, Tyler et al. (1979) found that these two tasks did not yield significant differences on a free recall retention test. However, their manipulation of cognitive effort revealed better retention for the high effort tasks in both levels-of-processing conditions. For the shallow processing task, effort was manipulated in how many letters were rearranged in the target word, (e.g., for the target word *DOCTOR*, the low effort anagram could be *DORTOC* whereas the high effort version could be *CROODT*). Similarly, for the deep processing task, the low effort correct responses were heavily implied by the sentence fragment (e.g., *The girl was awakened by her frightening DREAM*). The correct responses for the high effort sentence fragments were less determined by the sentence context (e.g., *The man was alarmed by the frightening DREAM*). The high effort versions of both tasks seem to require more cognitive resources than the low effort counterparts due, in part, to differences in difficulty. These findings provide evidence that higher levels of cognitive effort can lead to higher retention, regardless of the depth of processing.

The cognitive effort involved with note taking has been measured through a dual-task paradigm (Piolat, Olive, & Kellogg, 2005). Participants began with the primary task of taking notes (either from written text or an audio/visual lecture). At random intervals throughout the learning period, participants responded to a tone by pressing a key or mouse button or saying stop. The reaction time between the tone and the response was then compared to a single-task (without note taking) control condition. Cognitive effort was measured as the amount the primary task interferes with performance on the secondary task.

A comparison of the cognitive effort involved in note taking from both an audio lecture and the written transcript of the lecture found that note taking from an audio lecture requires significantly more cognitive effort than from a written text when taking free-form notes (Roussey & Piolat, 2003). This finding is partially due to the self-paced nature of reading, whereas the rate of presentation of a spoken lecture is fixed. Because the learner cannot control the rate of an audio lecture, there is more demand on working memory than during reading (Piolat et al., 2005). When the participants were provided with pre-planned note structures (instructor provided outlines), there were no differences in effort between the presentation modes, as measured by performance on the secondary task. Some studies have found that more involved note-taking styles (organized or matrix style notes) yield larger note-taking benefits than less effortful styles (linear or transcription style notes) (e.g. Bui et al., 2013). Furthermore, because cognitive effort is continuous, rather than binary, the range of effort might positively correlate with retention. For example, in my Master's research (Lalchandani, 2016; Lalchandani & Healy, 2017) there were three conditions that were assumed to differ in the level of cognitive effort (although effort was not measured), with generative notes (longhand notes) requiring the most effort, verbatim notes (computer notes) less effort, and no notes the least effort. It was found that even when students took verbatim notes they had better test performance than did students in the no note control condition. In addition, replicating previous work (Mueller & Oppenheimer, 2014), the condition that required the most effort (generative notes) retained over a short time interval significantly more conceptual information than did the moderate effort condition (verbatim notes). However, the amount of cognitive effort involved in different note-taking strategies compared to a control has yet to be assessed directly.

Overlapping Features

There appear to be some features that are common among these three hypotheses, including level of difficulty and the levels-of-processing framework. Note taking as a desirable difficulty also seems inherent in these hypotheses. Desirable difficulty is an umbrella term for a variety of phenomena that although they make acquisition of novel information more difficult, they result in improved retention and generalizability (Bjork, 1994). There are various qualities associated with both the learner and the task that are necessary for the additional difficulty to result in desirable outcomes (McDaniel & Butler, 2011). The additional task must involve different cognitive processes than the primary task. Secondary tasks that use the same processing as the primary task do not yield any benefits to retention. However, desirable difficulties also require transfer appropriate processing between the secondary task and the test or outcome task. As such, if the secondary task has no relevance to the test, then the student will not exhibit any benefits to retention or generalizability due to the increase in difficulty. In addition, the learner must also be able to perform the processing required by the secondary task. If the additional cognitive demands are too great for the individual, then the learner will not only fail to experience any benefits of the difficulty, but might instead exhibit detrimental effects of the difficulty. Although note taking creates a secondary task to learning the material, thereby increasing the level of difficulty, the level of difficulty appears to be in the optimal range for most students to positively influence retention. Note taking, as a desirable difficulty was evidenced in the Bui et al. (2013) study. Taking organized notes hindered the acquisition of novel information as the transcription note takers performed better at the time of the immediate retention test. However, organized note taking facilitated retention for the delayed retention test. In the generative processing hypothesis, note-taking benefits are the result of the active role of the learner constructing associations. This active processing could be considered more difficult than passive processing. Creating effective summaries of the material that select and modify pertinent information is also likely more difficult than passive processing. Likewise, cognitive effort seems nearly synonymous with difficulty. Although, difficulty is not directly related to the attention hypothesis, the inhibition of mind wandering or the executive control required for sustaining attention could be considered a desirable difficulty. Nevertheless, in order to fully understand note taking as a desirable difficulty, it is imperative to identify the cognitive processes involved in note taking that are not redundant with the primary task of learning and are matched with the cognitive processes elicited during the test.

Similar to increased difficulty, the cognitive effort hypothesis could underlie the other two hypotheses. Generative processing and summarization both require more cognitive effort than reproductive processing (McFarland, Frey, & Rhodes, 1980). As with desirable difficulties, effort is implied in increased attention, especially in relation to the cognitive antidote principle. Additional cognitive complications inherently increase difficulty in the primary task. However, even though it has been concluded that greater effort leads to greater recall and retention (Tyler et al., 1979), it seems likely that, as with difficulty, an optimal intermediate level of effort would make note taking most beneficial. Too much effort could lead to either no effect of note taking or possible detriments due to note taking. For example, Roussey and Piolat (2003) found that pre-planned notes mitigated the effect of presentation mode (i.e., text material, audio lectures, audio-visual lectures) on cognitive effort. However, providing note-taking aids such as outlines or illustrated diagrams, although presumably reducing effort, has been shown to improve retention compared to free-form notes (Bui & McDaniel, 2015). Nevertheless, increased cognitive effort, to some degree, is required for generative processing, summarization, and sustained attention.

Although the feature of cognitive effort has been distinguished independently from the levels-of-processing framework, in the context of note taking, at present, the two are thoroughly entangled. The amount of effort in note taking seems to vary, in part, as a function of the amount of generative processing. Generative processing, in turn, has been directly tied to a deeper level of processing. Additionally, the attention hypothesis is driven by the assumption that higher levels of attention lead to deeper levels of encoding (Kiewra, 1987; Peper & Mayer, 1986). Therefore, at least in regards to note taking, cognitive effort and the levels-of-processing framework appear to be inextricably intertwined.

Differentiating Features

Given that there are considerable overlapping features among the three hypotheses, identifying the unique qualities enables empirical testing of the different cognitive mechanisms contributing to the benefit of note taking. In the case that no differences in the note-taking effect can be found due to the unique aspects of the hypotheses, it can be surmised that the overlapping features account for the benefit. However, if one feature proves to be significant, it will provide evidence in favor of one hypothesis over the others. Furthermore, if multiple features are found significant, it would indicate that they work collectively to explain the note-taking effect. In any event, a more unified framework of the note-taking effect could be developed.

The first component of cognitive processing that was not included in all three hypotheses is sustained attention. Although attention has been related to the levels-of-processing framework, it has not been directly addressed in the generative processing or cognitive effort hypotheses. It is possible that note taking promotes encoding and retention only because the learner is forced to focus his or her attention to the material, regardless of increased effort, generative processing, or the effect of summarization. In fact, generative processing could lead to task related distraction to select relevant prior knowledge. Although this behavior would enhance generative processing effects, it could reduce the learners' attention to the presented material.

In relation to cognitive effort, there is likely some amount of effort that could be attributed to the act of paying attention and the inhibition of mind wandering. Focused attention alone might result in a greater benefit to learning via an optimized level of cognitive effort, compared to the relatively greater amount of effort associated with note taking. There is evidence that the cognitive effort due to free-form note taking yields detriments to retention compared to assisted note taking (Bui & McDaniel, 2015; Roussey & Piolat, 2003). On the assumption that free-form note taking requires more cognitive effort than assisted note taking, further reduction in cognitive effort might prove to be even more beneficial. However, the differences in learning between note taking and focused attention without note taking (e.g., motivating attention through rewards) have yet to be evaluated.

The second component of cognitive processing that is not incorporated in all the hypotheses is the role of generative processing, specifically the constructing of associations between the novel material and prior knowledge. Although cognitive effort might mitigate the extent to which the learner generates connections, cognitive effort does not necessitate generative processing. Cognitive effort can be increased during learning, for example by accelerating the rate of presentation, without eliciting generative processing. Similarly, an increase in cognitive effort could be due to summarization, the third and final component of cognitive processing, which has been confounded with generative processing. Also, as previously mentioned, generative processing could lead to task related mind wandering and therefore to decreased attention.

The generative processing hypothesis confounds the process of actively associating novel information to prior information and the process of summarization. Although, there are some types of summarization that might involve connecting concepts to the students' existing knowledge, summarization benefits have been shown primarily with selection summaries (Bretzing & Kulhavy, 1979). Writing selection summaries is thought to help students attend to and identify the broader conceptual framework of the to-be-learned material, thereby facilitating retention. Summarization necessarily involves attention, but does not necessarily require generative processing. Therefore, from the three hypotheses, the unique cognitive processes that have been identified are sustained attention, generative processing, and summarization.

Distinguishing Relevant Cognitive Processes

In order to determine which cognitive processes are involved in the note-taking effect, the present two experiments were designed in such a way that the different hypotheses make opposing predictions. Hence, these two experiments are able to discriminate between the distinct cognitive processes.

EXPERIMENT 1

The first experiment examines the unique effects of sustained attention and generative processing on retention by evaluating the amount of two different types of distraction the students reported (task-relevant and task-irrelevant distraction). If sustained attention is the relevant cognitive mechanism underlying the note-taking effect, the students taking notes should record fewer instances of distraction all together (i.e., both task-relevant and task-irrelevant) because the act of taking notes will force them to pay attention only to the lecture material. The students not taking notes would be predicted to indicate more task-irrelevant distraction, as they are not being forced to pay attention to the material and have no additional task to serve as a cognitive antidote. However, attention processing does not lead to any predictions about the amount of task-relevant distraction for students not taking notes. If generative processing underlies the note-taking effect, for this experiment, nearly the opposite pattern of distraction reporting would be expected. Students who take notes would be predicted to experience more task-relevant distraction than those not taking notes because note taking facilitates the students' ability to link the novel information to their prior knowledge. Although, generative processing leads to clear predictions for task-relevant distraction, it does not for task-irrelevant distraction for students in either condition. Nevertheless, the pattern of task-relevant distraction signifies one process over the other.

Method

Participants

The participants were 127 undergraduate college students from the University of Colorado, all of whom received partial course credit in an introductory course on General Psychology in the College of Arts and Sciences for their participation. There were no restrictions (e.g., with respect to gender, age, or ethnicity) imposed on recruiting those students. The student sample was composed of 90 female students and 37 male students with an average age of 19.10 years (sd = 1.47). With respect to race/ethnicity, the student sample was composed of 90 white/Caucasian students, 20 Asian students, 4 black/African American students, 2 Hispanic/Latino students, and 11 identified as other or preferred not to answer. The majority of the students were 1st year freshmen (67) and 2nd year sophomores (34). One student preferred not to answer the question regarding current year in the university. The remaining 26 students were 3rd year and beyond. One participant was excluded from analysis for failure to complete the full experimental session. Thus, analysis was conducted on data from 126 participants, which permitted full counterbalancing with 42 participants per experimental condition.

Design

A 3 x 2 mixed factorial design was used. Experimental condition varied between subjects. One third of the students took notes and self-reported their distracted thoughts (notes report). The remaining students did not take notes, however half of these students self-reported their distracted thoughts (no notes report). The other half did not report their distracted thoughts (no notes no report). Question type on the retention test (specific, conceptual) varied within subject. The first dependent measure in this experiment was performance on a retention test in order to verify the note-taking effect. There were two measures of distraction: task-relevant distraction and task-irrelevant distraction. The content of the notes was also evaluated

Materials

Two portions of lectures were used from the Massachusetts Institute of Technology (MIT) OpenCourseWare (OCW) database from the Introduction to Psychology course (Lecture A and Lecture B), taught by Professor John Gabrieli (2011). The OCW is an online database of nearly all MIT course content. Lecture A was approximately 10 min long, and Lecture B was approximately 18 min long. Lecture A was the full “Cognitive Dissonance”[1] portion of the class (from the Social Psychology II chapter), and Lecture B was the full portion of the class covering “Frontal Lobes and Thinking”[2] (from the Thinking chapter). The order of the two lectures was counterbalanced across participants in each note-taking condition. A pretest was previously administered to 80 students from the same population (students in the General Psychology course) as the participants, to check for preexisting knowledge. The level of preexisting knowledge (32.86% overall, 28.23% for conceptual questions, and 37.50% for specific questions) was significantly higher than chance (1 out of 4 choices, or 25%) (overall $t(79) = 6.352$, $p < .001$; conceptual $t(79) = 2.491$, $p = .015$; specific $t(79) = 6.931$, $p < .001$) but not high in absolute terms. Both lecture segments were chosen because the material was fully encompassed within the respective time frames and did not refer to material from previous lectures. Psychology content was chosen specifically to allow for greater control over the participants’ prior knowledge, which would be expected to be quite variable given content from other courses. Together the total lecture time was approximately one half hour. The students viewed the lecture capture videos available on the OCW, complete with the PowerPoint slides and views of the professor.

All students watched the lectures individually on Apple iMac desktop computers through the monitor speakers. The students in the notes report condition were provided with a legal pad and pen with which to take notes. The size of the lecture window was consistent across conditions. All participants then took a test written by the author (*see Appendix A*) and

---

[1] https://ocw.mit.edu/courses/brain-and-cognitive-sciences/9-00sc-introduction-to-psychology-fall-2011/social-psychology-ii/

[2] https://ocw.mit.edu/courses/brain-and-cognitive-sciences/9-00sc-introduction-to-psychology-fall-2011/thinking/

previously used in my Master's thesis research (Lalchandani, 2016; Lalchandani & Healy, 2017). The test consisted of two question types (specific and conceptual) following the taxonomy created by Butler (2010). The test was composed of 12 questions of each type (specific and conceptual) in multiple-choice format with four answer choices. Specific questions are questions where the correct answer is directly provided within the lecture material. For example, the correct answer to the question "Who coined the term cognitive dissonance?" can be found clearly in the lecture. Conceptual questions, on the other hand, require some abstraction from the lecture material to correctly answer the question. For example, the correct answer choice for the conceptual question "Which is an example of a beneficial utilization behavior?" requires the student to understand the definition of utilization behavior in order to derive the correct choice. The varying of question format facilitates the assessment of any transfer effects from note taking and the evaluation of both retention and comprehension. To increase the ecological validity of the paradigm, the test was administered using a paper and pen format like that most often used in classrooms.

Procedure

All students watched the two lectures. As mentioned earlier, the lecture order was fully counterbalanced across students in each between-subjects condition. Students in the notes report condition were instructed to take notes during the lecture using their normal note-taking styles. Students in both of the no note conditions (no notes report, no notes no report) were instructed just to pay attention to the lecture material. All students were told to pretend that they were in a normal class where they might take a test on the lecture material. The students in the notes report and no notes report conditions were instructed to indicate when they were distracted through key presses. For task-relevant distraction, students were told to signal when they had a thought that related the lecture material to their prior knowledge (e.g., something they had learned before or as the material was related to their personal lives) through a key press (pressed the “left arrow” key). When the students had thoughts that were completely unrelated to the lecture material, they recorded those thoughts as task-irrelevant by pressing a different key (pressed the “right arrow” key). Labels were placed above the keys to help the students remember which key indicated which type of thought. The students in the no notes no report condition did not have any secondary task. After both lectures were watched, the students’ notes were collected. They then participated in a battery of unrelated distractor tasks for 30 min: a letter detection task, a states and capitals crossword, and a simple mathematics worksheet. After the 30-min retention interval, all students then completed the multiple-choice retention test in pen and paper format (without any notes). After the test, all students were asked to complete a demographic and note-taking preference questionnaire. They were also asked to estimate the amount of time they were distracted during the lecture period.

The proportion of correct responses on the test were analyzed to verify that the note-taking effect was present. The amount of distraction (task-relevant and task-irrelevant) was analyzed to distinguish between the generative processing and increased attention hypotheses. The content of the notes was analyzed for (a) the amount of verbatim overlap with the transcripts from the lecture, (b) separate ratings of note style for summary style and for generative processing style, and (c) latent semantic analysis (LSA) to assess the degree of difference between the word choices in the notes and in the transcript of the lecture. The content analysis provided an exploratory view of whether students tend to take notes that involve generative processing or selection summaries when not given explicit instructions on what to include in their notes.

Results

Test Performance

A 2 x 3 mixed factorial analysis of variance, ANOVA, revealed a main effect of question type, $F(1,123) = 17.39$, $MSE = .013$, $p < .001$, $\eta_p^2 = .124$ (*see Figure 1*). Students performed better on the specific questions ($M = .67$, $SEM = .02$) than on the conceptual questions ($M = .61$, $SEM = .02$). This finding is unsurprising due to the difference in the level of difficulty between the question types. There was also a main effect of experimental condition, $F(2,123) = 3.29$, $MSE = .059$, $p = .041$, $\eta_p^2 = .05$. Tukey HSD post hoc test pairwise comparisons showed that the students who took notes and reported their mind wandering (notes report) ($M = .69$, $SEM = .02$) performed significantly better on the retention test than did those students who did not take notes but did report their distraction (no notes report) ($M = .59$, $SEM = .02$, $p = .034$). Although the students in the notes report condition performed numerically better than the control condition (no notes no report) ($M = .63$, $SEM = .02$), the difference was not significant ($p > .05$). There was no evidence of an interaction between experimental condition and question type ($F(2,123) = 1.90$, $MSE = .013$, $p = .153$, $\eta_p^2 = .03$).

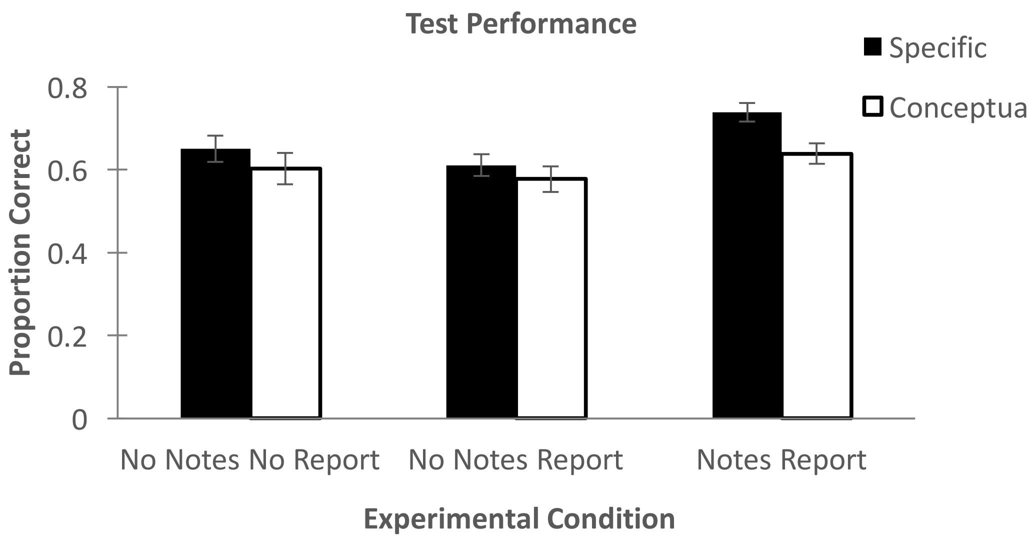

Figure 1. Mean proportion of correct test answers based on experimental condition and question type. The error bars reflect between-subjects standard errors of the mean.

Distraction

Because there was so much variability in the data for self-reported levels of distraction for both groups (notes report and no notes report), I conducted parallel analyses both on the raw data (i.e. the number of times each student indicated either task-relevant or task-irrelevant distraction) and on the data after using a logarithmic transformation. Because there were a few zeros in the data that would have yielded undefined scores, as there were no indications of any technical or comprehension issues, a constant, 1, was added to each raw score prior to taking the log base 10. The patterns of the results are very similar for both measures; however, the log transformation appeared to be a more sensitive measure, as evident in the reduced error bars.

A one-way ANOVA showed that although there was a numerical difference between the two groups (notes report: M = 20.93, SEM = 2.81; no notes report: M = 16.17, SEM = 2.12), there was no effect of note taking on the raw task-irrelevant distraction measure $F(1,82)$ = 1.83, MSE = 260.837, $p = .180$, $\eta_p^2 = .02$ (*see Figure 2*). However, the more sensitive measure, the log transformed data, yielded a main effect of note taking $F(1,82) = 5.01$, $MSE = .100$, $p = .028$, $\eta_p^2 = .06$ (*see Figure 3*), such that the students who took notes reported significantly less task-irrelevant distraction ($M = 1.10$, $SEM = .06$) than did those students who did not take notes ($M = 1.25$, $SEM = .04$).

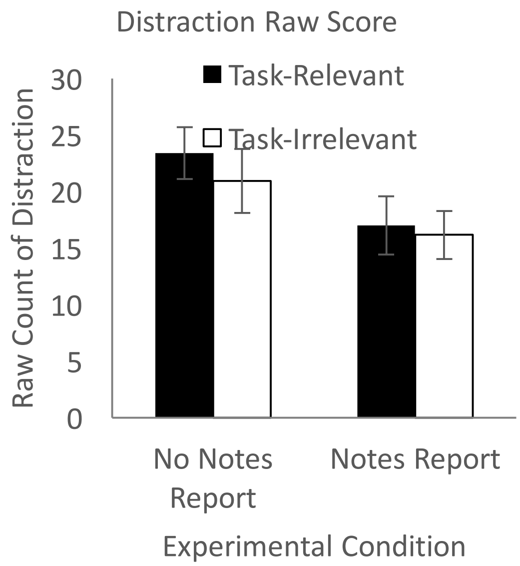

Figure 2.
Mean raw counts of each type of distraction reported by the students. Error bars represent between-subjects standard errors of the mean.

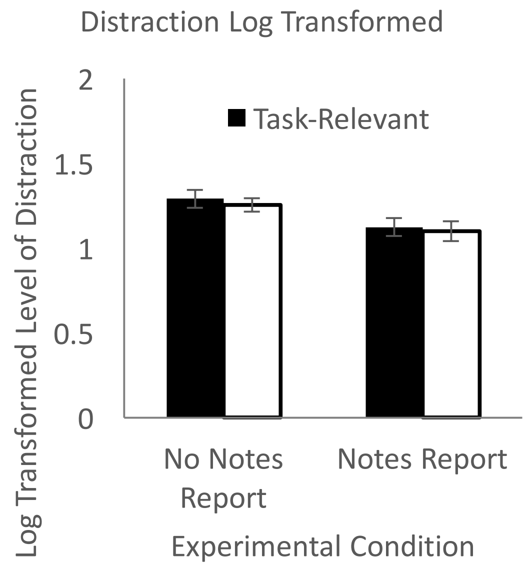

Figure 3.
Mean log transformed levels of each type of distraction reported by the students. Error bars represent between-subjects standard errors of the mean.

A similar pattern of results was found in the analysis of task-relevant distraction. A one-way ANOVA found a marginally significant effect of note taking on the raw measure of task-relevant distraction $F(1,82) = 3.44$, $MSE = 250.636$, $p = .067$, $\eta_p^2 = .04$. This finding was confirmed using the transformed data $F(1,82) = 4.97$, $MSE = .117$, $p = .029$, $\eta_p^2 = .06$. Students who took notes reported significantly less task-relevant distraction (raw: $M = 17.00$, SEM =

2.57; log: *M* = 1.12, *SEM* = .05) than did the students who did not take notes (raw: *M* = 23.40, *SEM* = 2.31; log: *M* = 1.29, *SEM* = .05).

A pair of 2 x 2 mixed factorial ANOVAs on both the raw data and the transformed data yielded no differences in the amounts of the two types of distraction (log: $F < 1$, raw: $F(1,82) = 1.23$, MSE = .046, $p = .27$, $\eta_p^2 = .02$). Students, overall, reported an equivalent number of task-relevant distractions (log: *M* = 1.21, *SEM* =.04; raw: *M* = 20.20, *SEM* = 1.75) as they did task-irrelevant distractions (log: *M* = 1.18, *SEM* = .04; raw: *M* = 18.55, *SEM* = 1.77).

Pearson product-moment correlation analyses were conducted to assess the relationship between test performance and distraction (*see Table 1*). Overall, there was a marginally significant positive relationship between the log-transformed, task-relevant measure of distraction and performance on both conceptual questions, $r = .19$, $n = 84$, $p = .078$, and overall test performance, $r = .20$ $n = 84$, $p = .075$.

Table 1.
Correlation matrix of test performance and distraction reports. Values below the diagonal are Pearson correlation coefficients. Above the diagonal are the probabilities associated with the corresponding correlations.

	Specific	Concept	Total	Raw On-Task	Raw Off-Task	Raw Total	Log On-Task	Log Off-Task	Log Total
Specific	1	$p < .001$	$p < .001$	$p = .615$	$p = .641$	$p = .986$	$p = .150$	$p = .447$	$p = .904$
Concept	.64	1	$p < .001$	$p = .358$	$p = .844$	$p = .693$	$p = .078$	$p = .896$	$p = .373$
Total	.89	.92	1	$p = .425$	$p = .721$	$p = .811$	$p = .075$	$p = .638$	$p = .561$
Raw On-Task	.06	.10	.09	1	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p < .001$
Raw Off-Task	-.05	-.02	-.04	.64	1	$p < .001$	$p < .001$	$p < .001$	$p < .001$
Raw Total	.00	.04	.03	.91	.91	1	$p < .001$	$p < .001$	$p < .001$
Log On-Task	.16	.19	.20	.86	.52	.76	1	$p < .001$	$p < .001$
Log Off-Task	-.08	-.01	-.05	.58	.85	.79	.60	1	$p < .001$
Log Total	.01	.10	.06	.79	.74	.84	.88	.88	1

Investigating this relationship further revealed that increases in the number of task-relevant distraction are correlated with increases in test performance for the students who did not take notes (*see Tables 2 and 3*). Specifically, there was a positive correlation between task-relevant distraction and performance on specific questions, $r = .33$, $n = 42$, $p = .032$, and overall test performance, $r = .34$, $n = 42$, $p = .025$. The relationship between conceptual question performance and task-relevant distraction was also marginally significant for students who did not take notes, $r = .30$, $n = 42$, $p = .056$. However, there was no evidence of any correlations between task-relevant distraction and any measure of test performance for students who did take notes.

Table 2.
Correlation matrix of test performance and distraction reports for only the students who *<u>did not</u>* *take notes. Values below the diagonal are Pearson correlation coefficients. Above the diagonal are the probabilities associated with the corresponding correlations.*

	Specific	Concept	Total	Raw On-Task	Raw Off Task	Raw Total	Log On-Task	Log Off-Task	Log Total
Specific	1	$p < .001$	$p < .001$	$p = .120$	$p = .403$	$p = .191$	$p = .032$	$p = .694$	$p = .161$
Concept	.65	1	$p < .001$	$p = .091$	$p = .447$	$p = .184$	$p = .056$	$p = .624$	$p = .140$
Total	.89	.92	1	$p = .072$	$p = .380$	$p = .145$	$p = .025$	$p = .624$	$p = .111$
Raw On-Task	.24	.26	.28	1	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p < .001$
Raw Off Task	.13	.12	.14	.57	1	$p < .001$	$p = .006$	$p < .001$	$p < .001$
Raw Total	.21	.21	.23	.86	.91	1	$p < .001$	$p < .001$	$p < .001$
Log On-Task	.33	.30	.34	.87	.42	.70	1	$p = .001$	$p < .001$
Log Off-Task	.06	.08	.08	.56	.91	.85	.50	1	$p < .001$
Log Total	.22	.23	.25	.87	.74	.90	.88	.83	1

Table 3.
Correlation matrix of test performance and distraction reports for only the students who did take notes. Values below the diagonal are Pearson correlation coefficients. Above the diagonal are the probabilities associated with the corresponding correlations.

	Specific	Concept	Total	Raw On-Task	Raw Off Task	Raw Total	Log On-Task	Log Off-Task	Log Total
Specific	1	p = .012	p < .001	p = .949	p = .234	p = .592	p = .319	p = .814	p = .844
Concept	.38	1	p < .001	p = .924	p = .278	p = .562	p = .402	p = .973	p = .600
Total	.81	.85	1	p = .981	p = .172	p = .501	p = .272	p = .911	p = .655
Raw On-Task	.01	-.02	.00	1	p < .001	p < .001	p < .001	p < .001	p < .001
Raw Off-Task	-.19	-.17	-.22	.73	1	p < .001	p < .001	p < .001	p < .001
Raw Total	-.09	-.09	-.11	.94	.92	1	p < .001	p < .001	p < .001
Log On-Task	.16	.13	.17	.83	.63	.79	1	p < .001	p < .001
Log Off-Task	-.04	.01	-.02	.57	.88	.76	.65	1	p < .001
Log Total	.03	.08	.07	.73	.80	.82	.89	.90	1

Note Content

In order to analyze the content of the notes taken by students in the notes report condition, I used three different measures: verbatim overlap, note style, and LSA. Verbatim overlap measures the amount of both 2-word and 3-word sequences in the notes that were taken directly from the transcript of the lecture. For example, in the lecture, the professor said "And so these patients seem to have this disconnection between thought and emotion" (Frontal Lobes and Thinking lecture, 12 min. 27 sec.). One participant wrote in the notes, "disconnection between thought and emotion." In this example, there are four 2-word verbatim sequences (i.e., disconnection-between, between-thought, thought-and, and-emotion) and three 3-word verbatim sequences (i.e., disconnection-between-thought, between-thought-and, thought-and-emotion). In order to account for differences in the amount written by each student, these frequencies were converted into the proportion of each student's notes that overlapped with the transcript of the lecture. In order to determine the style of the notes in terms of summary and generative processing, four undergraduate research assistants were recruited to independently rate the degree to which each set of notes followed directions for summary notes and generative processing notes (*see the Method of Experiment 2 and Appendix B for these directions*). The ratings used a 1 – 6 Likert scale where 1 indicated that the notes did not follow the instructions at all and 6 indicated that the notes fully followed the instructions. The average rating for each type of notes was then used for analysis. Latent semantic analysis (LSA) (Landauer & Dumais, 1997) was used to evaluate the terms used in each set of notes and compared them for semantic similarity to the terms used in the transcript of the lecture. The LSA method used extracted similarity cosines based on a semantic space containing three psychology textbooks at the university level (Laham, 1998).

Pearson product-moment correlations were computed to assess the relationship between the amount of verbatim overlap and test performance (*see Table 4*). Word count (i.e., the total number of words written in the students' notes) was positively associated with performance on specific questions, $r = .44$, $n = 42$, $p = .004$, conceptual questions, $r = .31$, $n = 42$, $p = .046$, and overall test performance, $r = .41$, $n = 42$, $p = .007$. The more notes the students took, the better they performed on the retention test. There was no evidence of a relationship between 2-word verbatim sequences and any measure of test performance. However, 3-word verbatim sequences were negatively associated with performance on conceptual questions, $r = -.43$, $n = 42$, $p = .005$, and overall test performance, $r = -.35$, $n = 42$, $p = .021$. Although the correlation between 3-word verbatim sequences and performance on specific questions was not significant, the trend was also negative, $r = -.20$, $n = 42$, $p = .202$. Higher proportions of verbatim overlap were associated with lower test performance.

Table 4.
Correlation matrix of test performance and the note content measures for the students who took notes. Values below the diagonal are Pearson correlation coefficients. Above the diagonal are the probabilities associated with the corresponding correlations.

	Specific	Concept	Total	Sum. Style	Gen. Style	Word Count	2-word Overlap	3-word Overlap	LSA
Specific	1	$p < .001$	$p < .001$	$p = .003$	$p = .003$	$p = .004$	$p = .221$	$p = .202$	$p = .138$
Concept	.64	1	$p < .001$	$p = .055$	$p = .028$	$p = .046$	$p = .221$	$p = .005$	$p = .134$
Total	.89	.92	1	$p = .008$	$p = .004$	$p = .007$	$p = .176$	$p = .021$	$p = .099$
Sum. Style	.45	.30	.41	1	$p = .187$	$p < .001$	$p = .043$	$p = .079$	$p = .028$
Gen. Style	-.45	-.34	-.43	-.21	1	$p = .168$	$p = .504$	$p = .520$	$p = .830$
Word Count	.44	.31	.41	.84	-.22	1	$p = .226$	$p = .114$	$p = .020$
2-word Overlap	.19	.19	.21	.31	-.11	.19	1	$p = .053$	$p = .162$
3-word Overlap	-.20	-.43	-.35	-.27	.10	-.25	-.30	1	$p = .173$
LSA	.23	.24	.26	.34	-.03	.36	.22	-.21	1

A one-way repeated measures ANOVA revealed an effect of note style on rating, $F(1,41) = 190.46$, $MSE = .059$, $p < .001$, $\eta_p^2 = .82$. The students were more likely to write notes using a summary style ($M = 3.34$, $SEM = .16$) than a generative processing style ($M = 1.04$, $SEM = .02$). Furthermore, Pearson product-moment analyses revealed that there were positive correlations between summary ratings and test performance. The higher the summary rating the better the students performed on specific questions, $r = .45$, $n = 42$, $p = .003$, and overall, $r = .41$, $n = 42$, $p = .008$. There was also a marginally significant positive correlation between summary rating and performance on conceptual test questions, $r = .30$, $n = 42$, $p = .055$. In contrast, the relationships between generative processing ratings and test performances were all negative. Students who rated higher for relating the material to themselves and their prior knowledge scored worse on specific questions, $r = -.45$, $n = 42$, $p = .003$, conceptual questions, $r = -.34$, $n = 42$, $p = .028$, and overall, $r = -.43$, $n = 42$, $p = .004$. Additionally, summary ratings were also positively correlated with word count, $r = .84$, $n = 42$, $p < .001$ and with 2-word verbatim sequences, $r = .31$, $n = 42$, $p = .043$.

Additional correlational analyses were conducted to assess the relationship between these measures of note content (i.e., word count, 2-word and 3-word verbatim sequences, summary style ratings, and generative processing style ratings) and the different measures of distraction (*see Table 5*). Summary style was negatively associated with task-irrelevant distraction, raw: $r = -.48$, $n = 42$, $p =.003$; log: $r = -.31$, $n = 42$, $p =.064$ and the raw measure of total distraction, $r = -.39$, $n = 42$, $p =.020$. Higher ratings of summary are associated with fewer reports of task-irrelevant distraction and maybe also with total distraction. Furthermore, the measure of more words written in the students' notes was related to fewer reports of task-relevant distraction, $r = -.29$, $n = 42$, $p =.087$, task-irrelevant distraction, $r = -.46$, $n = 42$, $p =.004$, and overall distraction, $r = -.41$, $n = 42$, $p =.011$[3]

[3] These findings neared or passed the threshold for significance for the raw measure of distraction. As seen in Table 5, the log transformed measures of distraction did not approach significance.

Table 5. *Correlation matrix of the note content measures and distraction measures for the students who took notes. Values below the diagonal are Pearson correlation coefficients. Above the diagonal are the probabilities associated with the corresponding correlation coefficients.*

	Raw On-Task	Raw Off-Task	Raw Total	Log On-Task	Log Off-Task	Log Total	Sum. Style	Gen. Style	Word Count	2-word Overlap	3-word Overlap	LSA
Raw On-Task	1	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p = .195$	$p = .742$	$p = .087$	$p = .945$	$p = .318$	$p < .001$
Raw Off-Task	.64	1	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p = .003$	$p = .289$	$p = .004$	$p = .236$	$p = .073$	$p < .001$
Raw Total	.91	.91	1	$p < .001$	$p < .001$	$p < .001$	$p = .020$	$p = .443$	$p = .011$	$p = .539$	$p = .123$	$p < .001$
Log On-Task	.86	.52	.76	1	$p < .001$	$p < .001$	$p = .419$	$p = .710$	$p = .124$	$p = .730$	$p = .532$	$p = .016$
Log Off-Task	.58	.85	.79	.60	1	$p < .001$	$p = .064$	$p = .877$	$p = .075$	$p = .877$	$p = .368$	$p = .007$
Log Total	.79	.74	.84	.88	.89	1	$p = .185$	$p = .854$	$p = .115$	$p = .709$	$p = .502$	$p = .006$
Sum. Style	-.22	-.48	-.39	-.14	-.31	-.23	1	$p = .096$	$p < .001$	$p = .015$	$p = .084$	$p = .009$
Gen. Style	.05	.17	.12	.06	-.03	.03	-.28	1	$p = .089$	$p = .386$	$p = .435$	$p = .310$
Word Count	-.29	-.46	-.41	-.26	-.30	-.26	.85	-.28	1	$p = .101$	$p = .118$	$p = .007$
2-word Overlap	.01	-.20	-.10	.06	.03	.06	.40	-.14	.27	1	$p = .037$	$p = .036$
3-word Overlap	.16	.29	.25	.10	.15	.11	-.29	.13	-.26	-.34	1	$p = .105$
LSA	-.56	-.66	-.67	-.37	-.42	-.42	.43	-.16	.44	.34	-.26	1

Pearson product-moment correlation coefficients were also calculated investigating the relationship between test performance, distraction levels, and LSA cosines (again see Tables 4 and 5). There were significant negative relationships between LSA cosines and task-relevant distraction, raw: $r = -.56$, $n = 42$, $p < .001$; log: $r = -.37$, $n = 42$, $p = .016$, and task-irrelevant distraction, raw: $r = -.66$, $n = 42$, $p < .001$; log: $r = -.42$, $n = 42$, $p = .007$. More similarity between the notes and the transcripts is associated with less distraction. There was a marginally significant positive correlation between LSA cosine and overall test performance, $r = .26$, $n = 42$, $p = .10$ (*again see Table 4*). Students with notes more similar to the transcript of the lecture performed better on the retention test.

Discussion

As outlined previously, the purpose of this first experiment was to identify the unique effects of generative processing and sustained attention in relation to note taking. In order to separate the two processes, the numbers of both task-relevant and task-irrelevant distracted thoughts the students had were examined. The two processes lead to predictions of nearly opposite patterns of distracted thoughts. Sustained attention processing leads to the prediction that the students taking notes would experience fewer overall distracted thoughts than would the students not taking notes. In contrast, generative processing leads to predictions of more task-relevant distractions for the students taking notes than for those who did not take notes.

Because instructing students to report their thoughts might impact their test performance, I included the no note no report control condition to evaluate the effect of the thought-reporting task. The students in the no note no report condition performed numerically better than the students who did not take notes but did report their thoughts, providing some evidence that the additional thought reporting task negatively influenced test performance. This trend is unsurprising given the literature regarding multitasking and task switching (e.g., Hembrooke & Gay, 2003). However, after controlling for the effect of reporting, the students in the note report condition did perform significantly better than those in the no note report condition, confirming the note-taking effect was successfully elicited in this experiment.

To test the implications of the two cognitive mechanisms, I first examined the effect of note taking on each type of distraction (task-relevant, task-irrelevant). As discussed previously, generative processing leads to the prediction that there would be more task-relevant thoughts for the students who took notes than for those who did not take notes because those students are actively constructing associations between the novel lecture material and their prior knowledge. However, the opposite result was found. Students who did not take notes reported more task-relevant distracted thoughts than those who did take notes. A similar result was found for the task-irrelevant measure of distraction. Students who took notes reported less task-irrelevant distraction. These findings provide evidence in favor of the sustained attention hypothesis. It seems that sustained attention was a more influential cognitive process underlying the note-taking effect than generative processing.

However, the conclusion from this experiment is not to say that generative processing does not facilitate learning. There is evidence that for the students who did not take notes, those who engaged in more generative processing, as measured by task-relevant distraction, tended to perform better on the retention test. However, with the additional task of taking notes, it is possible that the level of difficulty and effort required by generative processing pushed the task above the optimal range of difficulty and effort so there was no relation between generative processing and retention test performance for the students who did take notes.

An additional benefit of this experiment is the fact the students were given no instruction regarding their note-taking strategy. Exploratory analyses of the content of the notes provide insight into the strategies students naturally employ and can help distinguish between summarization and generative processing. When left to their own devices, student use a strategy more similar to summary writing than to generative processing notes as evidenced by the summary and generative processing ratings. In addition, it appears that students are more likely to use a selection summary note-taking style, (i.e., selection, deletion, and modification) (King, 1992), as opposed to the definition of summary note-taking that requires students to use their own words and connect across concepts (Wittrock & Alesandrini, 1990). This idea can be gleaned from the positive correlation between summary ratings and 2-word verbatim sequences. However, more research is needed to verify this conclusion.

EXPERIMENT 2

In the previous experiment, generative processing was directly contrasted with sustained attention. Nevertheless, the effects of generative processing and summarization must also be distinguished from each other. Therefore, in this second experiment, the act of relating information to prior knowledge and the act of summarizing were manipulated and compared to each other and to a no notes control condition. This manipulation directly ascertains the separate effects of generative processing and summarization and is important because those two processes are often confounded with each other. The students who wrote summaries did not explicitly construct associations between the lecture and their prior knowledge through their notes, and the students who took notes that required generative processing did not explicitly summarize the lecture material. Assuming that generative processing is the dominant cognitive mechanism responsible for the note-taking effect, the students relating the material to their own lives and their existing knowledge base would outperform the students summarizing the lecture on the retention test because high performance on the test would require associating new information to prior knowledge. In contrast, if summarization processing is the basis of the note-taking effect, the students summarizing the lecture would outperform those relating it to themselves because they would be focusing their attention on the most important information in the lecture. Thus, the differences in retention signify one cognitive process over the other.

Method

Participants

The participants were 129 undergraduate college students from the University of Colorado, all of whom received partial course credit in an introductory course on General Psychology in the College of Arts and Sciences for their participation. As in Experiment 1, there were no restrictions imposed on recruiting those students. The student sample for this experiment was composed of 94 female students, 34 male students, and 1 student who preferred not to answer this question, with an average age of 19.45 (sd = 2.76). With respect to race/ethnicity, the student sample was composed of 82 white/Caucasian students, 13 Asian/Pacific Islander students, 2 black/African American students, 10 Hispanic/Latino students, and 22 who identified as other or preferred not to answer. The majority of the students were

1st year freshmen (70) and 2nd year sophomores (29). One student preferred not to answer the question regarding current year in the university. The remaining 29 students were 3rd year and beyond. Three participants were excluded from analysis, one due to experimenter error and two due to not understanding the task instructions. Thus, analysis was conducted on data from 126 participants, which permitted full counterbalancing with 42 participants per experimental condition.

It is important to note that participants in Experiments 1 and 2 were tested concurrently, with participants randomly assigned to the two experiments. Therefore between-experiment comparisons are appropriate and statistically meaningful in this study.

Design

This experiment also employed a 3 x 2 mixed factorial design. The type of content included in the notes varied between subjects. In this experiment, two thirds of the students were instructed to take notes. The remaining third of the students did not take notes (no notes). Half of the students taking notes were instructed to take notes on how the material presented related to their prior knowledge, themselves, or their lives (generative notes). The other half of the students taking notes were told to write a selection summary of the lecture material (summary notes). As in Experiment 1, the retention test also included the within-subject variable of question type (specific, conceptual).

Materials

The materials in this experiment were the same lectures, distractor tasks, and tests that were used in Experiment 1.

Procedure

All students watched both lectures individually through the monitor speakers. As in Experiment 1, lecture order was fully counterbalanced across between-subjects conditions. One third of the students were instructed to take notes on how the material presented related to their prior knowledge, themselves, or their lives (generative notes) (*see Appendix B*). Specifically, the students were instructed to actively associate the lecture material to themselves. For example, in the lecture explaining the concept of cognitive dissonance, students might have written down an example from their own lives of when they experienced cognitive dissonance. The second third of the students were told to write a summary of the lecture material (summary notes). In order to write the summary, the students had to choose which information to record, what to omit, and how to manipulate the information to create a coherent synopsis. The remaining third of the students did not take notes during the lecture period, and were instructed just to pay attention to the lecture (no notes). All students were told to pretend they were attending a normal class where they might expect to be tested on the material. After both lectures were watched, the students' notes were collected. They then participated in a battery of unrelated distractor tasks for 30 min, the same as those that were used in Experiment 1. After the 30-min retention interval, all students completed the multiple-choice retention test in pen and paper format. After the test, all students then completed a demographic and note-taking preference questionnaire.

Results

Test Performance

A 2 x 3 mixed factorial ANOVA revealed a main effect of question type, $F(1,123) = 8.77$, $MSE = .015$, $p = .004$, $\eta_p^2 = .125$ (*see Figure 4*). Students performed better on the specific questions (*M* = .69, *SEM* = .02) than on the conceptual questions (*M* = .64, *SEM* = .01). As in Experiment 1, this finding is unsurprising due to the difference in the level of difficulty between the question types. There was also a main effect of instruction condition, *F*(2,123) = 4.97, *MSE* = .045, *p* = .008, η_p^2 = .08. A Tukey HSD post hoc test showed that the students who wrote summaries (*M* = .70, *SEM* = .02) performed significantly better on the retention test than did the students in the generative condition, who took notes relating the lecture material to their prior knowledge (*M* = .61, *SEM* = .02, *p* = .009). The students who did not take any notes (no notes control) (*M* = .68, *SEM* = .02) also performed marginally better than those in the generative condition (*p* = .055). However, there was no difference in test performance between students in the no notes control condition and the summary condition (*p* = .783). Furthermore, the interaction between instruction condition and question type was also nonsignificant (*F* < 1).

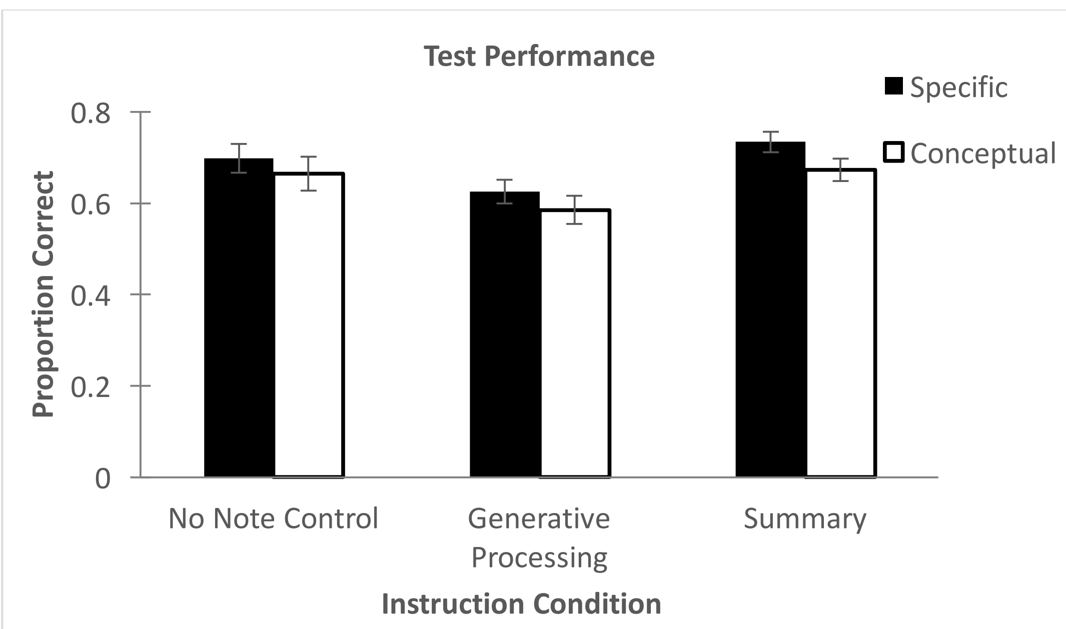

Figure 4. Mean proportion of correct test answers based on instruction condition and question type. The error bars reflect between-subjects standard errors of the mean.

A pairwise comparison between student test performance in the control condition in Experiment 1 (no note no report) and student test performance in the no notes control condition in Experiment 2 found that although there was a numerical difference in test performance between students in the two conditions, the difference was not statistically reliable, $p = .121$ (*see Figure 5*). This result is unsurprising because the two conditions are identical in procedure and data from the two experiments were collected concurrently. However, the students who summarized the lecture material in Experiment 2 performed significantly better than did the students in the no note no report control condition in Experiment 1, $p = .031$, providing some evidence that the note-taking effect was exhibited in Experiment 2.

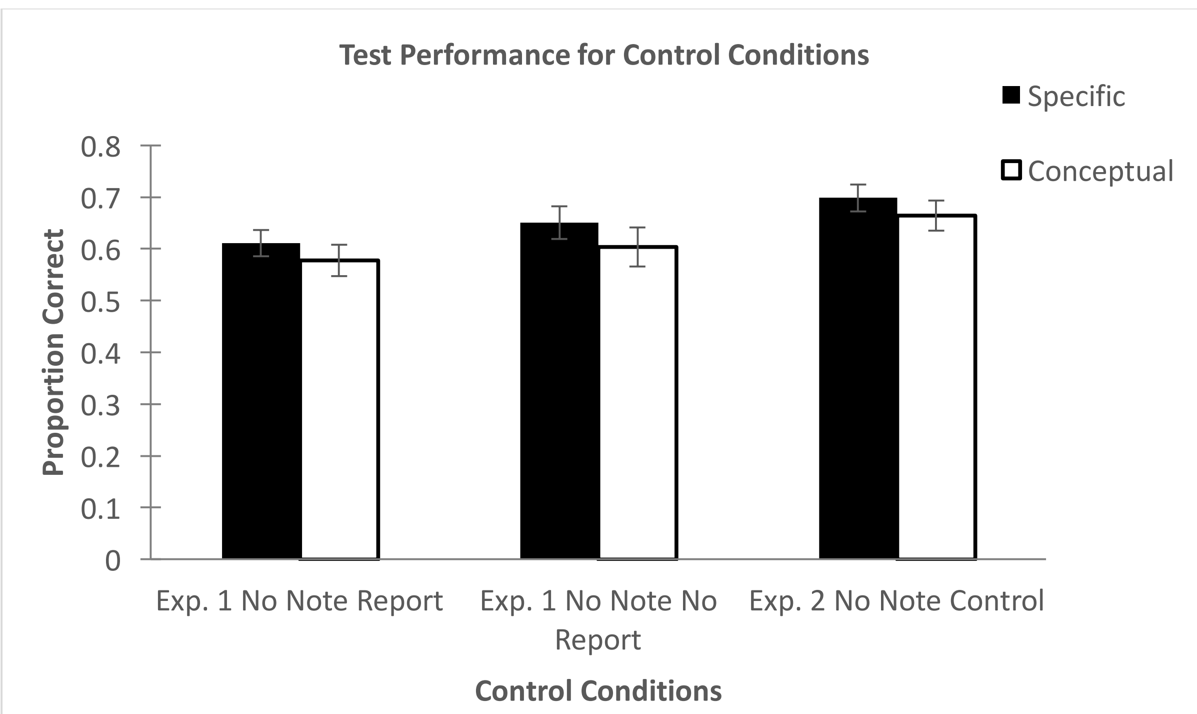

Figure 5. Mean proportion of correct test answers for the control conditions across experiments by question type. The error bars reflect between-subjects standard errors of the mean.

Note Content

The same three measures of note content from Experiment 1 were used in Experiment 2 (i.e. verbatim overlap, note style, and LSA). However, because this experiment manipulated what the students wrote in their notes, the content analysis was primarily used to verify that the manipulation was successful.

A series of separate one-way ANOVAs were conducted on the different measures of verbatim overlap (i.e., word count, 2-word verbatim sequences, and 3-word verbatim sequences) to examine the effects of instruction condition. There was a main effect of instruction condition on word count, $F(1,82) = 16.90$, $MSE = 8074.00$, $p < .001$, $\eta_p^2 = .17$ (*see Figure 6*). The students in the summary instruction condition wrote more words ($M = 252.98$, $SEM = 16.68$) than did the students in the generative processing condition ($M = 172.38$, $SEM = 10.31$). There were also main effects of instruction condition on both the proportion of 2-word verbatim sequences, $F(1,82) = 20.66$, $MSE = .046$, $p < .001$, $\eta_p^2 = .20$, and the proportion of 3-word verbatim sequences, $F(1,82) = 39.81$, $MSE = .002$, $p < .001$, $\eta_p^2 = .33$ (*see Figure 7*). Students who wrote summaries had significantly more verbatim overlap (2-word: $M = .75$, $SEM = .03$; 3-word: $M = .05$, $SEM = .01$) than did students who wrote generative notes (2-word: $M = .53$, $SEM = .04$; 3-word: $M = .12$, $SEM = .01$). These findings provide evidence that the instruction manipulation was successful.

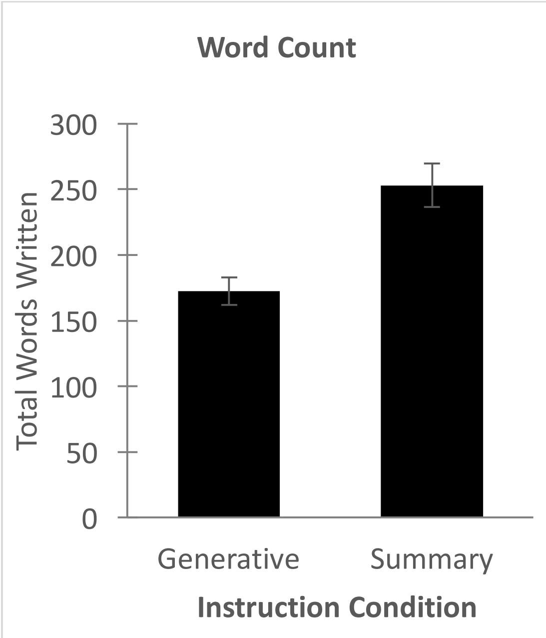

Figure 6.
Mean word count for each instruction condition. Error bars represent between-subjects standard errors of the mean.

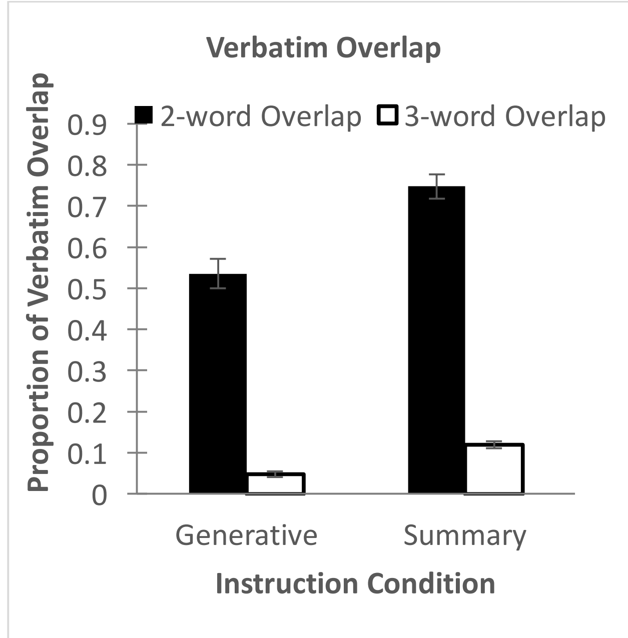

Figure 7.
Mean proportion of verbatim overlap for each instruction condition. Error bars represent between-subjects standard errors of the mean.

Pearson product-moment correlations were computed to assess the relationship between the amount of verbatim overlap and test performance (*see Table 6*). Word count had a marginally significant positive relationship with performance on specific questions, r = .18, n = 84, p = .100, and with overall test performance, r = .20, n = 84, p = .068. The more notes the students took, the better they performed on the specific questions on the retention tests. However, there was no evidence of a relationship between word count and performance on the conceptual questions, r = .17, n = 84, p = .119. There was also no evidence of a relationship between 2-word verbatim sequences and any measure of test performance. However, 3-word verbatim sequences were positively associated with performance on specific questions with marginal significance, r = .19, n = 84, p = .076, but not with either performance on conceptual questions, r = .11, n = 84, p = .340, or overall test performance, r = .17, n = 84, p = .125.

Table 6.
Correlation matrix of test performance and note content measures for all students. Values below the diagonal are Pearson correlation coefficients. Above the diagonal are the probabilities associated with the corresponding correlation coefficients.

	Specific	Concept	Total	Sum. Style	Gen. Style	Word Count	2-word Overlap	3-word Overlap	LSA
Specific	1	$p < .001$	$p < .001$	$p = .015$	$p = .048$	$p = .100$	$p = .219$	$p = .076$	$p = .392$
Concept	.54	1	$p < .001$	$p = .073$	$p = .091$	$p = .119$	$p = .865$	$p = .340$	$p = .036$
Total	.87	.89	1	$p = .017$	$p = .037$	$p = .068$	$p = .443$	$p = .125$	$p = .087$
Sum. Style	.26	.20	.26	1	$p < .001$	$p < .001$	$p = .001$	$p < .001$	$p = .477$
Gen. Style	-.22	-.19	-.23	-.79	1	$p = .090$	$p = .005$	$p < .001$	$p = .512$
Word Count	.18	.17	.20	.64	-.19	1	$p = .126$	$p = .283$	$p = .061$
2-word Overlap	.14	.02	.08	.36	-.31	.13	1	$p < .001$	$p = .138$
3-word Overlap	.19	.11	.17	.45	-.53	.12	.72	1	$p = .670$
LSA	-.09	-.23	-.19	.08	.07	.21	.16	.05	1

To investigate these relationships further, separate Pearson product-moment correlations were calculated between verbatim overlap and test performance for each condition. Performance on conceptual question was negatively correlated with both 2-word, $r = -.26$, $n = 42$, $p = .096$, and 3-word, $r = -.36$, $n = 42$, $p = .019$, verbatim sequences for students in the generative processing condition (*see Table 7*). Furthermore, there was also a marginally significant negative association between overall test performance and 3-word verbatim sequences, $r = -.27$, $n = 42$, $p = .086$. Thus, for students in the generative processing condition, more verbatim overlap was associated with worse test performance. In contrast, for students in the summary condition, there were marginally significant positive correlations between 3-word verbatim sequences and both performance on conceptual questions, $r = .26$, $n = 42$, $p = .098$, and overall test performance, $r = .29$, $n = 42$, $p = .058$ (see Table 8). Thus, higher proportions of verbatim overlap were associated with higher test performance for students in the summary condition.

Table 7.
Correlation matrix of test performance and note content measures for only the students instructed to take ***generative processing notes****. Values below the diagonal are Pearson correlation coefficients. Above the diagonal are the probabilities associated with the corresponding correlation coefficients.*

	Specific	Concept	Total	Sum. Style	Gen. Style	Word Count	2-word Overlap	3-word Overlap	LSA
Specific	1	$p < .001$	$p < .001$	$p = .945$	$p = .220$	$p = .191$	$p = .890$	$p = .445$	$p = .325$
Concept	0.60	1	$p < .001$	$p = .874$	$p = .478$	$p = .179$	$p = .096$	$p = .019$	$p = .122$
Total	0.89	0.89	1	$p = .899$	$p = .278$	$p = .137$	$p = .319$	$p = .086$	$p = .156$
Sum. Style	0.01	0.03	0.02	1	$p = .259$	$p = .434$	$p = .546$	$p = .985$	$p = .379$
Gen. Style	0.19	0.11	0.17	-0.18	1	$p < .001$	$p = .096$	$p = .567$	$p = .019$
Word Count	0.21	0.21	0.23	0.12	0.73	1	$p = .512$	$p = .683$	$p = .055$
2-word Overlap	-0.02	-0.26	-0.16	-0.10	0.26	0.10	1	$p < .001$	$p = .019$
3-word Overlap	-0.12	-0.36	-0.27	0.00	-0.09	-0.07	0.66	1	$p = .248$
LSA	.32	.12	.16	.38	.02	.05	.02	.25	1

Table 8.
*Correlation matrix of test performance and note content measures for only the students instructed to take **summary notes**. Values below the diagonal are Pearson correlation coefficients. Above the diagonal are the probabilities associated with the corresponding correlation coefficient.*

	Specific	Concept	Total	Sum. Style	Gen. Style	Word Count	2-word Overlap	3-word Overlap	LSA
Specific	1	*p* = .171	*p* < .001	*p* = .391	*p* = .476	*p* = .492	*p* = .960	*p* = .218	*p* = .720
Concept	.22	1	*p* < .001	*p* = .533	*p* = .706	*p* = .908	*p* = .528	*p* = .098	*p* = .093
Total	.68	.86	1	*p* = .360	*p* = .929	*p* = .657	*p* = .618	*p* = .058	*p* = .150
Sum. Style	-.14	-.10	-.14	1	*p* = .692	*p* < .001	*p* = .678	*p* = .270	*p* = .833
Gen. Style	.11	-.06	.01	-.06	1	*p* = .146	*p* = .408	*p* = .941	*p* = .295
Word Count	-.11	-.02	-.07	.79	.23	1	*p* = .181	*p* = .1781	*p* = .330
2-word Overlap	.01	.10	.08	-.07	.13	-.21	1	*p* < .001	*p* = .746
3-word Overlap	.19	.26	.29	-.17	.01	-.21	.61	1	*p* = .532
LSA	-.06	-.26	-.23	.03	.17	.15	-.05	-.10	1

A 2 x 2 mixed factorial ANOVA revealed an interaction between the instructions the students were given (summary, generative processing) and the type of style rating (summary rating, generative rating) on the magnitude of the rating, $F(1,82) = 497.80$, $MSE = 334.339$, $p < .001$, $\eta_p^2 = .86$ (*see Figure 8*). Students who were given the summary instructions had much higher summary style ratings ($M = 3.96$, $SEM = .14$) than generative style ratings ($M = 1.08$, $SEM = .04$). In contrast, students in the generative processing condition had much higher ratings for generative style ($M = 4.13$, $SEM = .18$) than for summary style ($M = 1.37$, $SEM = .07$). Because the style raters were blind to the participants' conditions, this finding confirms that the instruction manipulation was successful.

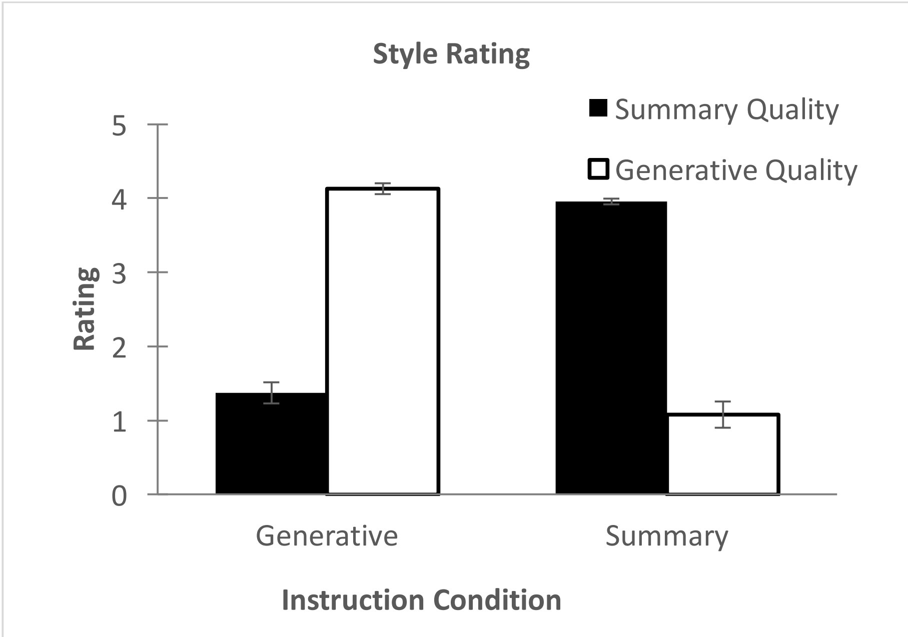

Figure 8. Mean note style ratings for each instruction condition. Error bars represent between-subjects standard errors of the mean.

However, a one-way ANOVA found no differences between the instruction conditions for the cosines extracted from the LSA, $F < 1$. Nevertheless, a Pearson product-moment correlation showed that there was a negative correlation between LSA cosine and performance on conceptual questions, $r = -.23$, $n = 84$, $p = .036$, and overall test performance, $r = -.19$, $n = 84$, $p = .087$ (*again see Table 6*). Students who had higher similarity cosines between their notes and the transcript of the lecture performed worse on the retention test.

Discussion

As mentioned before, the purpose of this second experiment was to disentangle the contribution of generative processing from the contribution of summarization to the note-taking effect. In this experiment, students who summarized the lectures were explicitly instructed to account for the main points of the lecture material and to be coherent (*see*

Appendix B). Students in the generative processing condition were instead directed to take notes on how the lecture material related to themselves, their lives, or their prior knowledge**.** If either of these two cognitive mechanisms underlies the note-taking effect, then students in at least one of the note taking conditions should perform better on a retention test than students in the no notes control condition. Summarization leads to the prediction that the students summarizing the lecture material will out perform the student in the other two conditions because the act of summarizing focuses their attention to the most pertinent information in the lecture. In contrast, generative processing leads to the prediction that the students in the generative processing condition will retain more information than those in the other two conditions because they are processing the information more deeply by relating it back to their prior knowledge.

Because the students were instructed to take notes using strategies that they might not use normally, the content of the notes was analyzed, using the same measures as in Experiment 1, to verify that the manipulation was successful. The significant interaction between the instructions given to the students and the type of style rating on the magnitude of the rating confirms that the students who were instructed to write summaries rated higher for summary style and those instructed to relate the material to their prior knowledge rated higher for generative processing. In addition, the students who wrote summaries wrote more words overall and had more verbatim overlap, both 2-word and 3-word sequences, than did the students in the generative processing condition. This finding is consistent with the students' use of selection summarization. Therefore, the instruction manipulation appears to have been successful.

Although the note-taking effect was not found comparing test performance of students in either of the note-taking conditions to performance of students in the no notes control condition, it appears that the sample of students in the no notes control condition was simply on the higher end of the distribution, although performance of students in this condition was not significantly different from performance of students in the no note no report control condition in Experiment 1. Because the data for Experiments 1 and 2 were collected concurrently from the same pool of participants, and there were no significant differences in performance between the procedurally identical control conditions, comparing the performance of the students in the summary condition to the performance of the students in the control condition in Experiment 1 does provide some evidence for a note-taking effect in the present experiment.

Furthermore, the students who summarized the lecture material retained more information than did those students who related the material back to their prior knowledge. This finding provides even more evidence that generative processing is not the pertinent cognitive mechanism underlying the note-taking effect. However, it does seem that the students in the generative processing condition who tried to relate the material to their prior knowledge and also include some verbatim information performed worse than did those who did not copy the information from the lecture, as suggested by the separate correlational analysis between the measures of verbatim overlap and test performance for students in each instruction condition. Specifically, higher levels of verbatim overlap were associated with lower test performance for students in the generative processing condition (although not for students in the summary condition).

GENERAL DISCUSSION

The purpose of the present investigation was to identify the pertinent cognitive processes underlying the note-taking effect. My Master's research (Lalchandani, 2016; Lalchandani & Healy, 2017) helped to define the scope of the current study. Because we had found that the better performance for longhand note takers was attributable to the act of taking notes and not the content of the notes, I chose to focus in the present investigation on the process function of note taking, not the product function. The note-taking effect was most clearly found after a short, one-half hour retention interval, as opposed to a long, 1-week retention interval. It was also most apparent when students took handwritten notes. Therefore, the present study used these conditions in an attempt to elicit the strongest note-taking effect in order to disentangle the unique cognitive mechanisms.

Of the research that has focused on cognitive processing during note taking, the findings fell under three primary hypotheses: the generative processing hypothesis, the attention hypothesis, and the cognitive effort hypothesis. After evaluating each hypothesis for overlapping and unique cognitive processes, I concluded that there were only three unique features among the different hypotheses: generative processing, summarization, and sustained attention. I conducted two experiments that were designed to distinguish the contributions of these three cognitive processes to the note-taking effect.

Generative Processing

The generative processing model of learning (Wittrock & Carter, 1975) presumes that information is processed at a deeper level (Craik & Lockhart, 1972) when the learner actively constructs associations between the to-be-learned material and his or her prior knowledge than when simply listening to or reading the novel material. In relation to note taking, the level of generative processing is thought to be influenced by the type of notes taken and the note-taking strategy used. Specifically, notes that require generative processing include summarizing, paraphrasing, outlining, or concept mapping (e.g., Bui et al., 2013; Lalchandani, 2016; Lalchandani & Healy, 2017; Mueller & Oppenheimer, 2014). However, this definition of generative processing confounds the distinct cognitive processes of generative processing (i.e., the process of associating novel information with prior knowledge) and summarization. Under the original model of generative processing, summaries required students to use their own words and connect across concepts (Wittrock & Alesandrini, 1990). However, most research regarding summarization examines selection summaries.

Because generative processing is the most cited as the mechanism underlying the note-taking effect, the present two experiments were designed to distinguish generative processing from the other two processes. Experiment 1 in the current investigation was designed to identify the effects of sustained attention and generative processing through the measurement of distraction. The current Experiment 2 separated generative processing and summarization by manipulating what the students wrote in their notes. The two experiments together tested both internally elicited generative processing through reported distraction and externally elicited generative processing with explicit note-taking instructions.

The results of both experiments lead to the conclusion that generative processing does not underlie the note-taking effect; in fact, there was evidence that generative processing is detrimental to the retention associated with note taking. Under the generative processing hypothesis, students taking notes should have reported more task-relevant distracted thoughts than students not taking notes in Experiment 1, and students who explicitly related the material to themselves and their prior knowledge should have performed better on the retention test than did students who wrote summaries of the lecture information in Experiment 2. Nevertheless, students who did not take notes reported more task-relevant distraction than did students who took notes, and students who wrote summaries performed better on the retention test than did those who actively associated the novel lecture material to their prior knowledge. These findings signify instead that note taking appears to help students focus their attention on the lecture material and that summarization facilitates retention. Therefore, it can be reasoned from these experiments that generative processing is not the pertinent cognitive process underlying the note-taking effect, despite being the most popular explanation for the benefits of note taking. However, although not responsible for the benefits of note taking, generative processing does seem to aid in retention for students not taking notes, as evidenced in the correlational analyses in Experiment 1. Specifically, there were significant positive associations between task-relevant distraction and performance both on specific test questions and on overall test performance, and there was also a marginally significant positive relationship between task-relevant distraction and performance on conceptual test questions, but only for the students who did not take notes in Experiment 1.

Summarization

The process of summarizing is thought to help students focus their attention on the most relevant information, thus improving both comprehension and retention (Anderson & Armbruster, 1984). Summarization, in this context, refers to summaries that involve the selection of relevant words and information from the lecture, the deletion of extraneous words and information, and the modification of the information to create a coherent synopsis. These summaries differ from other definitions that involve generative processing because they do not require the students to use their own words or connect across concepts (King, 1992; Wittrock & Alesandrini, 1990). Therefore, the present study removes the confounding between generative processing and summarization and examines each process for its unique effect on retention.

In Experiment 2, students either were instructed to write summaries that required the students to select, delete, and modify the information to produce a coherent summary or were instructed to explicitly relate the lecture material to themselves and their prior knowledge. Although students who summarized the lecture material performed better than did students who related the material to their prior knowledge, they did not outperform the students in the no notes control condition. However, it is possible that the sample of students in the control condition in Experiment 2 performed better than expected. A comparison of test performance between students in the control condition in Experiment 1 and the control condition in Experiment 2 found no differences between the two groups. However, the students who wrote summaries in Experiment 2 did perform better on the retention test than did the students in the control condition of Experiment 1. From this finding, one might infer that summarization underlies note-taking benefits.

Furthermore, the analysis of the content of the notes from Experiment 1 provides supporting evidence for summarization as a contributor to the note-taking effect when students take notes using their preferred note-taking strategies. Students were more likely to use a summary style of note taking when left to their own devices than a generative style. This finding is particularly informative because many of the investigations of the note-taking effect manipulated the type of notes the students took, thereby decreasing ecological validity (e.g., Bui et al., 2013). Moreover, the rating of summary style was positively correlated with test performance. These findings together lend support to the hypothesis that summarization is one of the primary cognitive processes underlying the note-taking effect.

Attention

The sustained attention hypothesis of the note-taking effect asserts that note taking forces the student to selectively focus on the lecture material and thereby process the information at a deeper level than just listening (Peper & Mayer, 1986). The cognitive antidote principle (Kole et al., 2008) functions in tandem with sustained attention. A cognitive antidote is an additional cognitive task that serves as an antidote to boredom. Both sustained attention and the cognitive antidote principle could be apparent through reduced mind wandering and/or increased task engagement.

Experiment 1 required students to report two different types of mind wandering, task-relevant and task-irrelevant. Task-irrelevant thoughts were those that in no way related to the lecture material. If students had thoughts that stemmed from the lecture material or related the material to themselves or their prior knowledge, they reported them as task-relevant thoughts. The findings of this experiment provide evidence that note taking functions as a cognitive antidote to boredom by decreasing the amount of mind wandering the student experiences. Moreover, there was no relationship between either type of distracted thought and test performance for those students who took notes. This finding is likely due to the fact that the students who took notes experienced fewer distracted thoughts overall, presumably because they were paying attention to the lecture material. Therefore, it can be surmised that sustained attention is also one of the pertinent cognitive processes underlying the note-taking effect.

Future Directions

Despite the strong conclusions from the present experiments, additional research must be conducted in order to distinguish the separate effects of summarization and sustained attention. Because generative processing was the most favored hypothesis explaining the note-taking effect, this investigation pitted it against the other two processes, presuming generative processing would be more influential than either summarization or attention. Unexpectedly, generative processing appeared to be detrimental to retention. Thus, the effects of both summarization and sustained attention need to be directly compared to each other. However, this comparison is not trivial as the two processes are thoroughly intertwined. One possible follow up experiment might include a condition providing breaks in the lecture period for student to summarize the information. This manipulation could separate summarization and sustained attention, as the students would exhibit summarization processing only when they were not focusing their attention on the lecture.

Further research also should investigate the effects of different types of selection summaries. In Experiment 2, students were asked to use complete sentences and to be coherent. Many of the students, although they focused on putting together coherent ideas, did not use complete sentences. Instead, they seemed to omit many high frequency words and articles. This type of summarization still requires students to select and delete from the lecture information, but it appears to necessitate less modification. Students in Experiment 1 seem to use a similar strategy as seen in the levels of verbatim overlap. One future experiment might directly compare the effects of the two types of selection summaries (one type that requires proper sentence structure and one type that does not) on retention and comprehension. Another possible experiment might directly compare the content of each of these types of selection summaries with notes taken “in the wild” (i.e., notes taken without instruction). This experiment would examine the effects of instructed summarization compared to an ecologically valid control. An experiment such as this one is important because the note-taking effect is apparent regardless of strategy. Therefore, it is essential to examine the moderating effect of note-taking instruction.

Nevertheless, the conclusions of this investigation have the potential to inform both pedagogical practices and the design of educational technology. As previously mentioned, most research has credited generative processing in explanations of the note-taking effect. However, it can be concluded that generative processing during note taking is detrimental to retention. Furthermore, teachers can now emphasize the importance of both sustained attention and summarization. Likewise, new technology could be designed to facilitate attention and/or summarization. There are ample digital note-taking applications in existence, nonetheless most of the most popular applications focus on the user interface and organizational features, not the effectiveness of the notes to be taken. This research provides an additional dimension for a developer to consider when designing new technology.

References

Anderson, T. H., & Armbruster, B. B. (1984). Studying. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research*, (Vol. 1, pp. 657-679). White Plains, NY: Lawrence Erlbaum Associates.

Bjork R.A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Bretzing, B. H., & Kulhavy, R. W. (1979). Notetaking and depth of processing. *Contemporary Educational Psychology*, *4*, 145-153. doi: 10.1016/0361-476X(79)90069-9

Bretzing, B. H., & Kulhavy, R. W. (1981). Note-taking and passage style. *Journal of Educational Psychology*, *73*, 242-250. doi: 10.1037/0022-0663.73.2.242

Bui, D. C., & McDaniel, M. A. (2015). Enhancing learning during lecture note-taking using outlines and illustrative diagrams. *Journal of Applied Research in Memory and Cognition*, *4*, 129-135. doi: 10.1016/j.jarmac.2015.03.002

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

Butler, A. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1118-1133. doi: 10.1037/a0019902

Chapman, M. J., Healy, A. F., & Kole, J. A. (2016). Memory load as a cognitive antidote to performance decrements in data entry. *Memory*, *24*, 1182-1196. doi: 10.1080/09658211.2015.1086380

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

Crawford, C. (1925). Some experimental studies of the results of college note-taking. *Journal of Educational Research, 12,* 379-386. doi: 10.1080/00220671.1925.10879612

Di Vesta, F. J., & Gray, G. S. (1972). Listening and note taking. *Journal of Educational Psychology, 63,* 8-14. doi: 10.1037/10.1037/h0032243

Gabrieli, J. (2011). 9.00SC Introduction to Psychology, Fall 2011. Massachusetts Institute of Technology: MIT OpenCourseWare. http://ocw.mit.edu (Accessed). License: Creative Commons BY-NC-SA

Healy, A. F., Jones, M., Lalchandani, L. A., & Tack, L. A. (2017). Timing of quizzes during learning: Effects on motivation and retention. *Journal of Experimental Psychology: Applied*, *23*, 128. doi: 10.1037/xap0000123

Hembrooke, H., & Gay, G. (2003). The laptop and the lecture: The effects of multitasking in learning environments. *Journal of Computing in Higher Education*, *15*, 46-64. doi: 10.1007.BF02940852

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kiewra, K. A. (1985). Investigating notetaking and review: A depth of processing alternative. *Educational Psychologist, 20,* 23–32. doi: 10.1207/s15326985ep2001_4

Kiewra, K. A. (1987). Notetaking and review: The research and its implications. *Instructional Science*, *16*, 233-249. doi: 10.1007/BF00120252

Kiewra, K. A. (2016). Note taking on trial: A legal application of note-taking research. *Educational Psychology Review*, *28*, 377-384. doi: 10.1007/s10648-015-9353z

King, A. (1992). Comparison of Self-Questioning, Summarizing, and Notetaking-Review as Strategies for Learning from Lectures. *American Educational Research Journal, 29*, 303-323. Retrieved from http://www.jstor.org/stable/1163370

Kobayashi, K. (2005). What limits the encoding effect of note-taking? A meta-analytic examination. *Contemporary Educational Psychology, 30,* 242–262. doi: 10.1016/j.cedpsych.2004.10.001

Kole, J. A., Healy, A. F., & Bourne, L. E. (2008). Cognitive complications moderate the speed-accuracy tradeoff in data entry: a cognitive antidote to inhibition. *Applied Cognitive Psychology*, *22*, 917-937. doi: 10.1002/acp.1401

Laham, D. (1998). Latent Semantic Analysis @ CU Boulder, http://lsa.colorado.edu/

Lalchandani, L. A. (2016). *Disentangling the effects of note-taking strategy: Generation and summarization* (Unpublished master's thesis). University of Colorado, Boulder, CO.

Lalchandani, L. A. & Healy, A. F. (2017). *A note-taking effect and differences in note-taking* *media*. Manuscript submitted for publication.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, *104*, 211-240. doi: 10.1037/0033-295X.104.2.211

McCurdy, M. P., Leach, R. C., & Leshikar, E. D. (2017). The generation effect revisited: Fewer generation constraints enhances item and context memory. *Journal of Memory and Language, 92,* 202–216. doi: 10.1016/j.jml.2016.06.0

McDaniel, M. A., & Butler, A. C. (2011). A contextual framework for understanding when difficulties are desirable. In A.S. Benjamin (Ed.) *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 175-198). New York & London: Psychology Press.

McFarland, C. E., Frey, T. J., & Rhodes, D. D. (1980). Retrieval of internally versus externally generated words in episodic memory. *Journal of Verbal Learning and Verbal Behavior*, *19,* 210-225. doi: 10.1016/S0022-5371(80)90182-6

Mueller, P. A., & Oppenheimer, D. M. (2014). The Pen Is Mightier Than the Keyboard. *Psychological Science*, *25*, 1159–1168. doi: 10.1177/0956797614524581

Mueller, P. A., & Oppenheimer, D. M. (2016). Technology and note-taking in the classroom, boardroom, hospital room, and courtroom. *Trends in Neuroscience and Education*, *5*, 139-145. doi: 10.1016/j.tine.2016.06.002

Peper, R. J., & Mayer, R. E. (1986). Generative effects of note-taking during science lectures. *Journal of Educational Psychology*, *78*, 34-38. doi: 10.1037/0022-0663.78.1.34

Peverly, S. T., Garner, J. K., & Vekaria, P. C. (2014). Both handwriting speed and selective attention are important to lecture note-taking. *Reading and Writing*, *27*, 1-30. doi: 10.1007/s11145-013-9431-x

Piolat, A., Olive, T., & Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology, 19,* 291-312. doi: 10.1002/acp.1086

Roussey, J. Y., & Piolat, A. (2003). Prendre des notes et apprendre. Effet du mode d’accès à l’information et de la méthode de prise de notes [Taking notes and learning: Effects of the access mode to information and to the note taking procedure]. *Arob@ se 7*, 47-68.

Seward, S.S. (1910). Note-taking. Allyn and Bacon: Chicago, IL.

Slamecka, N. J. & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 592–604.

Stacy, E. M., & Cain, J. (2015). Note-taking and handouts in the digital age. *American Journal of Pharmaceutical Education*, *79*, 107. doi: 10.5688/ajpe797107

Szpunar, K. K., Jing, H. G., & Schacter, D. L. (2014). Overcoming overconfidence in learning from video-recorded lectures: Implications of interpolated testing for online education. *Journal of Applied Research in Memory and Cognition*, *3*, 161-164. doi:10.1016/j.jarmac.2014.02.001

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 110,* 6313-6317. doi: 10.1073/pnas.1221764110

Tyler, S. W., Hertel, P. T., McCallum, M. C., & Ellis, H. C. (1979). Cognitive effort and memory*. Journal of Experimental Psychology: Human Learning and Memory, 5,* 607-617. doi: 10.1037/0278-7393.5.6.607

Wittrock, M. C., & Alesandrini, K. (1990). Generation of summaries and analogies and analytic and holistic abilities. *American Educational Research Journal*, *27*, 489-502. Retrieved from http://www.jstor.org/stable/1162933

Wittrock, M. & Carter, F. (1975). Generative processing of hierarchically organized words. *American Journal of Psychology, 88,* 489-501. doi: 10.2307/1421780

Appendix A

Retention Test Used in Both Experiments

Specific questions are *italicized* and conceptual questions are in plain text. Correct answers are in **bold**.

1. A person with a frontal lobe injury may have trouble doing what in the Wisconsin Card Sorting test?
    a. Verbalizing their choice
    b. Having an emotional response to a family member
    c. Switching rules
    d. Beginning a course of action

2. Who coined the term cognitive dissonance?
    a. Festinger
    b. Cornell
    c. Gabrielli
    d. Brehm

3. Patients with orbitofrontal lesions have what kind of galvanic skin response to pictures of family when compared to pictures of strangers?
    a. Low response
    b. High response
    c. The same (high response)
    d. The same (no response)

4. A person with ventromedial lesions walks into a building and sees a nail, a hammer, and a picture. This person hangs the picture. This is what kind of behavior?
    a. Perseverative errors
    b. Utilization behavior
    c. Functional fixedness
    d. Modality effect

5. A student rates a series of items on a scale from 1-10 of how much he/she likes them. The researcher picks out two items that were similarly ranked near the middle and allows the student to choose one. When asked to re-rate the items, how would the student re-rate that item he/she picked?
    a. Rated better than before
    b. Rated worse than before
    c. Rated more in the middle
    d. No change in rating

6. *What does the professor describe as the business end of the brain?*
    a. Temporal Lobe
    b. Parietal Lobe
    c. Occipital Lobe
    d. Frontal Lobe

7. According to cognitive dissonance theory, what do people seek to avoid?
    a. Loud noises
    b. Excessive stimulation
    c. Contradictions between behaviors and beliefs
    d. Negative criticism

8. A student is accepted into CSU for psychology, but is not accepted into CU, the student's top choice school. When asked to re-rate the schools, how would you predict the student to rate CSU?
    a. Rated better than before
    b. Rated worse than before
    c. Rated more in the middle
    d. No change in rating

9. *What measure has been used in polygraphs?*
    a. Wisconsin Card Sorting Test
    b. Perseverative errors
    c. Galvanic skin response
    d. ERPs

10. A person with lesions in the prefrontal cortex may have difficulty with mental flexibility and what?
    a. Functional Fixedness
    b. Ambiguity Error
    c. Mental Filtering
    d. Perseverative errors

11. *What psychological aspect can be measured using sweat gland response?*
    a. Utilization behaviors
    b. Psychopathy
    c. Emotional reactivity
    d. Lateral hypothalamic function

12. *In Brehm's experiment, when asked to re-rate the wedding gifts, how did the rating of the item the subject chose change?*
    a. **Rated better than before**
    b. Rated worse than before
    c. Rated more in the middle
    d. No change in rating

13. The decision to help someone is made in which part of the brain?
    a. Temporal Lobe
    b. Parietal Lobe
    c. Occipital Lobe
    d. **Frontal Lobe**

14. *In the Festinger's original experiment, who rated the task as truly most enjoyable?*
    a. True enjoyment was not measured
    b. They were the same
    c. **The people who were paid $1**
    d. The people who were paid $20

15. Which is an example of a beneficial utilization behavior?
    a. Injecting yourself at the doctor's office
    b. Hammering in nails
    c. Sweating
    d. **Walking**

16. *What were the women in the 1959 study told they must do in order to qualify for the research study?*
    a. **Read aloud a list of obscene words**
    b. Pay for the study
    c. Attend a very boring lecture about the mating habits of animals
    d. Show up

17. A student rates a series of items on a scale from 1-10 of how much he/she likes them. The researcher picks out one item to give to the student. When asked to re-rate the items, how would the student re-rate that particular item?
    a. Rated better than before
    b. Rated worse than before
    c. Rated more in the middle
    d. **No change in rating**

18. *What is an example of a behavioral task testing stopping a course of action?*
    a. **Wisconsin Card sorting test**
    b. Stroop task
    c. Memory test
    d. N-Back task

19. A student rates a series of items on a scale from 1-10 of how much he/she likes them. The researcher picks out two items that were similarly ranked near the middle and allows the student to choose one. When asked to re-rate the items, how would the student re-rate that item he/she not pick?
    a. Rated better than before
    b. **Rated worse than before**
    c. Rated more in the middle
    d. No change in rating

20. Psychopaths would have what kind of galvanic skin response to the sound of an air horn when compared to a control group?
    a. High response
    b. Low response
    c. **The same (high response)**
    d. The same (no response)

21. *Being unable to stop following a rule that worked initially, even though it no longer applies is called what?*
    a. Mental Filtering
    b. **Perseverative errors**
    c. Ambiguity Error
    d. Functional Fixedness

22. A patient who extremely over estimates or extremely underestimates how many elephants are in Utah may have what kind of brain damage?
    a. Hemispheric damage
    b. Parietal damage
    c. Thalamic damage
    d. **Frontal cortical damage**

23. *Our feelings about the gap between our attitudes and our actions is commonly* known as what?
    a. The daffodil effect
    b. **Cognitive dissonance**
    c. The gap effect
    d. The testing effect

24. Mental flexibility depends on which part of the brain?

a. Hippocampus
b. Prefrontal cortex
c. Thalamus
d. Cerebellum

Appendix B

Instructions for Each Note-Taking Condition in Experiment 2

Generative Processing Notes

In this experiment, you will watch 2 lectures taken from a psychology class at MIT. During the lecture time, please take notes on how the lecture material relates to you, your life, or your prior knowledge (e.g., your own experiences or something you learned in a class).

Please pay attention to the lecture material as if this was a regular class where you might take a test on the material.

Summary Notes

In this experiment, you will watch 2 lectures taken from a psychology class at MIT. During the lecture time, please take notes by summarizing the lecture on the provided paper. Your summary should account for the main points of each lecture, use full sentences, be somewhat brief, and be coherent.

Please pay attention to the lecture material as if this was a regular class where you might take a test on the material.