On the mnemonic benefits of drawing

by

Jeffrey D. Wammes

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Psychology

Waterloo, Ontario, Canada, 2017

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Dr. Morgan Barense Associate Professor
Supervisor(s)	Dr. Myra Fernandes Professor
Internal Member	Dr. Michael Dixon Professor
	Dr. Daniel Smilek Professor
Internal-external Member	Dr. Ben Thompson Associate Professor

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Researchers have long sought to explore means through which the amount, and quality of information one is able to remember can be improved. Attention has been paid variously to multimodal encoding, visual learning, and deep elaborative processing of to-be-remembered information in pursuit of this goal. Drawing is a relatively easily implemented task that engages many of the foregoing, and therefore is an excellent candidate task for systematic exploration. In this dissertation, the efficacy of drawing is explored as a potentially potent, memory-bolstering encoding strategy, and possible explanatory mechanisms are explored. In Chapter II, drawing at encoding is compared to writing, and leads to better free recall. Drawing is also contrasted with other tasks, indicating that the benefit cannot be fully explained by elaborative encoding, visual imagery, or picture superiority. Drawing led to a large boost in free recall in both within- and between-subjects designs, thereby ruling out within-list distinctiveness as a comprehensive explanatory mechanism. In Chapter III, drawing and writing are again compared, this time using recognition task variants, revealing that the memory benefits of drawing are driven primarily by detailed, recollection-based memory. A component model of the benefit drawing provides to memory is proposed, which holds that the effect of drawing is driven by the confluence of three beneficial components required in drawing – an elaborative component, to derive an internal concept of what to draw based on a target word; a motor component, to translate that internal image to the page using a specific manual motor program; and a pictorial component, derived from one's newly created depiction of the original target word. In Chapter IV, this component model is tested using tasks which systematically do or do not require each of the requisite components of the drawing process. Results indicate that each component is an important contributor to the benefit, and that memory performance scales up as the number of components increases. Taken together, the experiments presented herein indicate that drawing is a potent encoding task which improves memory through the provision, and integration of rich contextual information from at least three multisensory components.

Acknowledgments

First and foremost, I am indebted to my supervisor, Dr. Myra Fernandes, for guiding me when I was a fledgling graduate student in the beginning, and trusting my judgment once I had things (slightly) more figured out. I also owe thanks to my collaborators on some of this work – Melissa Meade and Tanya Jonker - whose perspectives made this research easier, and much more fun, to conduct.

I would like to thank the research assistants – Iulia Banica, Brady Roberts, Robin Chang, Erin Kim, Liat Kofler, Terri Middleton, Grace Sim – who invested many hours assembling materials and collecting data in support of these experiments. I would also like to thank Dr. Britt Anderson, for not only provision of his own personal laptop for my later implementations of drawing experiments, but also for being generous with his time in helping me design the corresponding program. In this latter arena, I'd also like to thank Dr. Alex Filipowicz, Andriy Struk, and Syaheed Jabar for consulting with me in the early stages of programming. Finally, I owe considerable thanks to Marshall Taylor, one of my oldest friends, whose foray into drawing rather than writing his notes for undergraduate courses was the first inspiration for this work.

This research was funded by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) awarded to Myra A. Fernandes. It was also supported by a post-graduate scholarship from NSERC, awarded to author Jeffrey D. Wammes.

The experiments reported in Chapter II of this dissertation were published by myself, Melissa E. Meade and Myra A. Fernandes in the *Quarterly Journal of Experimental Psychology* (Wammes, Meade & Fernandes, 2016, see references). The experiments reported in Chapter III of this dissertation were published by myself, Melissa E. Meade and Myra A. Fernandes in

Journal of Experimental Psychology: Learning, Memory and Cognition (Wammes, Meade & Fernandes, in press, see references). The experiments reported in Chapter IV were conducted by myself, Tanya R. Jonker and Myra A. Fernandes, and will be submitted for publication shortly. Portions of this dissertation have been derived from or inserted verbatim from these manuscripts.

Table of Contents

List of Tables ix
List of Figures x
CHAPTER I: Introduction 1
Image Generation as an Encoding Tool 1
Image Generation in Educational Research 3
Evidence from Extant Encoding Strategies 5
The Present Work 7
CHAPTER II: Drawing Improves Free Recall 9
Experiment 1 10
Method 11
Results and Discussion 16
Experiment 2 18
Method 19
Results and Discussion 21
Experiment 3 23
Method 24
Results and Discussion 26
Experiment 4 27
Method 29
Results and Discussion 30
CHAPTER III: Drawing and Recollection 33
Experiment 5A 36
Method 38
Results and Discussion 39
Experiment 5B 43
Results and Discussion 45
Experiment 6 47
Method 48
Results and Discussion 49
CHAPTER IV: The Components of Drawing 53
Experiment 7 55

Method 56
Results and Discussion 59
Experiment 8 62
Method 63
Results and Discussion 65
CHAPTER V: General Discussion 70
Context and the Integrated Components Model 71
Relation to Skilled Drawing 74
Drawing Inferences Beyond Memory 75
Limitations and Future Directions 77
Conclusions 81
Footnote 83
References 84
Appendix 106

List of Tables

Table 1. Mean (Standard Error in Brackets) Accuracy, Hit Rate, False Alarm Rate, Proportion Remember (R) Responses, Sensitivity (d'), Overall, for R Responses Only, and for K Responses Only for Each Item Type, and Process Estimates for Each Item Type in Experiment 5A and 5B……………………………………..41

Table 2. Mean (Standard Error in Brackets) Hit Rate, False Alarm Rate, Sensitivity, and Response Time (RT) for Each Item Type in Experiment 6……………………….50

Table 3. Trial Types in Experiments 7 and 8, Indicating the Presence (Shaded) or Absence (White) of the Elaborative (Elab), Motor (Mot), and Pictorial (Pic) Components, as Well as a Short Description of Each Task………………………………………55

Table 4. Mean (Standard Error in Parentheses) Hit Rate, False Alarm Rate, Sensitivity, and Response Time (RT) for Each Item Type in Experiment 7……………………….60

Table 5. Mean (Standard Error in Parentheses) Hit Rate, False Alarm Rate, Sensitivity, and Response Time (RT) for Each Item Type in Experiment 8……………………….66

List of Figures

Figure 1. Examples of drawings created in Experiments 1A and 1B using pencil and paper…………………………………………………………………………......15

Figure 2. Proportion of words recalled out of 15 in Experiments 1A and 1B, from the draw and write trial types. Error bars represent the standard error of each mean………17

Figure 3. Proportion of words recalled out of 10 in Experiments 2A, 2B and 2C, from the draw, list, visualize, view and write trial types. Error bars represent the standard error of each mean………………………………………………………………..23

Figure 4. Mean proportion of words recalled out of 33 for each trial type in Experiment 3 (Mixed-lists) and out of 66 for each trial type in Experiment 4 (Pure-lists). Error bars represent the standard error of each mean…………………………………...31

Figure 5. Sample drawings in Experiments 5B, 6, 7 and 8, created using a stylus on the touchscreen tablet ………………………...…………………………………...44

Figure 6. Recollection process estimates in Experiments 5A and 5B in the Draw and Write trial type. Error bars represent standard error of the mean………………………46

Figure 7. Sensitivity (d') in recognition memory decisions in Experiment 6 when responses were at control pace, relative to speeded. Error bars represent standard error of the mean…………………………………………………………………………...51

Figure 8. Sensitivity (d') in recognition memory decisions in Experiments 7 and 8, for the Draw, Blind, Trace, Imagine, View and Write trial types. Error bars represent standard error of the mean……………………………………………………..67

Figure 9. The integrated components model of the drawing effect, with Experiment 8's trial types overlaid. In this model, the beneficial effects of drawing, over and above basic verbal memory ('v') are driven by the integrated contributions of elaborative, motoric, and pictorial information. …....…………………………………………72

CHAPTER I: Introduction

> "Whatever is experienced simultaneously or in immediate succession is conceived in one act of consciousness and by that means its elements are united."
> (Ebbinghaus, 1885, p. 91)

Much research has been devoted to the exploration of techniques, actions or tasks through which one might bolster the amount, vividness and accuracy of information one can commit to memory. This often entails modifying intrinsic stimulus qualities, such as including associated images or sounds alongside words (Wheeler & Buckner, 2003, 2004; Luo, Hendriks & Craik, 2007), or asking that participants study emotional (Charles, Mather & Carstensen, 2003; Kensinger & Corkin 2003; Kleinsmith & Kaplan, 1964) or bizarre words (Jacoby & Craik, 1979; Waddill & McDaniel, 1998; Stein, Littlefield, Bransford, & Persampieri, 1984). However, because the nature of the information we want to remember is typically beyond our control – for instance, we cannot decide whether an item is a picture or not – attention has also been paid to encoding tasks or strategies which might bolster our ability to encode and maintain information, regardless of its intrinsic qualities, in memory. Previous work has indicated that many tasks performed at encoding, including production (MacLeod, Gopie, Hourihan, Neary & Ozbuko, 2010), generation (Slamecka & Graf, 1978), and enactment (Guttentag & Hunt, 1988), reliably improve later memory performance, relative to their more passive respective baselines. Pulling on the same thread as these previous researchers, who explored the anecdotal notion that there is an inherent advantage to learning by *doing*, I sought to determine whether drawing could also provide a measurable memorial advantage, over and above that of writing or passive note-taking.

Image Generation as an Encoding Tool

While there are countless popular science books that will extoll the virtues of using drawing to more effectively think, learn, remember and convey one's ideas (e.g. Agerbeck, 2012;

Brown, 2014; Roam, 2013), there is surprisingly little controlled psychological research testing the merit of drawing as a mnemonic. An early allusion to the possible benefit of drawing in bolstering memory was delineated by Paivio and Csapo (1973) in two experiments exploring the role of dual-coding in driving the picture superiority effect (Hockley, 2008; Kinjo & Snodgrass, 2000; Maisto & Queen, 1992; Paivio & Foth, 1970; Paivio, Rogers, & Smythe, 1968; Paivio, 1969, 1971; Snodgrass & McLure, 1975). The picture superiority effect is the finding that pictures are reliably better remembered than words, and the dual-coding theory holds that this effect manifests because pictures can be encoded as both the visual image, and its verbal label (Paivio, 1971). Accordingly, when the time arrives for one to recall the encoded information, both the verbal and visual codes can be used as cues for recall. In one of their experiments, Paivio and Csapo (1973) tested memory for both words and pictures that were either drawn, or their verbal label written, at encoding. Their results indicated that all of the conditions which included either viewing or drawing a picture led to better later recall than simply writing a word. This led the authors to conclude that picture superiority effects are best explained by the additive contributions of a verbal code and a visual or imaginal code to memory (Paivio & Csapo, 1973).

Later work (Peynircioglu, 1989) reincarnated the study of drawing as a facilitator of later memory, with a focus on testing memory for images of objects, scenes or nonsense figures. Interestingly, in lieu of more typical retrieval tasks, participants were required to later draw the studied pictures from memory, and the accuracy of these images was scored. The baseline conditions used were variously rating the quality of a presented image, copying a presented image, or connecting the dots in numbered order to create an image. In all cases, Peynircioglu (1989) demonstrated beneficial effects associated with generating one’s own drawing during encoding. While this finding seems to indicate that drawing improves later memory, the effects may be attributable to transfer appropriate processing (TAP; Morris, Bransford & Franks, 1977). That is, the accuracy of the image one produced at recall was the yardstick by which memory was measured. Based on the principles of TAP, re-drawing items at recall that were originally drawn at encoding would likely lead to a memory boost simply as a result of the similarities between the encoding and retrieval tasks, irrespective of the fact that they were drawn.

Image Generation in Educational Research

More recent research on the matter has largely been constrained to the educational literature, where the merit of image generation has been explored extensively as a learning tool. This is not particularly surprising, as it has long been known that presenting information in the pictorial rather than, or in addition to, the verbal modality leads to boosts in memory (e.g. McBride & Dosher, 2002; Paivio et al., 1968; Weldon & Coyote, 1996; Whitehouse, Mayberry, & Durkin, 2006). This benefit for individual items has also been extended to more complex educational materials, where the inclusion of conceptual diagrams leads to reliably better retention of the associated material (Ainsworth & Loizou, 2003; Arnold & Dwyer, 1975; Carney & Levin, 2002; Mayer & Moreno, 2003; Reid, 1984; Rigney & Lutz, 1976; Scaife & Rogers, 1996; Shah & Hoeffner, 2002). If the inclusion of graphical information improves student retention, it follows that the active *creation* of graphical information might further improve retention. Many educational studies have explored this idea. However, the studies vary greatly in their methodologies, and often include complex manipulations which lead to difficulties in ascribing performance benefits solely to drawing in its most basic form (i.e. drawing without additional manipulations or instructions). As an example, some studies have incorporated conditions wherein first-graders are provided with backgrounds and cutouts, which are then to be assembled into a representation of a provided narrative (Lesgold, DeGood & Levin, 1977;

Lesgold, Levin, Shimron & Guttman, 1975). In this particular case, retention was only improved if the representation was assembled *for* the student by a facilitator, or if they were given *only* the necessary cutouts. With this in mind, these studies do not provide particularly strong evidence for the profitable use of *free-form* image generation.

Other work has precipitated mixed outcomes with respect to image generation, finding either benefits (Alesandrini, 1981; Dean & Kulhavy, 1981; Greene, 1989; Schwamborn, Mayer, Thillman, Leopold & Leutner, 2010; Van Meter, 2001), costs (Leutner, Leopold & Sumfleth, 2009), or no effects at all (Hall, Bailey & Tillman, 1997; Kulhavy, Lee, Caterino, 1985; Rasco, Tennyson & Boutman, 1975; Snowman & Cunningham, 1975) as a result of drawing or image creation manipulations. Accordingly, there is no clear consensus within this literature, and it would be difficult to derive a firm conclusion given the massive variety that exists in manipulations, instructions, and control conditions. Potentially influential additional factors in these studies include providing the basic elements of a potential drawing and instructing students to assemble them rather than free-form draw in a manner of their choosing (Lesgold et al., 1975; 1977; Schwamborn et al., 2010; Van Meter, 2001), comparing drawing to a control condition of silent reading alone (Dean & Kulhavy, 1981), or asking participants to generate only one image for a lengthy passage (Alesandrini, 1981; Dean & Kulhavy, 1981; Hall et al., 1997; Rasco et al., 1975). As a result, few of the studies in this literature incorporate a *pure* drawing manipulation. Of those that did include something resembling pure drawing, one showed a small benefit, but only when participants were instructed to make their drawings holistic instead of analytic in nature (Alesandrini, 1981). In another, participants were exposed to diagrams within the studied text, and provided with extensive training on how these diagrams should be drawn (Van Meter, 2001). Their training included rehearsal of rules and conventions for creating good diagrams, practice trials, and extensive experimenter feedback on their diagrams. Complicating things further, participants spent nearly twice as much time completing the ‘draw’ condition as they did completing the ‘read’ control. This difference in encoding time is problematic, as the amount of time spent with to-be-remembered information is a factor that is known to boost memory substantially in itself (Cooper & Pantle, 1967).

Evidence from Extant Encoding Strategies

With the foregoing in mind, one of my aims in the current dissertation was to explore the efficacy of drawing in a relatively basic form (with no extra implements, instructions, or emphasis), as a tool for improving memory. Many tasks which use active manipulations to elicit some additional processing of an item at encoding demonstrably improve performance on later retrieval tasks. A classic example illustrating this premise comes from the deep levels of processing (LoP) framework and its associated experimental evidence. Specifically, Craik and Lockhart (1972) proposed that any encoding task that requires a greater amount of semantic processing (i.e. a deep LoP) would lead to superior performance when contrasted with a ‘shallower’ perceptual task (that is, a task that does not require semantic analysis). As an example, deciding whether a word fits into a provided sentence should lead to better memory than deciding whether the word was written in upper or lower case. This framework was later rigorously tested with a number of task combinations (e.g. pleasantness ratings compared to counting syllables), and the results were consistent with the ideas presented in the LoP framework (Walsh & Jenkins, 1973). In a similar approach, research has shown that memory is superior for words that are generated from a cue - often a definition of or a synonym for the target – relative to words that are simply read (Slamecka & Graf, 1978). This phenomenon is labelled ‘the generation effect’ and has been replicated hundreds of times in the years since its seminal observation (see Bertsch, Pesta, Wiscott, & McDaniel, 2007, for a review).

In addition to deep LoP and generation, there have been several other notable tasks that improve memory, simply through encouraging some active form of encoding that participants would not otherwise engage in. Among these is enactment, where participants perform a motor action in response to a presented word, phrase, or picture (Guttentag & Hunt, 1988; Engelkamp & Zimmer, 1997; Engelkamp, Zimmer & Biegelmann, 1993; Knopf, 1991; Knopf & Neidhardt, 1989; Senkfor, 2008). It has been suggested that through incorporating a movement, enactment brings online a motor code, which would not otherwise be available in the encoding of words that are simply read (Engelkamp, 2001), or alternatively, that binding the studied item with the action in one integrated encoding event is of critical importance in driving the benefit (Kormi-Nouri & Nilsson, 2001). Similarly, visual imagery reliably improves memory (Elliott, 1973; Kieras, 1978; Lupiani, 1977; Winnick & Brody, 1984) and the underlying mechanism is thought to be the same dual-coding that drives the picture superiority effect (e.g. Paivio & Csapo, 1973; Marschark, Richman, Yuille & Hunt, 1987). More recently, research has also shown that saying a word aloud, relative to reading it silently will improve memory (Bodner & MacLeod, 2016; Bodner & Taikh, 2012; MacLeod et al., 2010; Ozubko, Major & MacLeod, 2014). This ‘production effect’ (MacLeod et al., 2010) is thought to be driven by the extra ‘aloud-ness’ providing extra, distinctive information that aids in later recall.

While each of these seemingly disparate effects tends to invoke their own proposed explanatory mechanisms, there does seem to be a common thread that unites them in the same ‘family’. The common theme seems to be enhanced item processing, which involves some or all of the following: (1) processing the semantics of the item deeply, (2) adding additional multimodal information that would not otherwise be engaged by just viewing or reading an item, (3) bringing online imaginal or pictorial information. It follows from these common themes that drawing should also improve memory. I propose that this would be the case, because drawing a picture based on a word encourages the encoding of a deeply integrated memory trace including (1) elaborative information about not only the semantics of the word, but the appearance of its real-world referent, (2) motoric information resultant from translating that internal image to the page using coordinated hand movements, and (3) pictorial information that comes from viewing and processing one's final composition. Thus, these three components would be bound or integrated into a seamless engram, leading to improved memory. The correspondence between the components, and the proposed mechanisms underlying many other beneficial memory strategies is clear, so it stands to reason that drawing should be a particularly potent encoding strategy.

The Present Work

In this dissertation, I test the foregoing hypothesis, exploring whether drawing will improve memory substantially. Following that, I also systematically test several candidate mechanisms, including my own 'integrated components' model. In Chapter II, I contrast free recall performance following drawing, relative to following writing at encoding. This is done across multiple instruction variants, encoding times, and list lengths, as well as in both within- and between-participants designs. I also contrast drawing with other potent encoding manipulations (elaborative encoding, visual imagery, and pictorial presentation), to discern whether LoP, imagery, or picture superiority are sufficient as explanatory mechanisms. In Chapter III, I test whether drawing selectively improves recollection, as opposed to familiarity-based memory, to determine whether participants are drawing upon contextual information from the encoding experience when they attempt retrieval. Furthermore, as a byproduct of the design, I determine whether the beneficial effects of drawing generalize to recognition paradigms. In Chapter IV, I test the 'integrated components' model, to determine whether drawing enhances memory by integrating various sources of multimodal information that are manifest in the requisite stages of the drawing process.

CHAPTER II: Drawing Improves Free Recall

In Chapter II, I present my seminal studies testing whether drawing, relative to simply writing words out, leads to improved performance on a later free recall test. Drawing is a complicated, integrated process which requires one to create an elaborative internal image, and eventually produce a corresponding pictorial representation of an item, implemented using an intricate set of hand movements coordinated by the motor system. Not only that, but these three steps or processes must continuously and seamlessly interact with one another to successfully complete a drawing. Based on classic work in levels of processing (Craik & Lockhart, 1972), dual-coding (Paivio, 1969), and enactment (Guttentag & Hunt, 1988), and because drawing requires a confluence of steps or stages – all of which are analogous to these classic encoding tasks – I predicted that memory would be improved substantially by drawing. Such an idea has found preliminary support from previous work in dual-coding research (Paivio & Csapo, 1973), in copying pictures from memory (Peynircioglu, 1989), and more recently the idea has been co-opted in the educational literature, albeit with mixed results (see Van Meter & Garner, 2005, for a review). My first aim in this line of work was to determine whether drawing, performed at a relatively basic level (i.e. with no intensive instructions or training), could improve free recall performance substantially relative to writing information.

In the event that drawing did indeed improve memory, my aim was to rule out a mechanism that tends to account for the beneficial effects of many active encoding strategies. As an example, work in the production effect has often invoked a ‘distinctiveness account’ in explanation of the memorial benefits that reading aloud confers. Specifically, this account suggests that the benefit occurs only because produced items are rendered distinctive *relative* to the more weakly encoded, silently read items (Conway & Gathercole, 1987; Dodson & Schacter, 2001; Gathercole & Conway, 1988; Hopkins & Edwards, 1972; MacLeod et al., 2010). In other words, in a mixed list, reading one word aloud provides extra information, rendering it distinct from words that *were not* read aloud. Support for this account is provided by countless studies of subject-performed tasks wherein the beneficial effects of the mnemonic are manifest in within-participants, mixed-list designs, but tend to be reduced in blocked lists (Bertsch et al., 2007; Bodner, Taikh, & Fawcett, 2014), and smaller, notoriously difficult to detect, or absent entirely in between-participants designs (Begg & Roe, 1988; Huff, Bodner & Fawcett, 2015; Hunt & Elliot, 1980; Hunt & Mitchell, 1982; Jonker, Levene & MacLeod, 2014; Karpicke & Zaromb, 2010; McDaniel & Bugg, 2008; McDaniel & Einstein, 1986; Peterson & Mulligan, 2013; Serra & Nairne, 1993; Slamecka & Katsaiti, 1987). This also tends to be true of effects that depend more on stimulus properties, including bizarreness and orthographic atypicality (Hunt & Elliott, 1980; Hunt & Mitchell, 1982; McDaniel & Bugg, 2008; Waddill & McDaniel, 1998). Because many of the most prominent encoding strategies are blunted by a relatively commonplace design change, I also tested whether the beneficial effects of drawing could be neatly situated with its cousins under this same 'distinctiveness' umbrella. In this chapter, the experiments first establish the beneficial effects of drawing (Experiment 1), test several of the most obvious candidate mechanisms to explain the effect (Experiment 2). I then test whether the benefit applies to lengthier lists and shorter encoding times (Experiment 3) and finally, test the efficacy of 'distinctiveness' as an explanatory mechanism by manipulating drawing between-participants (Experiment 4).

Experiment 1

The aim of Experiment 1 was to determine whether, in a controlled and relatively simple paradigm, drawing would provide a benefit to later memory performance, relative to writing. The basic framework of the paradigm remains consistent across all experiments in this chapter, with only deviations in task instructions, list length, and timing. In this initial experiment, recall of incidentally encoded words was tested after a brief retention interval. In Experiment 1A, during incidental encoding, words were either drawn in detail, or written out multiple times. Based on previous work (Paivio, 1969; Paivio & Csapo, 1973), I predicted that despite the inherent advantage that production would afford written words (Forrin, MacLeod & Ozubko, 2012), drawn words would be better recalled than written ones.

I was concerned that instructing participants to add detail to their drawing, but not their writing, was implicitly biasing participants toward favoring retention of the drawn items. Further, it is possible that semantic satiation, or the loss of meaning that occurs with repeated exposure, would occur with repeated writing (Balota & Black, 1997). To control for this, the instructions were crossed over in Experiment 1B, such that the participants were to draw each item repeatedly, while they were to write only once with the instruction to add detail to the writing until time had elapsed. I predicted that the advantage of drawing would still manifest under these instructions.

Method

Participants. Participants for Experiment 1A were 30 undergraduate students (19 female), and for 1B were 25 undergraduate students (9 female) at the University of Waterloo, who completed the experiment for course credit or monetary remuneration. Participants ranged in age from 18 to 47 (M = 20.67, SD = 4.15), with between 14 and 27 years of education (M = 16.72, SD = 2.22). All participants had normal or corrected to normal vision, and learned English before the age of seven.

Materials.

Target items. An 80-item word list (See Appendix A) was created from a selection of the verbal labels for Snodgrass images (Snodgrass and Vanderwart, 1980), to ensure that all words could be easily drawn. Complex drawings were avoided (e.g. clown) in favor of simpler items (e.g. apple). This measure was taken to reduce the time it would take participants to create each of the drawings; every word could be drawn in the time provided based on a pilot study, and no item required excessive visual detail to be discernable. Words ranged in frequency between 1 and 25 (M = 8.23, SD = 6.44), in length between 3 and 11 letters, (M = 5.56, SD = 1.79), and in number of syllables from 1 to 4 (M = 1.63, SD = .72).

Filler task. A continuous reaction time task (CRT) was created by making sound files representing low-, medium- and high-pitched tones. This was done using Audacity software (Mazzoni & Dannenberg, 2000), such that each sine wave tone was exactly 500 ms long, at frequencies of 350, 500, and 650 Hz respectively.

Questionnaires. Participants were also asked to complete the Vividness of Visual Imagery Questionnaire, created by Marks (1973), and three questions regarding participants' history of drawing. The VVIQ is a short questionnaire that assesses individual differences in ability to create a mental image of an item or scene. Individuals are provided with four scenarios and for each, are asked to rate how clearly they can visualize nuanced aspects of each of the scenarios on a 5-point scale. There were no reliable correlations with either questionnaire throughout the course of study. Accordingly, more detailed methodological information and data regarding these correlations are omitted from this dissertation.

Procedure. Participants completed the experiment individually in a testing room. Stimulus presentation and response recording was controlled using E-prime v2.0 software (Psychology Software Tools Inc., Pittsburgh, PA) via an IBM computer with 17 inch monitor. Instructions were presented in English both on-screen and were also read aloud by the experimenter. Participants were told that, depending on the 'prompt' word they saw, they were to either 'draw' or 'write' the subsequent word on the pad of paper (14 cm x 21 cm) provided. In Experiment 1A, a prompt of 'draw' meant the participant was to draw a picture illustrating the word on the screen, and to continue adding detail until their allotted time was exhausted. A prompt of 'write' meant the participants were to clearly and carefully write out the word multiple times. In Experiment 1B, the instructions were crossed over, such that 'draw' meant they were to repeatedly draw the item presented, while 'write' meant they were to continue adding detail to a single instance of writing the word. This instruction was unorthodox, but was included to potentially shift emphasis to the write condition. Participants' responses to this instruction included using block letters, adding decorative flourishes to their letters, shading, or incorporating some elements of the concept into the writing (e.g. music notes on the word 'harp'). They were informed of time constraints for each item and that they would hear a tone to warn them that the next item would appear. Participants were not told that their memory would be tested.

Encoding. Participants underwent a brief practice phase in order to familiarize them with the encoding phase, after which the experiment began. Participants were not informed that they would be required to complete a later memory test. This incidental encoding paradigm was selected to reduce the possibility that participants would develop a strategy of preferentially focusing on drawn items in anticipation of later testing. From the list of 80 words, 30 were randomly selected to be studied, a list unique for each participant. Of these 30, 15 were randomly selected to be drawn, and 15 written (See Figure 1 for samples of the outcome of each trial type). This set of words was then presented in a randomized order, such that drawn and written items were randomly intermixed. On each trial, the prompt appeared in the center of the screen for 750 ms, followed by a 500 ms fixation, after which the word to be encoded appeared for 750 ms. Participants then had 40 s to perform the encoding task, either draw or write. A 500 ms tone alerted them that the next item was forthcoming, after which they had 3 s to flip their pad of paper to the next page in preparation for the next prompt.

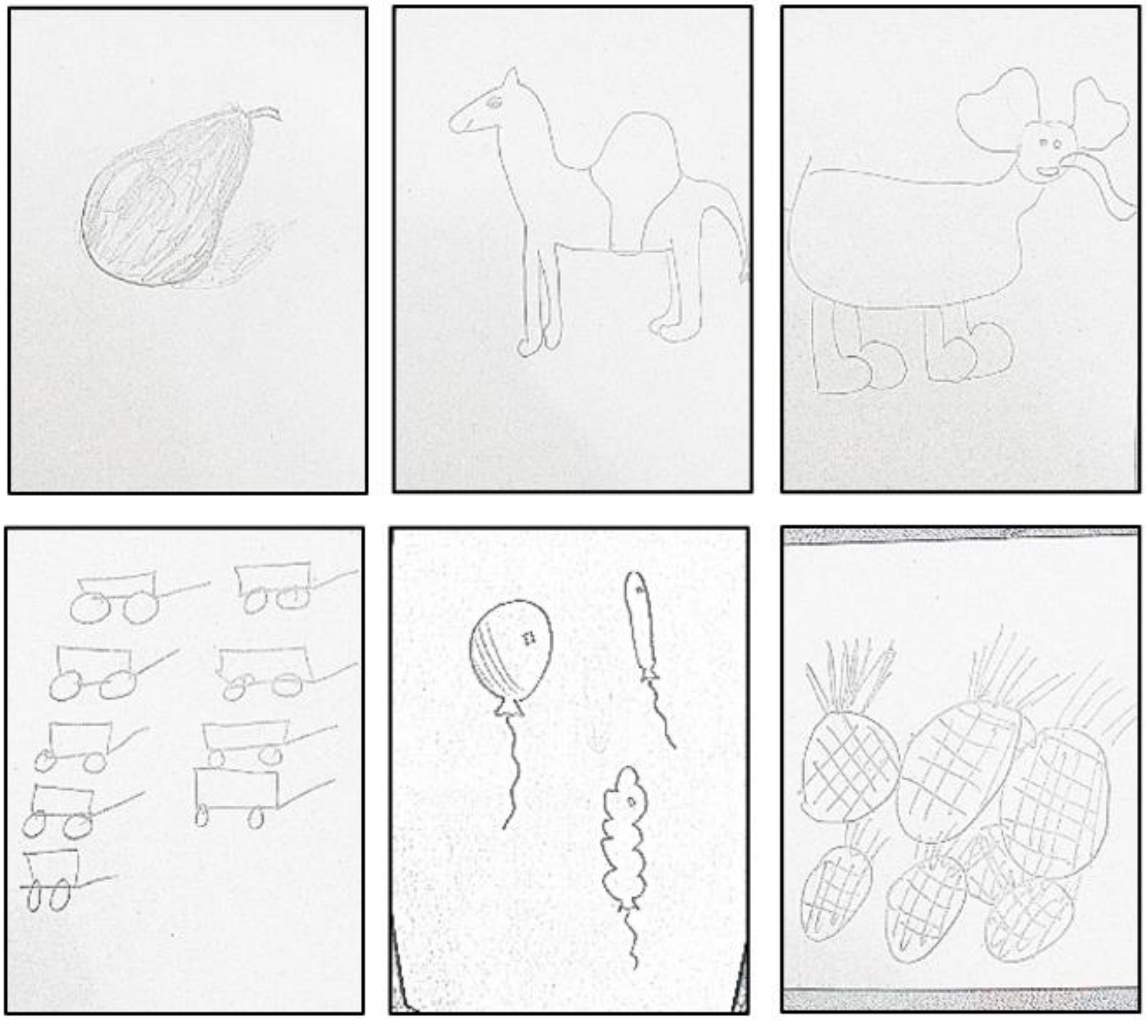

Figure 1. Examples of drawings created in Experiments 1A and 1B using pencil and paper.

Retention. Following the encoding trials, participants were asked to perform the CRT as a filler task. Tones were to be classified as low, medium or high, by pressing the 1, 2 or 3 key on a small response pad. After hearing samples of each kind of tone, participants proceeded to classify 60 tones, selected at random. For each trial, the tone was played for 500 ms, after which participants had 1500 ms to make their response, for a total of 2000 ms per trial. Thus the retention interval was two minutes.

Retrieval. In the next phase of the experiment participants were asked to freely recall as many words as they could, in any order, either written or drawn, from earlier in the experiment. They were given 60 s to complete their recall, which was spoken aloud by the participant, and recorded.

Questionnaires. Immediately following the retrieval phase, participants completed a version of the VVIQ (Marks, 1973), and the three questions pertaining to drawing experience and ability.

Results and Discussion

Participants' recall output was sorted into the number of recalled words that were drawn at encoding and the number that were written. Each of these values was divided by the total number of words studied within each encoding trial type (15), to create a proportional recall score. Data were analyzed in a 2x2 mixed measures ANOVA, with Experiment (1A, 1B) as a between-subjects variable and trial type (Draw, Write) as a within-subjects variable. Analyses showed a significant main effect of trial type, F (1, 53) = 82.83, MSE = .02, p < .001, η^2 =.61, such that drawn words were recalled better than written words (See Figure 2). The average number of times participants wrote out the word in Experiment 1A (M = 16.71, SD = 5.19) or drew the word in Experiment 1B (M = 2.96, SD = 3.85) was not correlated with the number of recalled drawn items, recalled written items, total recalled items, or the magnitude of the benefit of drawing, *p*s > .21. Twenty-six of the thirty participants in 1A showed the pattern of improved recall for drawn, relative to written items (t (29) = 8.29, SE = .03, p < .001, d = 1.51), while twenty of the twenty-five did in 1B (t (24) = 4.81, SE = .04, p < .001, d = 0.96). Hereafter, I use the term 'drawing effect' to refer to this clear advantage of drawing words relative to writing them out in terms of later memory performance.

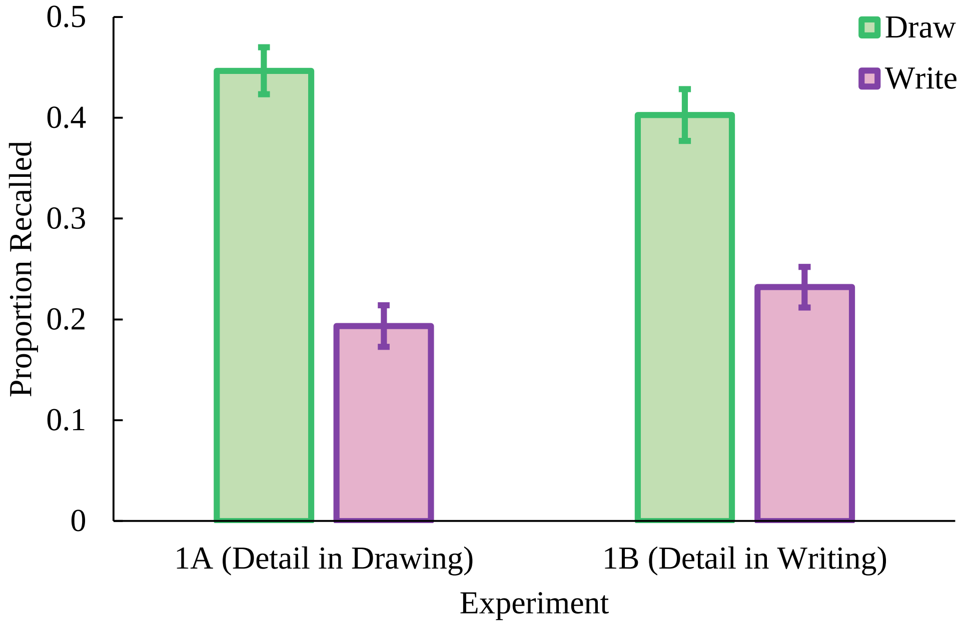

Figure 2. Proportion of words recalled out of 15 in Experiments 1A and 1B, from the draw and write trial types. Error bars represent the standard error of each mean.

The main effect of Experiment, F (1, 53) = 0.02, MSE = .00, p = .90, η^2 =.00, and the interaction were not significant, F (1, 53) = 3.15, MSE = .02, p = .08, η^2 =.06, allowing the conclusion that first, there was no notable overall difference in memory performance depending on whether participants were instructed to add detail to one trial type or the other, and second, that the benefit of drawing did not differ depending on the instruction to add detail or not. Such a finding rules out the possibility that the instruction to add detail on drawn trial types (Exp 1A) could account for the beneficial effect of drawing on subsequent memory, since the instruction to add detail to the written trial types did not confer the same benefit (Exp 1B).

First and most importantly, my results indicated there was a significant recall advantage for words that were drawn during incidental encoding as compared to those that were written. Participants recalled more than twice as many drawn than written words in Experiment 1A, nearly twice as many drawn words in 1B, and the majority of participants recalled more drawn than written words. This finding expands upon previous work suggesting some special advantage of drawing (Paivio & Csapo, 1973; Peynircioglu, 1989). Contrary to previous work however, I carefully controlled the time the participants were given to complete their drawings, thereby allowing sufficient time for the creation of one's drawing, and also equated the time provided for drawing, with the time given for writing trials. Unlike in previous work, in which participants wrote the word once then subsequently sat idle, I created trial types wherein the entirety of the time allotted for each trial type was most likely to be used. The design thus increased the likelihood that participants were processing the word throughout the entire encoding trial, rather than engaging in task unrelated thoughts.

Experiment 2

Having established that drawing improves free recall, and ruling out the alternative account that the instruction to add detail was producing the drawing effect (Experiment 1B), my aim in Experiment 2 was to determine whether the observed benefit resulting from drawing relative to writing words could be simply explained by one of three readily apparent candidate mechanisms: deep LoP (e.g. Craik & Lockhart, 1972), visual imagery (e.g. Elliott, 1973), or picture superiority (e.g. Paivio & Csapo, 1973). Indeed, because drawing requires one to engage with a word at a deep semantic level, to develop a mental image to draw, and to create a picture, it could be the case that the beneficial effect of drawing can be explained away using these well-established mechanisms. To test whether this was the case, trial types designed to elicit each of these acts or processes were designed, and contrasted directly with drawing. In Experiment 2A, drawing and writing were compared with a 'list' trial type, wherein participants must process the word at a deep semantic level in order to generate a list of physical characteristics of the word's referent. In Experiment 2B, the 'list' trial type was replaced with a 'visualize' trial type, wherein participants were required to engage in visual imagery to create a mental image of the word's referent. Lastly, in Experiment 2C, the 'visualize' trial type was replaced with a 'view' trial type, where participants simply viewed an image of the word's referent, thereby introducing potential picture superiority effects. I predicted that performance for drawn trials would surpass all other included trial types. This is because drawing is seemingly a much richer multimodal process than any of the others, including not only semantic elaboration and imagery, but also motoric information required to undergo the process of moving the pencil to create the image.

Method

Participants. Participants for Experiment 2A, 2B and 2C were 47 (40 female), 28 (21 female), and 37 undergraduate students (29 female) respectively at the University of Waterloo. These were unique groups of participants for each experiment, so all were naïve subjects. They all completed the experiment in return for course credit or monetary remuneration. Participants ranged in age from 17 to 51 (M = 19.86, SD = 3.42), with between 13 and 19 years of education (M = 15.25, SD = 1.28). All participants had normal or corrected to normal vision, and learned English before the age of seven.

Materials. Word stimuli and tones were the same as in previous experiments. For the 'view' trial type, a picture was chosen for each word stimulus from a picture set containing thousands of unique images of objects (Brady, Konkle, Alvarez & Oliva, 2008). If the object was not found in this stimulus set, one was retrieved using a Google image search. All images were cropped to the same dimensions (256 x 256 pixels), and converted to grayscale using ImageMagick software (ImageMagick Studio LLC, 1999-2013).

Procedure. The procedure was identical to that used in Experiment 1A, except that a third trial type was added during encoding. Participants were instructed to either ‘draw’ or ‘write’ the word being presented to them. Instructions for ‘draw’ and ‘write’ were identical to previous experiments. In each experiment, a third trial type was added. In Experiment 2A, 2B and 2C, these new trial types were ‘list’, ‘visualize’ and ‘view’ respectively. The instructions were as follows:

For ‘list’ trials:

> If the prompt is 'list' we ask that you generate a list of associated words, adjectives or characteristics describing the word. Continue adding words to your list until time is up.

For ‘visualize’ trials:

> If the prompt is 'visualize', we ask that you mentally picture what the word represents. Maintain your focus on the image and continue adding detail to it until your time is up.

For ‘view’ trials:

> If the prompt is 'view', we ask that you view a picture of what the word represents until your time is up.

Because there were now three different prompts in each experiment instead of two, the randomly selected list of 30 words was divided into three lists of 10 words each (10 to be drawn, 10 to be listed/visualized/viewed, 10 to be written) instead of two lists of 15, as in the prior experiments. This set of words was then presented in a randomized order, such that drawn, visualized and written items were randomly intermixed. Apart from these modifications, the experimental protocol was identical to Experiment 1A.

Results and Discussion

All data are analyzed using repeated measures ANOVA, with Encoding trial type as a within-subjects variable, and followed up with paired samples t-tests using Bonferroni adjusted alpha levels of .0167 per test (.05/3). In Experiment 2A, data from one participant were excluded as they performed the incorrect encoding instruction for 9 of the 30 words. There was a significant main effect of Encoding trial type (draw, list, write) in the remaining participants, F (2, 92) = 21.20, MSE = .02, p < .001, η^2 =.32. This was driven by significantly better recall for words in the draw than both the list, t (46) = 5.16, SE = .03, p < .001, d = 0.75, and write trial types, t (46) = 5.89, SE = .03, p < .001, d = 0.86. The difference in recall between words from the list and write trial types was not significant, t (46) = 0.39, SE = .03, p = .70, d = 0.06 (See Figure 3). In the ‘list’ trial types participants listed a mean of 6.37 characteristics (SD = 1.65) per word, but the number of characteristics listed was not correlated with the number of words recalled overall, or in any of the three encoding trial types, ps > .11. This finding indicates that drawing cannot be characterized as simply another iteration of classic deep LoP findings. Further, listing a greater number of characteristics for any particular word did not lead to better memory for that word, suggesting that regardless of how deeply words were encoded in the listing task, the benefit of drawing as an encoding strategy remained superior.

In Experiment 2B, there was a significant main effect of Encoding trial type, F (2, 54) = 8.23, MSE = .04, p < .005, η^2 =.23, driven by significantly better recall for drawn trial types than for write trial types, t (27) = 4.62, SE = .04, p < .001, d = 0.87, and marginally better recall for drawn than visualized trial types, t (27) = 2.20, SE = .06, p = .037, d = 0.42. The difference in recall between visualize and write trial types was not significant, t (27) = 1.58, SE = .05, p = .126, d = 0.30 (See Figure 3). While the contrast between drawing and visualizing failed to meet Bonferroni-corrected thresholds, the general pattern indicates that drawing led to better memory than a visual imagery manipulation.

In Experiment 2C, there was a significant main effect of Encoding trial type, F (2, 72) = 7.24, MSE = 0.03, $p < .005$, η^2 =.17, driven by significantly better recall for words in the draw than both the view, t (36) = 2.56, SE = .04, $p < .0167$, d = 0.42, and write trial types, t (36) = 4.08, SE = .04, $p < .001$, d = 0.67. The difference in recall between words from the view and write trial types was not significant, t (36) = 1.07, SE = .04, p = .293, d = 0.18 (See Figure 3). This pattern of results indicates that the effect of picture superiority alone cannot be invoked as an explanation for the observed benefit of drawing, as drawing led to better recall than viewing images.

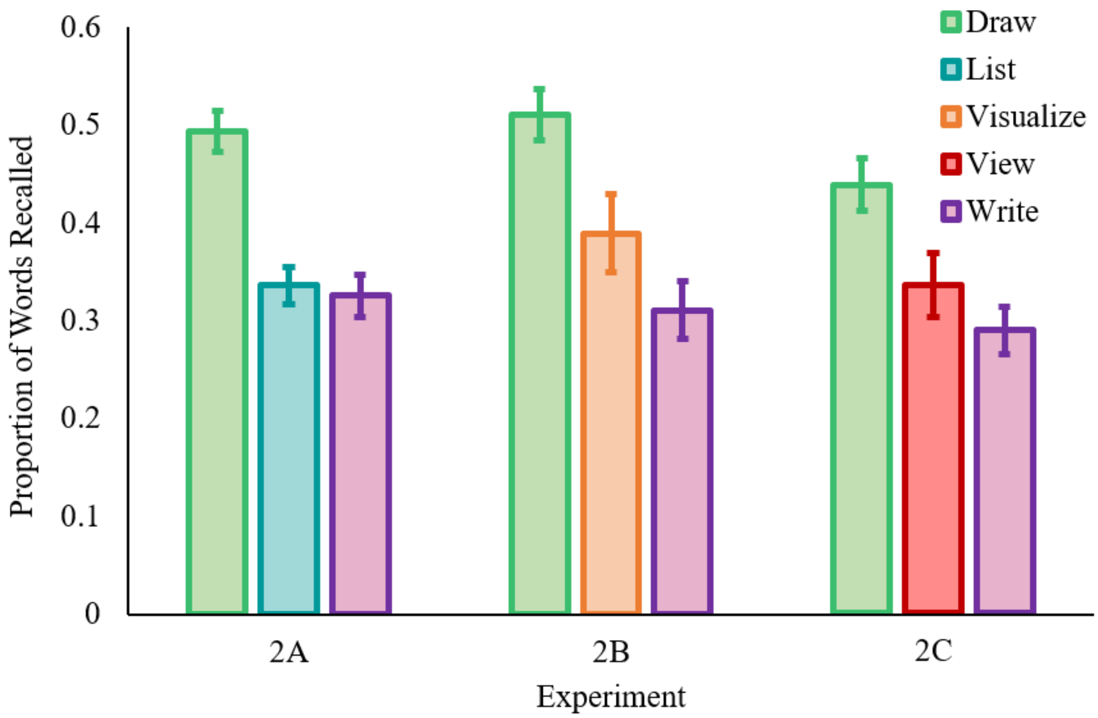

Figure 3. Proportion of words recalled out of 10 in Experiments 2A, 2B and 2C, from the draw, list, visualize, view and write trial types. Error bars represent the standard error of each mean

Experiment 3

While drawing did improve memory relative to writing in Experiments 1 and 2, it was possible that because the list length was so short, and the encoding time so long, that the manipulation was simply unduly shifting the emphasis to item information rather than order information (see McDaniel & Bugg, 2008), thus providing an advantage to drawn items. While we did not explicitly test memory for order in these experiments, it is known that relational information about the order in which stimuli were presented benefits free recall in other paradigms (e.g. Engelkamp & Dehn, 2000; Jonker et al., 2014). Thus, shifting the emphasis toward item information (i.e. as a result of longer encoding times per stimulus) might hinder memory in the ‘write’ condition. In this experiment, it was my aim to test whether this was the case by reducing the amount of time allowed for drawing (and writing) each item, and increasing the number of items. In addition, from a purely practical point of view, providing participants with 40 seconds to draw is *much* longer than the time which is typically allowed for many encoding tasks (e.g. MacLeod et al., 2010; Nairne, Thompson & Pandeirada, 2007; Slamecka & Graf, 1978). It is possible that the beneficial effects of drawing would not occur given a shorter time frame allowed for encoding/drawing, given that drawing can be such a time-consuming task.

A borrowed explanation from the production effect literature suggests that ‘lazy reading’ (Forrin, Jonker & MacLeod, 2014; Macleod et al., 2010) may be contributing to the beneficial effects of production. That is, participants might not be properly engaging with the encoding task when they are instructed to read silently. Conversely, when they read aloud, there is audible evidence of their compliance with the manipulation, so participants are less likely to be loafing. The drawing effect has clear parallels to production in this instance: because participants were provided with 40 seconds to draw and write in Experiment 1, it is possible that participants were loafing, or ‘lazy writing’ for a subset of the allotted encoding time. If the drawing effect was due to ‘lazy writing’ over the course of this extended 40 second encoding duration, then the drawing effect should not be present when encoding duration is drastically reduced. Any drawing created within a 4 second time period would likely contain fewer details, and thus fewer contextual cues for later retrieval, than a drawing created in 40 seconds. For this reason, I predicted that the shorter encoding time might lead to a drawing effect that was smaller in magnitude than that previously observed.

Method

Participants. Participants were 28 undergraduate students (22 female) at the University of Waterloo, who completed the experiment in return for course credit or monetary remuneration. Participants ranged in age from 18 to 25 (M = 20.64, SD = 1.90), with between 13 and 22 years of education (M = 16.41, SD = 1.96). All participants had normal or corrected to normal vision, and learned English before the age of seven.

Materials. Word stimuli were the same as in Experiment 1. The rapid presentation of stimuli that was to occur in this Experiment necessitated a methodological change. Our previous implementation, wherein drawings were created using pencil and paper, required participants to flip the page of their notepad over between stimuli, and thus was impractical given the new trial structure. Accordingly, I switched to using a Fisher-Price Doodle Pro®. Similar to an Etch-a-Sketch®, the Doodle Pro® is a small drawing pad which uses magnetic drawing technology to allow participants to create drawings on a surface roughly the same size as Experiment 1's drawing pads (11 cm x 16 cm). Unlike an Etch-a-Sketch®, the Doodle Pro® comes with a small stylus to draw, rather than bimanual knobs, making it more analogous to the paper and pencil technique used in my earlier experiments. The Doodle Pro® also features a sliding knob on the top which quickly wipes clean the drawing pad in preparation for the next trial.

Procedure. The procedure was very similar to Experiment 1A, with the exception of some timing and presentation changes to accommodate more rapid presentation of to-be-remembered stimuli. Additionally, in order to facilitate comparison with investigations of other encoding strategies thought to depend on distinctiveness (McDaniel & Bugg, 2008), this experiment employed intentional, rather than incidental encoding. Like Experiment 1A, participants were tested individually in a testing room. Participants studied 66 words instead of 30, split equally between the write and draw encoding trial types. These 66 words were then presented with drawn and written items randomly intermixed. During encoding, rather than the typical protocol of showing a participant a prompt (750 ms), a fixation (500 ms), the word (750 ms) and then asking them to complete the task (40 s), participants were simply shown the word. Words were presented on a black background in either white or red font, and participants were instructed that the font color indicated whether they were to draw or write the item (with colors counterbalanced across participants). Each word was presented for 4 seconds, during which time the participant was to complete the task indicated by the font color of the word. Between trials, a blank screen was shown for 1 second to give participants time to use the sliding knob to reset their drawing pads.

The 2 min filler task from the previous experiment was replaced with a 5 min tone classification, in addition to a 5 min visual CRT task, in which participants identified within which three numbered boxes (left, middle or right) an asterisk was presented, using the keys corresponding to the numbers presented above the boxes (1, 2, 3).

Participants were then given two minutes to recall as many words as they could from the encoding phase, by typing them into an input field. Participants heard a tone to notify them when there were 20 seconds remaining in the recall test. The VVIQ or the drawing related questions were not completed in this or any of the subsequent experiments.

Results and Discussion

First, the data were analyzed with cue color (red or white) as a between-subjects factor. There were no main effects or interactions with cue color, so the data were collapsed across this variable. Data were analyzed in a paired-samples t-test, with encoding trial type (Draw, Write) as a within-participants variable. Analyses showed a significant main effect of trial type, t (27) =

12.01, *SE* = .02, $p < .001$, $d = 2.27$, such that drawn words were better recalled than written words (See Figure 4).

These results indicate that despite a lengthier study list (33 per trial type instead of 15), a dramatically shorter encoding duration (4 s instead of 40 s), and an extended retention interval between study and test (10 min instead of 2 min), drawn words were remembered much better than those that were written. Further, the effect was substantially more pronounced (see Figure 4) than in previous experiments (See Figures 2 and 3), suggesting that the drawing encoding strategy is potentially even more potent at shorter encoding durations. Comparisons between this and previous experiments however, may be problematic because the paradigms differ in more ways (e.g. list length, intentional learning) than simply encoding duration. For instance, it is also possible that writing was simply a *less* potent strategy at shorter encoding strategies, or that this larger effect size was an artefact of the change from incidental to intentional encoding.

Experiment 4

Having ruled out emphasis on adding detail to a particular encoding trial type (the instruction to continue adding detail to drawing, or writing in Experiment 1), lazy writing (Experiment 3) and several individual candidate explanations for the effect (elaboration, visual imagery, and picture superiority in Experiments 2A-C), I turned my focus to examining whether the drawing effect can be accounted for, like many other item-specific encoding strategies (e.g. McDaniel & Bugg, 2008; Nairne, Riegler & Serra, 1991), by a distinctiveness account. I proposed that drawing exerts its memorial benefits through integration of a number of different multisensory memory codes into one cohesive trace. The distinctiveness account however, is a plausible alternative account, in that drawing may be exerting its effects by making some items more distinct than others, at the expense of those others. The current work draws obvious parallels with the production effect, where the explanatory mechanism often proposed is a distinctiveness account (Bodner & Taikh, 2012; MacLeod et al., 2010; McDaniel & Bugg, 2008; Ozubko & MacLeod, 2010). In other words, in a mixed-list paradigm, some items are read aloud (or generated from a cue, or enacted), while others are simply read silently (or enacted by experimenters while the participant passively views) (Engelkamp & Zimmer, 1997; Slamecka & Graf, 1978). Accordingly, the trials that are more unique, in which a participant completes a novel or bizarre encoding task, have some distinctive information that can later be used as an extra retrieval cue. The empirical support for this account comes from the observation that most of these effects are absent, or in some cases reversed, when trial types are completed between-participants, in pure lists, as opposed to in mixed lists within-participants. This is true of production (MacLeod et al., 2010; but see Bodner et al., 2014), generation (Slamecka & Katsaiti, 1987), enactment (Engelkamp & Dehn, 2000), and many other commonly cited mnemonic strategies (McDaniel & Bugg, 2008; Mulligan, 2002). In contrast with most of the foregoing findings, which were borne out of recognition paradigms, the experiments in the current work used free recall tests. In contrast to recognition, where some previous work has found a significant between-participants production effect (e.g. Bodner et al., 2012), no such between-participants production effect has been observed in free recall (Jones & Pyc, 2014; Jonker et al., 2014). What this tells us is that the benefit of many of the foregoing effects manifests as a *relative* enhancement, rather than an absolute one. Specifically, distinctive items are better encoded at the expense of the less distinctive ones. In an attempt to control for this throughout all of the experiments thus far, I employed writing as a baseline condition with which to compare drawing. This baseline was chosen in part because writing has been previously described as a version of the production effect, and as a task that is distinctive in its own right (Forrin et al., 2012). Further, in Experiment 2, I compared drawing with alternate encoding trial types of visual imagery, elaborative encoding and picture superiority, which might also be categorized under the distinctive encoding umbrella.

However, in order to conduct the most direct experimental exploration of a distinctiveness account, in Experiment 4, I compared drawing and writing encoding strategies between-participants in pure lists, a manipulation that typically eliminates the effects of other encoding strategies that are purported to derive their benefits from relative distinctiveness (McDaniel & Bugg, 2008; Mulligan, 2002). Thus, if the drawing effect was driven solely by distinctiveness, one would predict a minute or absent effect in a between-participants design. If distinctiveness can only partially explain the findings, however, the results should indicate an extant, albeit smaller boost as a result of drawing items, relative to writing them. I predicted the latter, based on the magnitude and reliability of the effect in empirical investigations thus far.

Method

Participants. Participants were 47 undergraduate students (37 female) at the University of Waterloo, who completed the experiment in return for course credit or monetary remuneration. Participants ranged in age from 18 to 24 (M = 19.51, SD = 1.51), with between 13 and 21 years of education (M = 15.55, SD = 1.58). All participants had normal or corrected to normal vision, and learned English before the age of seven. Participants were randomly assigned to be in the Pure draw (n = 24) or Pure write (n = 23) conditions.

Materials. All materials were the same as in Experiment 3.

Procedure. The procedure was identical to Experiment 3, including the intentional encoding, except that participants were randomly assigned to one of two conditions. In one condition, participants drew *all* of the words (Pure draw), and in the other, they wrote *all* words (Pure write) presented during study. Accordingly, there was no need for color coding of stimuli, and all study words were presented in white font on a black background. Participants still studied 66 total words, but rather than 33 falling under each Encoding trial type, the entire set of 66 were either drawn or written.

Results and Discussion

Data were analyzed using an independent-samples t-test, with trial type (Draw, Write) as a between-subjects variable. Analyses showed a significant main effect of trial type, t (45) = 3.63, SE = 0.02, p < .005, d = 1.08, such that those in the Pure draw condition recalled significantly more words than those in the Pure write condition (Figure 4). Interestingly, the effect is likely underestimated in this analysis, as one participant in the pure write condition recalled a striking 33 words, which is more than 3.5 standard deviations above the mean. When analyzed without this outlier, the effect size was substantially larger, t (44) = 5.54, SE = 0.02, p < .001, d = 1.63. Overall, these results suggest that a distinctiveness account is not sufficient to explain the benefit that drawing affords recall. Other effects (e.g. production, generation) depend on a relative boost at the expense of the less distinct words, resulting in a lack of any memorial effect when the encoding strategy is presented in pure lists, between participants (McDaniel & Bugg, 2008). These data indicate that even when compared between participants, drawing allows for a significant boost in memory performance.

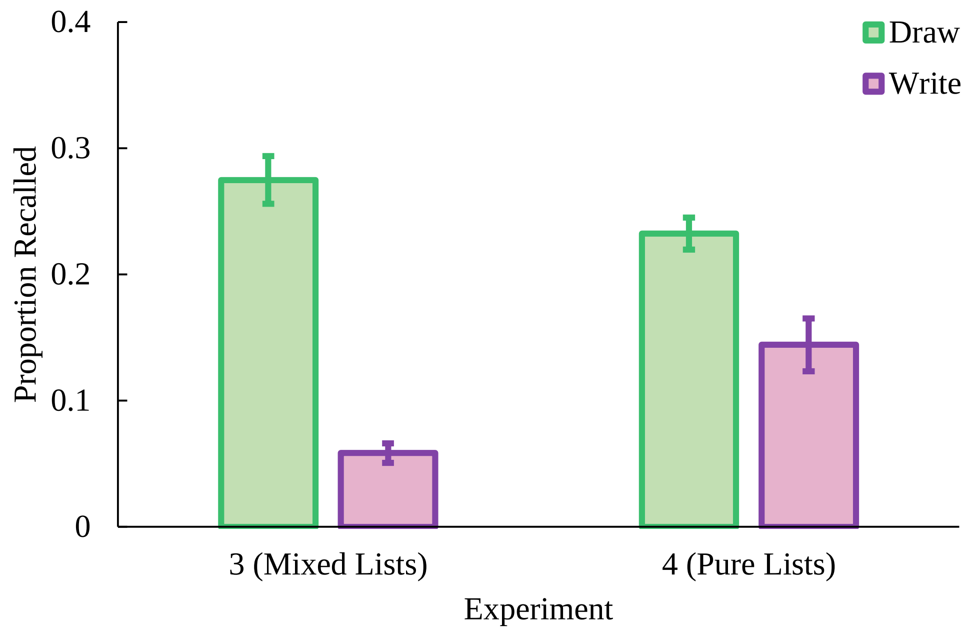

Figure 4. Mean proportion of words recalled out of 33 for each trial type in Experiment 3 (Mixed-lists) and out of 66 for each trial type in Experiment 4 (Pure-lists). Error bars represent the standard error of each mean.

It should be noted that by inspection of the means, we see that the proportion of written words that were recalled in the mixed lists experiment is quite low relative to the higher proportion in the pure list presentation. For drawn words, while there was a higher proportion of words recalled in the mixed list relative to pure list presentation, the difference was only marginal and does not appear as dramatic as the increase in proportion of written words. An analysis designed to contrast within and between-participants comparisons of the same conditions was performed, as outlined by Erlebacher (1977). This analysis is designed to explore the interaction of design type with a dependent variable of interest[1]. Results indicated that there was a significant Design Type by Encoding Trial Type interaction, F (1,44) = 19.63, $p < .001$, η^2

= 0.09. Such a result indicates that the effect of drawing is augmented in mixed lists, but this is due to lower recall of written words, rather than a massive boost in recall of drawn words. Because of well documented issues with output interference in free recall paradigms (e.g. Roediger, 1974), it is also unclear whether the poorer memory for written words is due to the moderate distinctiveness advantage for drawn words, or due to the fact that output interference, especially with long lists, prevents output of written words later in the allotted recall time. Similar to my work, Bodner et al. (2014) found that memory for words read aloud during encoding was not substantially improved in a within-participants, relative to a between-participants design, as would be expected given the distinctiveness account (MacLeod et al., 2010). Their findings were comparable to ours, suggesting that a distinctiveness account might not be applicable to production, or to drawing. However, their findings were in a recognition paradigm, and thus are not readily generalizable to our current work in free recall. This is especially noteworthy, given that there is no evidence to support a between-participants production effect in free recall (Jones & Pyc, 2014; Jonker at al., 2014). Thus, it is clear that the drawing effect cannot be fully explained by a relative distinctiveness account, and can, at least by one important criterion, be differentiated from other item-specific encoding strategies thought to rely on distinctiveness (McDaniel & Bugg, 2008). The most important point is that there remains at least a 9% (or a nearly 6-word) benefit of drawing in the pure-lists design, such that 23% of drawn words, but only 13-14% of written words were later recalled. This shows that even if distinctiveness can account for some of the advantage provided by drawing, it is far from an exhaustive mechanism.

CHAPTER III: Drawing and Recollection

In Chapter II, I established that drawing is a relatively quick, relatively easily implemented strategy that is efficacious in improving free recall performance. I also established that the benefit drawing affords memory cannot easily be entirely explained through simple deep LoP, picture superiority or visual imagery mechanisms alone. This was demonstrated by contrasting drawing directly with trial types designed to capture the underlying mechanism driving these classic effects. Furthermore, the effect cannot be explained by the abundance of encoding time, or 'lazy writing' or loafing in our baseline condition. In Chapter II, I also ruled out distinctiveness as a mechanism that can wholly explain 'the drawing effect'.

In this chapter, I explore perhaps the most important of the questions that were left unanswered following my previous work: how does drawing exert its memorial benefit? I previously proposed that drawing exerts its beneficial effects by promoting a better-integrated memory trace (Wammes et al., 2016). That is, I suggested that drawing leads to better memory than other comparable encoding strategies because it facilitates the integration of a basic verbal memory trace with other multisensory information derived from the encoding experience. I argued that the memory trace created by drawing consists of three integrated components: the semantic/elaborative information required in deciding what to draw, motoric information from physically producing the image, and pictorial information from the final drawing that was created. I also provided preliminary evidence to suggest that the estimated contributions of these three components, when each was isolated through subtractions and then added together, still fell slightly short of achieving the level of recall performance attained by drawing alone (see Wammes et al., 2016).

For this integrated trace hypothesis to be plausible however, participants must be able to effectively retrieve specific contextual information from the initial encoding experience to a greater extent when they had drawn items, relative to when they had written them. In other words, the participant must experience a 'recollection' of the drawing experience. Within dual-process models of recognition memory, theorists posit that successful retrieval can be driven by two processes: recollection and familiarity (Gardiner, 2001; Perfect, 1996; Tulving, 1983; 1985; Yonelinas, 2002). Recollection-based memory encompasses the rich, vivid experiences where one is able to consciously bring to mind specific contextual details about the initial encoding event. Conversely, familiarity-based memory is a general phenomenological feeling of familiarity or perceptual fluency, wherein one has the sense that they have been exposed to an item before, but cannot attribute any specific contextual details to that feeling (Yonelinas, 2002). While there are several different conceptualizations of dual-process models (e.g. Atkinson & Juola, 1973; Graf & Mandler, 1984; Jacoby & Dallas, 1981; Juola, Fischler, Wood, & Atkinson, 1971; Mandler, 1980; Yonelinas, 1994; 2002), it is not my aim to arbitrate. Rather, I focus on the commonalities between models as they pertain to the current work's goals. Most importantly for this chapter, it is generally agreed upon across models that recollection is accompanied by specific contextual details from the encoding experience, and that is a slower-unfolding process than familiarity (Atkinson & Juola, 1973; Yonelinas, 2002).

I proposed that drawing improves memory by integrating elaborative, motoric and pictorial components of a memory trace. A necessary precondition for this integration hypothesis would be a demonstration that drawing improves recollection in particular. To inform predictions, we can consider the relation these proposed components might have with recollection and familiarity through the lens of their most analogous respective literatures (i.e.

deep levels of processing (LoP), enactment effect, picture superiority). Deep LoP manipulations, which require that participants attend to and elaborate upon the *meaning* of the studied information, tend to selectively increase recollection (e.g. Gardiner, 1988; Gardiner, Java & Richardson-Klavehn, 1996; Rajaram, 1993) as evidenced the generation effect being smaller in magnitude when responses are speeded (Mulligan & Hirshman, 1995). There is evidence to suggest that enactment manipulations, which improve memory by instructing participants to perform an action in response to a presented item, also target recollection-based memory exclusively (Engelkamp & Dehn, 1997; Lövdén, Rönnlund, & Nilsson, 2002; Manzi & Nigro, 2008). Findings however, are less clear with respect to whether picture superiority is driven by recollection or familiarity. Some findings suggest a role for both of these processes (e.g. Defeyter, Russo & McPartlin, 2009; Dewhurst & Conway, 1994; Wagner, Gabrieli, & Verfaellie, 1997), while others indicate that recollection is most critical (e.g. Boldini, Russo, Punia, & Avons, 2007; Curran & Doyle, 2011; Rajaram, 1996). Taken together though, this constellation of findings suggests that drawing might serve to increase recollection primarily, but also could boost familiarity due to the influence of pictorial information.

While I make my predictions based on the foregoing, it is as yet unclear whether one's phenomenological experience when retrieving drawn items is consistent with a general feeling of familiarity or perceptual fluency, or with a rich, recollective experience, specific in time and place. To answer this question, in this chapter I employed two variants of recognition memory tasks to determine whether drawing to-be-remembered words produces a contextually strong, recollective memory of an item, as opposed to a more familiar or gist-based memory. To measure the contribution of recollection to the beneficial effects of drawing on memory, I approached the question using two different paradigms that have been employed in the literature to first differentiate between recollection and familiarity, then to isolate recollection. I took this approach because each paradigm independently carries its own strengths, weaknesses, and theoretical criticisms. To overcome these potential methodological issues, I examined the same phenomenon using both paradigms, and predicted that the convergence of evidence across tasks would support one common conclusion. In this way, any findings would not be an artefact of the failings of one particular paradigm. Specifically, in Experiments 5A and 5B, I employed a traditional remember/know/new (RKN) paradigm. I predicted better overall memory, as well as a greater proportion of recollection-driven responses for drawn items. In Experiment 6, I used a response deadline procedure applied to a basic old/new recognition task. If memory for drawn items is driven largely by recollection-based responses, then speeding the decision with a response-deadline should lead to a smaller drawing effect than would more typical relaxed timing constraints. Thus, across experiments, it was my aim to test whether drawing improves memory by providing a rich, immersive setting at encoding, providing a wealth of contextual information which can later be used to facilitate retrieval. My previous work, both in this dissertation and elsewhere (Wammes et al., 2016; in press) is suggestive of this mechanism, as drawing aided free recall, which is thought to rely primarily on recollection, and memory for encoding modality (i.e. source memory) for both images and words was superior for drawn, relative to written items (Wammes et al., in press).

Experiment 5A

In Experiment 5, I sought to exhaustively sample the contribution of recollection to the subjective state of participants' memory using the remember/know paradigm (Tulving, 1985; Gardiner 1988). One advantage of the remember/know paradigm is that it provides an inclusive measure of recollection (as opposed to source memory decisions, for instance). Namely, it is not limited by what specific contextual information the experimenter requests, which is the critical shortcoming of the source memory decision employed in my previous work (Wammes et al., in press). The remember/know task was developed to allow participants to explicitly report whether their memory was based on recollection processes (i.e. they 'Remember' the item), or on familiarity processes (i.e. they 'Know' the item was present). Results from this paradigm typically converge on those using other methods of capturing recollection and familiarity (Yonelinas, 2002). The instruction encourages participants to focus on whether or not they remember specific episodic information about the encoding experience, and to answer 'Remember' if they do. Thus, unlike source memory decisions, which only capture recollections that are associated with knowing what task they performed in response to the studied item, the remember/know paradigm should capture nearly *all* recollection-based hits.

The remember/know paradigm has been used extensively as a proxy for the recollection and familiarity processes that underlie recognition memory. However, there are differing opinions on whether or not the paradigm actually measures separable psychological processes (Rotello & Zeng, 2008; Starns & Ratcliff, 2008). Specifically, the exclusivity model holds that a given recognition memory response can be based on recollection *or* familiarity, but not both (Gardiner & Parkin, 1990; Jones, 1987; c.f. explicit vs. implicit information: Nelson, Schreiber, & McEvoy, 1992, but see Wais, Mickes, & Wixted, 2008 for evidence to the contrary). The redundancy model suggests that all successfully recognized items are familiar, whether they are recollected or not (Joordens & Merikle, 1993; Knowlton & Squire, 1995; Tulving, 1985). Lastly, the independence model suggests that 'Remember' responses are based on recollection, while 'Know' responses are based only on the subset of familiarity-driven memories where no recollection occurred (Jacoby, Yonelinas & Jennings, 1997; Yonelinas & Jacoby, 1995).

Importantly however, this work is more concerned with whether drawing improves recollection of studied items, a question for which remember/know is particularly well suited. Thus in the current experiment, I used the remember/know paradigm under the assumption that 'Remember' responses contain more contextual detail than 'Know' responses, and aimed to determine the effects of drawing (relative to writing) during encoding on the subjective mnemonic experience of the rememberer. I predict that drawn items, relative to written items will be associated with an increase in 'Remember' responses, as well as an increase in overall memory performance.

Method

Participants. Participants were 35 undergraduate students (23 female) at the University of Waterloo, who completed the experiment for course credit. Participants ranged in age from 17 to 22 (M = 18.74, SD = 1.24).

Materials. Materials were the same as in Experiments 1 through 3.

Procedure. Participants were tested in groups in a lecture hall, as previous work showed no impact of testing in group setting or individually on the drawing effect (Wammes et al., 2016). The encoding sequence, including prompts, instructions and timing were identical to Experiment 1. However, participants now studied 40 words (20 drawn, 20 written) instead of 30 words. The distracting task completed during the retention interval was also the same as in Experiment 1A, though responses were provided using a response sheet rather than a button press.

At the Retrieval stage, the memory test was an RKN paradigm. Participants were told that a "Remember" (R) response meant that they had a conscious recollection of specific contextual information about their initial encounter with the word, such as hearing a sound in the hall, or what they did when the word was presented. They were told that a "Know" (K) response meant that they had only a feeling of familiarity; that they believed that the word had been seen recently, but could not remember specific details from their initial viewing of the word. Lastly, a "New" (N) response meant that they did not encounter the word in the previous phase. Participants were given a Scantron sheet, typically used in multiple choice examinations, with which to indicate their responses. These sheets consist of a set of columns, each with ten rows containing a number (e.g. 11-20), followed by the options A through E, each with a circle surrounding the letter. Letter responses are indicated by filling in the corresponding circle for each numbered question. Participants indicated whether an item was one that they would classify as R, K or N by shading in a bubble on their score sheet (A, B or C respectively). The 40 studied words were presented amongst 40 lures, and each word was presented centrally on screen for 3 s. The test trials were numbered consecutively from 1 to 80, and this number was displayed prominently on the top left of the screen during the test. This was done so that in the event the participant missed an item, they could easily reorient themselves to the appropriate numbered row on their Scantron sheet. The response mappings (i.e. ‘Remember (A) Know (B) New (C)) were also displayed at the bottom of the screen for the duration of the test. Following each trial, a blank screen was presented for 1 s.

Results and Discussion

Data from this experiment were analyzed in a few different ways. First, I compared overall recognition memory performance (collapsing remember and know responses), then compared R and K responses separately. Hit rates were calculated out of 20 possible items, while false alarm rates were out of 40 possible items. Accuracy (hit rate minus false alarm rate) and sensitivity (d’) were also computed overall and within each response type. Because hit rate and accuracy followed the same pattern as d' in all analyses, I only report d' in full, while the other values are presented in Table 1. I also computed and analyzed process estimates for recollection and familiarity based on calculations outlined in previous work (e.g. Yonelinas, Kroll. Dobbins, Lazzara, & Knight, 1998). However, because the contextual quality of memory was so high in some participants (that is, their hits were exclusively 'Remember' responses), I was unable to reliably estimate the contribution of familiarity in these individuals. Thus, these analyses are based on only the subset of subjects wherein the estimates could be computed for familiarity in both trial types.

Overall recognition. Paired-samples t tests indicated that overall d' was higher for drawn words than for written words, t (34) = 8.00, SE = .10, d = 1.35, p < .001 (WSR Z = 4.64, p < .001, PSS Z = 4.93, p < .001). Accuracy and hit rate were also higher for drawn than written words, ps < .001.

Remember responses. d' for R responses was significantly higher for drawn words than written words, t (34) = 8.40, SE = .14, d = 1.42, p < .001 (WSR Z = 4.86, p < .001, PSS Z = 5.39, p < .001). Accuracy and hit rate were also higher for drawn than written words, ps < .001.

Know responses. d' for K responses was significantly *lower* for drawn words than written words, t (34) = 5.41, SE = .12, d = 0.91, p < .001 (WSR Z = 3.99, p < .001, PSS Z = 4.35, p < .001). Accuracy and hit rate were also lower for drawn than written words, ps < .001.

Table 1

Mean (Standard Error in Brackets) Accuracy, Hit Rate, False Alarm Rate, Proportion Remember (R) Responses, Sensitivity (d'), Overall, for R Responses Only, and for K Responses Only for Each Item Type, and Process Estimates for Each Item Type in Experiment 5A and 5B.

	Trial Type and Response					
	Draw			Write		
	Overall	Remember	Know	Overall	Remember	Know
Experiment 5A						
Accuracy	0.87 (.04)	0.89 (.03)	-0.02 (.04)	0.70 (.04)	0.58 (.05)	0.12 (.05)
Hit Rate	0.97 (.01)	0.91 (.03)	0.06 (.03)	0.80 (.03)	0.60 (.05)	0.20 (.03)
False Alarm Rate	0.10 (.04)	0.02 (.01)	0.08 (.03)	0.10 (.04)	0.02 (.01)	0.08 (.03)
d'	3.49 (.17)	3.53 (.16)	0.18 (.18)	2.65 (.17)	2.35 (.16)	0.84 (.18)
Experiment 5B						
Accuracy	0.84 (.03)	0.85 (.04)	-0.02 (.03)	0.60 (.04)	0.51 (.04)	0.10 (.02)
Hit Rate	0.96 (.01)	0.90 (.03)	0.06 (.03)	0.73 (.03)	0.55 (.03)	0.18 (.02)
False Alarm Rate	0.13 (.03)	0.05 (.02)	0.08 (.02)	0.13 (.03)	0.05 (.02)	0.08 (.02)
d'	3.11 (.15)	3.25 (.18)	-0.08 (.13)	2.00 (.14)	2.02 (.14)	0.61 (.09)
		Recollection	Familiarity		Recollection	Familiarity
Process Estimates						
Experiment 5A		0.82 (.04)	1.77 (0.42)		0.55 (.05)	1.66 (0.30)
Experiment 5B		0.87 (.03)	1.51 (0.25)		0.52 (.04)	1.18 (0.12)

Note. Process estimates are based on the subset of participants (3A = 18/35, 3B = 33/37) for whom these estimates could be computed. If one scores 100 percent 'Remember' hits, there is no way to estimate familiarity.

Process estimates. Nearly half of the sample gave exclusively 'Remember' responses during recognition of items that were drawn at encoding. As such, there was no opportunity for them to give 'Know' responses, making it impossible to calculate the influence of familiarity, in line with the underlying assumptions of the independence model upon which the equations are based (Yonelinas et al., 1998). Accordingly, I present here only the data from the 18 participants who provided at least one 'Know' response, thereby rendering the contribution of familiarity calculable. Because it is based on only a subset of the sample, the results should be interpreted with caution. That said, it appears that familiarity contributed little to the difference in recognition between drawn and written items. Of particular relevance is the absence of ‘Know’ responses to drawn items in the participants who were not included in the analysis. In any case, the formal results indicated that the contribution of recollection (See Figure 6) was greater for drawn, relative to written items, t (17) = 2.11, SE = .05, d = 1.34, p < .001 (WSR Z = 3.52, p < .001, PSS Z = 3.87, p < .001), while there was no difference in the contribution of familiarity between drawn and written items, t (17) = 0.27, SE = .42, d = 0.07, p = .79 (WSR Z = 0.31, p = .76, PSS Z = 0.53, p = .59).

The results from this experiment indicate that drawing not only improves overall recognition performance, but does so primarily by improving recollection. Specifically, d’ within R responses for drawn items was higher than within R responses for write items. While there was also a difference between trial types in d’ for K responses, it was in the opposite direction, such that d’ was higher for written items. However, the remember/know procedure likely underestimates the actual contribution of familiarity to recognition memory, because K responses, as taken from the task instructions, typically indicate that an item is familiar *and* not recollected (Yonelinas, 2002), rather than just familiar. Accordingly, using the independence remember/know method (e.g. Koen & Yonelinas, 2010; Yonelinas & Jacoby, 1995; Yonelinas et al., 1998), the estimated contribution of familiarity to overall recognition was calculated, based on extrapolation from the subset of items that were not recollected. Using this method, the estimated contribution of familiarity to memory was similar across drawn and written items, however, the contribution of recollection was substantially larger for drawn relative to written items. The current work is concerned with testing whether drawing improves memory by creating a rich and vivid recollection-based memory trace, and the evidence from this experiment suggests that this is indeed the case. However, because participants recollected so many drawn items, these estimates were only derived from a non-randomly selected subset of the sample. Accordingly, in Experiment 5B, I sought to reduce recollection in order to obtain a complete set of process estimates.

Experiment 5B

Because participants' level of recollection was so high in drawn items, it is difficult to interpret the process estimates from Experiment 5A, which were derived from only a subset of the sample. In Experiment 5B, I sought to bring performance down by increasing the number of studied items, and decreasing the encoding time, in order to derive a more exhaustive set of process estimates.

Participants. Participants were 37 undergraduate students (31 female) at the University of Waterloo, who completed the experiment for course credit. Participants ranged in age from 18 to 27 (M = 19.74, SD = 1.79).

Materials. The materials were the same as in Experiments 1 through 4A.

Procedure. For this experiment, I eschewed the pen and paper format, in favor of an entirely computer-based paradigm, and participants were all tested individually. Stimulus presentation and response recording were controlled using Python and displayed via an Acer One 10 touchscreen 2-in-1 notebook with 10.1-inch monitor and a stylus. Rather than 40 items, as in previous experiments, 80 items (40 drawn and 40 written) were now studied, and the encoding time was reduced from 40 to 20 seconds per item. In addition, there were now 60, rather than 40 lures (See Appendix for items). Because of the change in apparatus, participants now used a stylus to draw directly on the same screen where the words were presented (See Figure 5 for samples of the outcome of each trial type). The computer was converted into tablet mode during the encoding phase so that the screen was flat on the desk, analogous to the sheet of paper which was lying flat on the desk in previous experiments. Because page-flipping was no longer required, the idle time that was previously filled by page-flipping between trials was omitted. Instead, when the trial was over, the screen was cleared of all content, and the next prompt simply appeared where the participant had previously been drawing or writing.

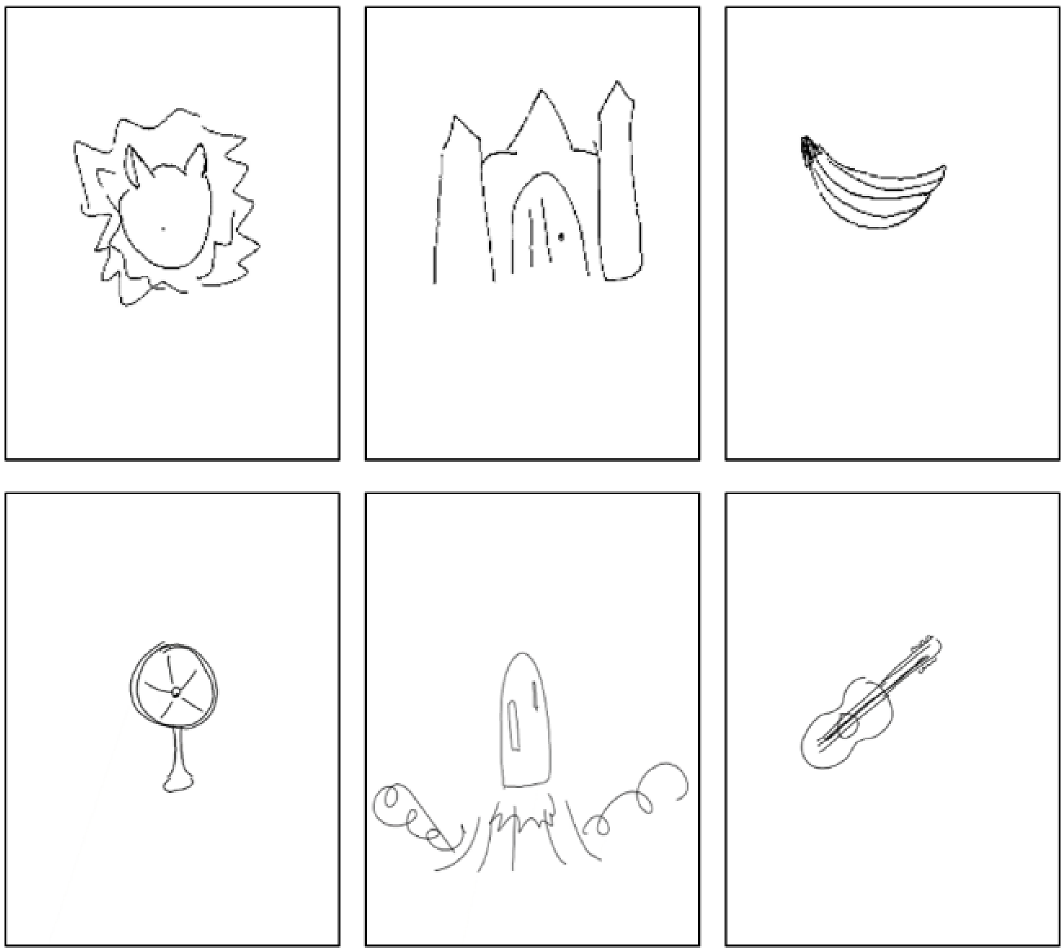

Figure 5. Sample drawings in Experiments 5B, 6, 7 and 8, created using a stylus on the touchscreen tablet

The filler task and retrieval phase were also completed using the computer. Rather than classifying the CRT tones on a sheet of paper, participants classified tones as low, medium or high using number keys (1, 2 and 3 respectively). The retrieval task required participants to press one of three keys to indicate their response, instead of filling in a bubble sheet as in Experiment 5A.

Results and Discussion

Overall recognition. Paired-samples t tests indicated that overall d' was higher for drawn words than for written words, t (36) = 12.57, SE = .09, d = 2.07, p < .001. Accuracy and hit rate were also higher for drawn than written words, ps < .001 (WSR Z = 5.23, p < .001, PSS Z = 5.83, p < .001).

Remember responses. d' for R responses was significantly higher for drawn words than written words, t (36) = 8.91, SE = .14, d = 1.52, p < .001. Accuracy and hit rate were also higher for drawn than written words, ps < .001 (WSR Z = 4.75, p < .001, PSS Z = 5.50, p < .001).

Know responses. d' for K responses was significantly *lower* for drawn words than written words, t (36) = 5.81, SE = .12, d = 1.01, p < .001. Accuracy and hit rate were also lower for drawn than written words, ps < .001 (WSR Z = 4.47, p < .001, PSS Z = 4.50, p < .001).

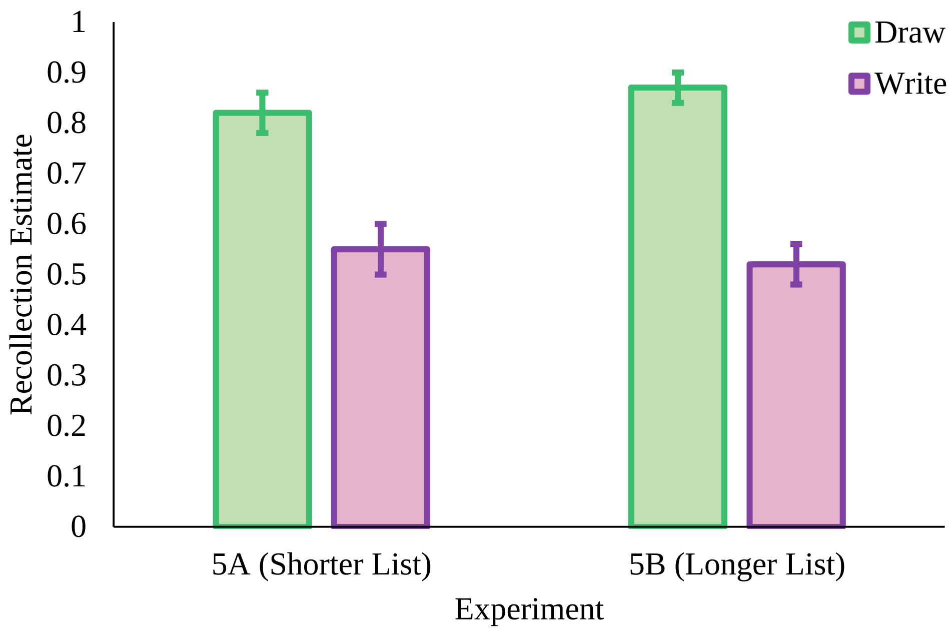

Figure 6. Recollection estimates in Experiments 5A and 5B in the Draw and Write trial type. Error bars represent standard error of the mean.

Process estimates. In this iteration, only 4 participants gave exclusively 'Remember' responses during recognition of items that were drawn at encoding, leaving 33 subjects where process estimates were calculable. The contribution of recollection was greater for drawn, relative to written items, t (32) = 8.76, SE = .04, d = 1.53, p < .001 (WSR Z = 4.73, p < .001, PSS Z = 5.22, p < .001), while the difference in the contribution of familiarity was not significant, t (32) = 1.28, SE = .26, d = 0.26, p = .21 (WSR Z = 0.74, p = .46, PSS Z = 0.35, p = .73).

The data here replicate the findings from Experiment 5A completely, demonstrating that drawing leads to significantly better memory performance when looking at overall recognition, as well as d' within exclusively R responses. Similarly, when calculating process estimates, drawing led to an increase in the contribution of recollection, but not familiarity to recognition performance. Taken together, the findings from these experiments suggest that drawing provides rich contextual information which later selectively improves recollection.

Experiment 6

If, as the evidence thus far indicates, the benefit that drawing affords memory is most strongly driven by recollection, it follows that this benefit should be reduced or eliminated when the opportunity to engage in the recollection process is limited or taken away. In Experiment 6, I employed a response-deadline procedure to test this hypothesis. As discussed in the introduction to this line of experiments, one of the main points of agreement, and the prevailing view in dual-process models, is that familiarity is relatively fast, and recollection is slower (e.g. Yonelinas, 2002). Evidence for the temporal dissociation of these processes comes from event-related potential (ERP) studies demonstrating that each is uniquely associated with a different component. Specifically, recollection, including source memory decisions, is associated with a LPC originating in the parietal lobe (Luo et al., 2007; Senkfor & Van Petten, 1998), while familiarity is associated with an earlier frontal and temporal component (the FN400, Addante et al.,2012; Meng, Ye, & Gonsalves, 2014). Thus, because recollection is associated primarily with a later component (Mecklinger, 2000; Rugg and Curran, 2007; Wilding and Ranganath, 2011; Woodruff, Hayama, & Rugg, 2006), it follows that recollection unfolds more slowly than familiarity.

The temporal difference between the two processes has also been supported by behavioral work demonstrating that speeding recognition responses leads to a greater reliance on familiarity (e.g. Yonelinas & Jacoby, 1994), or to responses based solely on familiarity (Besson, Ceccaldi, Didic, & Barbeau, 2012; Besson, Ceccaldi, Tramoni, Felician, Didic, & Barbeau, 2015). Furthermore, there is evidence from a rich neuroimaging literature that has explored the time course of these processes and converged on the same point: that familiarity is a faster process than recollection (e.g. Brown & Aggleton, 2001; Staresina, Fell, Do Lam, Axmacher, & Henson, 2012). Response-deadline tasks are those that exploit this temporal distinction to determine the contribution of familiarity to memory (e.g. Sauvage, Beer & Eichenbaum, 2010). Evidence has shown that when a recognition task is speeded, relative to when it is not, responses based on recollection are offset or eliminated, while familiarity remains intact (Benjamin & Craik, 2001; Boldini, Russo, & Avons, 2004; Bowles et al., 2007; Duzel, Yonelinas, Mangan, Heinze, & Tulving, 1997; Sauvage et al., 2010; Toth, 1996; Yonelinas & Jacoby, 1994; but see, Gardiner, Ramponi, Richardson-Klavehn, 1999). With this, and the findings reported in the current work indicating that the benefit of drawing relies on recollection, I expect that the drawing effect will be substantially reduced or eliminated as a result of speeding retrieval responses, but will remain when a longer response window is allowed.

Method

Participants. Participants were 70 (35 per group) undergraduate students (51 female) at the University of Waterloo, who completed the experiment for course credit. Participants ranged in age from 17 to 33 (M = 20.11, SD = 2.27). Performance in two of the participants in the speeded group however, was below chance, meaning that they had an equal or greater amount of false alarms than they did hits. Accordingly, these participants were replaced with additional participants to compose the full sample of 70[2].

Materials. Materials were the same as in Experiment 1 through 5A.

Procedure. For this experiment, I again employed an entirely computer-based paradigm, and participants were all tested individually. Stimulus presentation and response recording were controlled using Python and displayed via a Toshiba Portege M750 touchscreen laptop/tablet with 12.5 inch monitor and a stylus. The timing and display of prompts and words was identical to Experiment 1, and the manner in which the act of drawing and the filler task was implemented was identical to Experiment 4B.

For retrieval, participants completed an old/new recognition task, incorporating a response deadline procedure. Participants were evenly divided into two groups. In the Speeded group, a response deadline of 800 ms was imposed for each trial on the recognition test, while participants in the Control group had 2400 ms to respond to each trial. In either case, they were told that if an item was old, meaning that they had previously studied it, that they were to press the number 1, while if it was new, they were to press the number 0. Participants in each group were informed of how much time they had to respond to each trial, and that they would hear a short, shrill tone if they missed the deadline in either condition. They were also told that this tone served to remind them to respond more quickly on the next trial.

Results and Discussion

Hit Rate. Hit rate was analyzed using a 2 x 2 mixed measures ANOVA with Group (Speeded, Control) as a between-subjects factor and Trial Type (Draw or Write) as a within-subjects factor. There was a significant main effect of both Group, F (1, 68) = 35.66, MSE = .05, p < .001, η^2 =.34, and Trial Type, F (1, 68) = 59.85, MSE = .01, p < .001, η^2 =.47, as well as a significant interaction, F (1, 68) = 19.84, MSE = .01, p < .001, η^2 =.23. In unpacking the interaction, results showed that drawing led to significantly more hits than did writing in both groups, though the benefit was smaller in the speeded group, t (34) = 2.71, SE = .02, d = .51, p < .025 (.05/2) (WSR Z = 2.44, p < .05, PSS Z = 1.95, p = .052), than in the control group, t (34) = 7.66, SE = .03, d = 1.32, p < .001 (WSR Z = 4.98, p < .001, PSS Z = 4.97, p < .001). There was a large significant decrease in both drawn hit rate, t (68) = 8.62, SE = .04, d = 2.06, p < .001 (WRS Z = 6.74, p < .001), and in write hit rate, t (68) = 3.01, SE = .05, d = 0.72, p < .01 (WRS Z = 3.01, p < .01), in the Speeded, relative to the Control group.

Table 2

Mean (Standard Error in Brackets) Hit Rate, False Alarm Rate, Sensitivity, and Response Time (RT) for Each Item Type in Experiment 6

	Draw	Write	New
Speeded			
Hits and false alarms	0.63 (.03)	0.57 (.03)	0.24 (.02)
Sensitivity (d')	1.09 (.12)	0.95 (.11)	
Correct RT	612.83 (16.98)	587.93 (20.83)	576.89 (8.93)
Control			
Hits and false alarms	0.94 (.03)	0.72 (.03)	0.03 (.01)
Sensitivity (d')	3.53 (.12)	2.54 (.11)	
Correct RT	861.80 (16.98)	918.60 (20.83)	908.14 (22.99)

Sensitivity (d'). Sensitivity was analyzed in the same manner as hit rate. There was a significant main effect of both Group, F (1, 68) = 193.59, MSE = 0.74, p < .001, η^2 =.74, and Trial Type, F (1, 68) = 81.75, MSE = 0.14, p < .001, η^2 =.55, as well as a significant interaction, F (1, 68) = 44.92, MSE = 0.14, p < .001, η^2 =.40. The interaction was driven by a larger drawing effect in the control group, t (34) = 8.94, SE = .11, d = 1.52, p < .001 (WSR Z = 4.93, p < .001, PSS Z = 4.97, p < .001), than in the speeded group, t (34) = 2.47, SE = .06, d = .42, p < .025 (.05/2) (WSR Z = 2.24, p < .05, PSS Z = 1.95, p = .052).

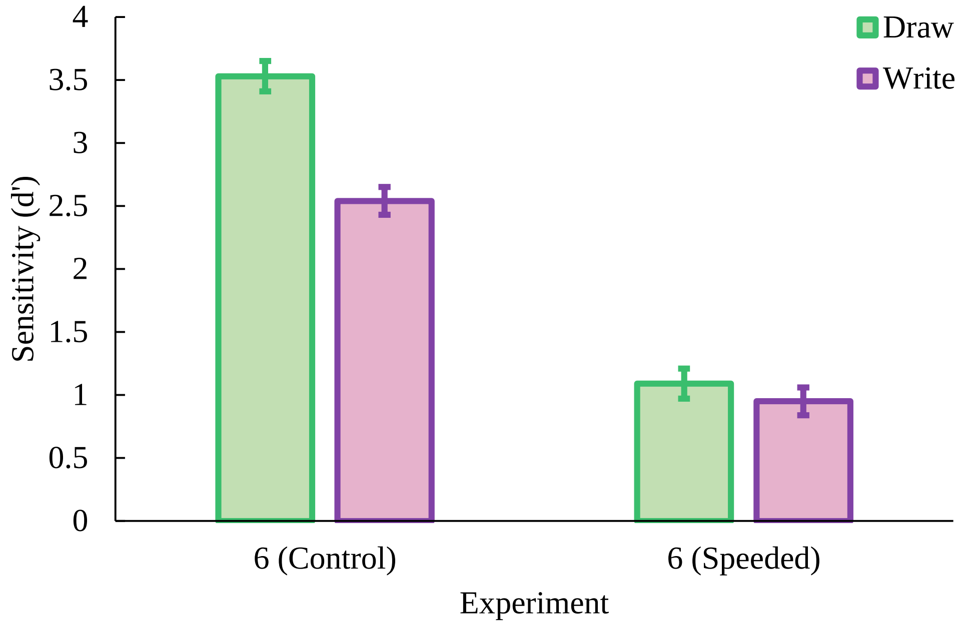

Figure 7. Sensitivity (d') in recognition memory decisions in Experiment 6 when responses were at control pace, relative to speeded. Error bars represent standard error of the mean.

Response time. RTs for hits only were included in the analysis. RTs were analyzed in a 2 x 2 mixed-measures ANOVA with Group (Speeded, Control) as a between-subjects factor and Trial Type (Draw, Write) as a within-subjects factor. There was a main effect of Group, F (1, 68) = 139.34, MSE = 21099.38, $p < .001$, η^2 =.67, as well as a significant interaction, F (1, 68) = 13.95, MSE = 4187.58, $p < .001$, η^2 =.17. The main effect of Trial Type was not significant, F (1, 68) = 2.13, MSE = 4187.58, p = .15, η^2 =.03. The interaction was driven by faster response times for drawn than written items in the Control group, t (34) = 2.68, SE = 21.18, d = 0.47, $p < .025$ (0.05/2) (WSR Z = 2.31, $p < .05$, PSS Z = 2.57, $p < .05$), and slower response times for drawn than written items in the Speeded group, t (34) = 4.56, SE = 5.46, d = 0.77, $p < .001$ (WSR Z = 3.67, $p < .001$, PSS Z = 3.38, $p < .01$). RTs for false alarms were analyzed using an independent samples t test. RTs were slower in the Control than the Speeded group, t (68) = 13.43, SE = 24.66, d = 3.51, p < .001 (WRS Z = 7.19, p < .001).

As predicted, when the contribution of recollection to memory was reduced through speeding memory responses (e.g. Besson et al., 2012; Yonelinas & Jacoby, 1994), the drawing effect was also substantially reduced, though it remained significant. While performance on both draw and write trial types showed decreased memory performance in the speeded, relative to the control condition, the data indicate that the draw trial type suffered more from the imposed response-deadline. This suggests that the benefit that drawing affords memory relies largely on recollection of specific episodic details from the initial encoding experience, which likely require more time to unfold before exerting their influence on memory decisions (e.g. Besson et al., 2012; Mandler, 1980). Interestingly, there remained a relatively small benefit of drawing lingering in the speeded condition, suggesting one of two possibilities. The first is that familiarity also contributes in part to the benefit to memory that participants enjoy as a result of drawing, while the second is that even in the speeded condition, there was some early influence of recollection. In support of the latter possibility, analyses showed that drawn items were responded to more slowly than were the written items in the speeded condition, suggesting that in some cases, participants might have delayed their response just long enough for the beginnings of a recollective experience to occur. Nevertheless, the overall findings from this experiment are consistent with those from Experiments 5A and 5B, suggesting that the benefit that drawing provides to recognition memory performance is driven primarily by the creation of a recollection-based memory trace, allowing for a small contribution of familiarity as well.

CHAPTER IV: The Components of Drawing

In previous chapters of this dissertation and in other work (Wammes et al., 2016; submitted), my findings have highlighted that drawing is an efficacious strategy, effective in improving memory (relative to writing) across a variety of encoding materials (words, pictures), lengths (words, definitions; Wammes, Meade & Fernandes, submitted), retrieval tasks (free recall, short answer, source memory, remember-know; Wammes, Meade & Fernandes, in press) and populations (healthy older adults, MCI, and AD; Meade, Wammes & Fernandes, submitted; in progress). Moreover, the benefit seems resilient regardless of the amount of time provided for drawing (as little as 4 seconds), and seemingly regardless of self-reported artistic talent (Wammes et al., 2016). Most importantly, it is an effective manipulation in between-participants designs as well (Chapter II, Experiment 4), indicating that distinctiveness alone is not enough to account for the beneficial effects of drawing on memory. An encoding task of such potency is in limited company within the memory literature, and thus is an intriguing target for applied research, but also a task whose underlying mechanism is important to understand. Chapter III demonstrated that while drawing may improve both recollection and familiarity, recollective processes are especially critical in driving the benefit that drawing affords memory. This finding is perhaps not surprising, as drawing is an especially rich activity to engage in. However, testing whether drawing improved recollection was a necessary precondition prior to investigating my prediction that the benefit of drawing is driven by integrated multimodal components. To explain my proposed mechanism in more detail, consider the act of creating a drawing derived from a presented target word: One must first generate an internal representation of what it is they intend to draw, then implement a specific manual motor program to translate that image to the page, and finally are left with a pictorial representation.

In line with this description of the process, the first component I argue contributes to the beneficial effects of memory is the ‘elaborative component’. I define this component as the generative processes one must engage in upon viewing a word to bring to mind an internal representation of what that item looks like. This would require deep semantic processing of the item, in concert with visual imagery. The second component is the ‘motoric component’, which I define as the manual motor program one must implement to translate one’s internal image from mind to paper. While the act of writing, for instance, requires the motor system as well, the movements are not directly associated with the meaning of the word, which is a necessary aspect of my proposed motoric component. The third and final component is the ‘pictorial component’, which is simply the visual processing of the picture that one is creating or has created.

Each of these ‘components’ of the drawing process fosters a clear analogy with a rich corresponding literature. That is, the elaborative component would be something akin to a hybrid of deep levels of processing and visual imagery (Craik & Lockhart, 1978). Similarly, the motor component draws clear parallels with the enactment effect literature, wherein producing a motor action in response to a word improves memory (Cohen, 1983; Engelkamp & Zimmer, 1989). Lastly, the pictorial component would be well captured by the classic picture superiority effect (Paivio, 1971; Paivio et al., 1968). Together these components provide three, largely independent, multimodal sources of contextual information one could retrieve to reconstitute the encoding experience. In the current work, I test whether memory performance improves as more components are added during encoding. Furthermore, I test the extent to which each of these components is important in driving the benefit that drawing affords to memory. Across two experiments, I systematically dissociate the components from one another using contrasts between sets of encoding instructions. Across these tasks, I manipulate the presence of absence of the three proposed components of the drawing process. In so doing, I can estimate the relative contribution of each proposed component, as well as gain a sense of the importance of the *number* of sources of information in driving memory increases more generally (Table 3).

Table 3

Trial Types in Experiments 7 and 8, Indicating the Presence (Shaded) or Absence (White) of the Elaborative (Elab), Motor (Mot), and Pictorial (Pic) Components, as Well as a Short Description of Each Task.

	Component			
Trial Type	Elab	Mot	Pic	Task Description
E7				
Draw	■	■	■	An image is drawn of what the word represents
Imagine	■		■	An internal visual image of the word is created
Trace		■	■	A faded image of what the word represents is traced over
View			■	An image of what the word represents is viewed
E8				
Blind	■	■		Image is drawn, screen does not display what is being drawn
Draw	■	■	■	An image is drawn of what the word represents
Imagine	■			An internal visual image of the word is created
Trace		■	■	A faded image of what the word represents is traced over
View			■	An image of what the word represents is viewed
Write				The word is written out repeatedly

Experiment 7

In Experiment 7, I created four tasks with the aim of determining whether, in line with an integrated components mechanism, memory performance would scale linearly with the number of components a particular trial type incorporates. I opted to fix a task engaging the pictorial component alone as the baseline condition. Next, I ensured that the pictorial component was included in all other tasks, such that the key differences between trial types would be the presence or absence of both the motor and the elaborative component. Accordingly, in this experiment the baseline trial type was ‘view’, wherein participants simply viewed a pictorial representation of the word. The baseline was compared to three additional trial types – one which added only the motor component, one which added only the elaborative component, and ‘draw’, which added both. As a secondary analysis. I aimed to derive an estimate of the relative contribution of each proposed component through subtractions and contrasts.

Method

Participants. Participants were 40 undergraduate students (31 female) at the University of Waterloo who completed the study in two parts, in exchange for partial course credit. One participant did not provide their age, but the remainder of the sample ranged in age from 17 to 23 ($M = 19.49$, $SD = 3.60$).

Materials. The word list, created based on the names of Snodgrass images (Snodgrass & Vanderwart, 1980) was expanded to 130 words in length. A comprehensive list is provided in the Appendix. A random subset of 50 of the words were selected as lures. The remaining 80 were divided into four separate sets of 20, matched on frequency, length and number of syllables. This fully counterbalanced design ensured that each set was allocated to each of the four trial types once. Within each of these counterbalancing assignments, two different pseudo-random orders or mixed presentation schedules were implemented.

In a somewhat archaic implementation, 80-page notepads were painstakingly preassembled to match the pseudo-random orders. Pages that corresponded to a trial where a participant was to ‘draw’ or ‘imagine’ were left blank. Pages that corresponded to ‘view’ trials, had a printed image of what the word represents taped to them, and pages that corresponded to ‘trace’ trials, had the printed image taped on the underside of the page, such that the participant could view the item through the paper to trace it. I assure the horrified reader that the materials and the approach improved to one more reflective of the current century in Experiment 8.

Participants were provided with a sharpened pencil with which to draw, and a backup pencil was provided in the event of any catastrophic graphite mishaps.

Presentation of words, and collection of eventual recognition data was achieved using E-Prime 2.0.

Procedure. Participants completed the experiment in two separate sessions, where the second session occurred two days following the first session. The impetus for the introduction of this delay was simply the potency of drawing as a mnemonic, which has led to difficulties driving participants' performance away from ceiling. In piloting, I attempted various design alterations, including decreased encoding time, inclusion of more than 100 items, and packing up to 20 minutes of demanding verbal and visual distracting tasks into the retention interval. However, participants remained at ceiling for many trial types, precluding reliable estimates of performance in every trial type. Accordingly, I decided to conduct this experiment in two parts, which proved to be a fruitful method of reducing performance and avoiding ceiling.

Encoding. In session 1, participants were seated in front of a computer screen and presented with one of the preassembled pads. They were instructed that they would view written prompts, followed by words and images, and that they would have 15 s to perform the task that the prompt indicated. There were four tasks/trial types and instructions were as follows:

For 'draw' trials:

> If the prompt is 'draw' we ask that you use the pad of paper provided to draw a picture illustrating the word on the screen. You won't see a picture of the object in this case. Instead, you draw the picture of the object yourself.

For 'trace' trials:

If the prompt is 'trace' we ask that you use the pad of paper and pencil provided to trace the picture, which you will be able to see through the page.

For ‘imagine’ trials:

If the prompt is 'imagine' we ask that you create a mental picture of what the word represents. In other words, you should think about what that thing is and how it looks. Create an image in your mind's eye.

And for ‘view’ trials:

If the prompt is 'view' we ask that you view the picture of what the word represents, which will be presented on the pad of paper provided. In other words, just look at the picture on the page.

To clarify the instructions, the experimenter produced a ‘practice pad’, which was only four pages long, with an example of what the pad would look like for each trial type. The experimenter walked through the practice pad with the participant to ensure understanding of the instructions. Once any questions had been addressed, the experiment began. On the computer screen for each trial, participants were shown a prompt word in blue font (1.5 s), immediately followed by a fixation (0.5 s) then a target word in black font above an image of what it represents (1.5 s). They then had 15 s to perform the task that the prompt indicated. When the time to complete the prompted task had elapsed, participants heard a short tone, then had 4 s to flip their pad to the next page and prepare for the subsequent item. Participants studied a mixed list of 80 words, 20 in each trial type. This encoding session took approximately 35 minutes, including instructions.

Retrieval. Participants returned two days later to complete session 2. Their memory for the studied items was tested using a basic old-new recognition task, wherein the 80 studied items were presented, intermixed with 50 lures. Studied words (without associated images) were presented one at a time in the center of the screen, and participants were instructed to press 1 for old items, and 0 for new items. It was made clear to participants that old items included items that were traced, drawn, imagined, or viewed in the initial session. The response options ('Old (1), New (0)') were also explicitly presented on the screen throughout the retrieval task, at the bottom of the screen. Responses were self-paced, and the next trial was presented upon response to a trial.

Results and Discussion

Overall recognition. Repeated-measures ANOVA indicated a main effect of trial type, F (3, 117) = 28.08, MSE = .17, p < .001, η_p^2 =.42, and a massive linear effect, F (1, 39) = 57.03, MSE = .23, p < .001, η_p^2 =.59. This main effect was driven by d' being significantly higher in draw, relative to trace, t (39) = 4.82, SE = .10, d = 0.77, p < .001, imagine, t (39) = 8.16, SE = .08, d = 1.31, p < .001, and view, t (39) = 7.23, SE = .11, d = 1.14, p < .001, trial types. d' was also higher in the trace than in the view trial type, t (39) = 3.51, SE = .09, d = 0.56, p < .01. While d' was numerically higher in trace than imagine trial types, t (39) = 2.23, SE = .09, d = 0.35, p = .03, and imagine higher than view, t (39) = 1.57, SE = .08, d = 0.25, p = .12, these comparisons failed to reach Bonferroni-corrected significance thresholds of 0.008, and the former comparison was only marginal even with an uncorrected threshold. Hit rate and accuracy analyses followed the identical pattern, though the contrast of trace and imagine trial types was closer to significance, p = 0.057. The structure of the data is readily identifiable in Figure 8.

Table 4

Mean (Standard Error in Parentheses) Hit Rate, False Alarm Rate, Sensitivity, and Response Time (RT) for Each Item Type in Experiment 7

	Hits	False alarms	Accuracy	Sensitivity (d')	Correct RT
Trial Type					
Draw	0.88 (0.02)	0.09 (0.01)	0.79 (0.02)	2.72 (0.13)	1278.98 (56.50)
Trace	0.76 (0.03)		0.67 (0.03)	2.27 (0.12)	1466.79 (60.80)
Imagine	0.70 (0.02)		0.61 (0.02)	2.06 (0.12)	1456.46 (58.70)
View	0.65 (0.03)		0.56 (0.03)	1.93 (0.13)	1471.51 (73.11)

Response time. Repeated-measures ANOVA indicated a main effect of trial type, F (3, 117) = 6.14, MSE = 56580.31, p < .01, η_p^2 =.14, as well as a significant linear pattern, F (1, 39) = 14.51, MSE = 47635.85, p < .001, η_p^2 =.27. Drawn items were responded to more quickly than traced, t (39) = 3.68, SE = 48.24, d = 0.58, p < .01, imagined, t (39) = 3.62, SE = 51.84, d = 0.57, p < .01, and viewed items, t (39) = 3.77, SE = 51.02, d = 0.63, p < .01. All other comparisons failed to reach significance, ps > 0.77.

Component estimates. The trial types were designed with the intention of systematically varying the presence or absence of the three proposed components of drawing (i.e. pictorial, elaborative, motor). Specifically, all trial types in this experiment contain a basic verbal memory trace, plus the benefit of the pictorial component because a picture was presented after the word at encoding. The imagine trial type *also* had the elaborative component, but did not have the motor component. In contrast, the trace trial type *also* had the motor component, but did not have the elaborative component (See Table 3). Following this logic, one can subtract accuracy in the view condition from accuracy in the imagine trial type to derive a rough estimate of the relative contribution to memory of the elaborative component. Similarly, one can subtract view from trace to obtain an estimate of the motor contribution. Now, by adding performance in the view condition to these subtractions, I can obtain a measure of the added contributions of the three proposed components, and determine how they compare to performance in the draw condition. These operations were performed, and a paired samples t-test revealed that d' was higher for the draw (M = 2.72, SE = 0.13) trial type than for the additive components (M = 2.39, SE = 0.14), t (39) = 3.25, SE = 0.10, d = 0.51, p < .01.

Findings from Experiment 7 replicated previous findings, which demonstrated that drawing led to better memory than imagining items or viewing pictures (Wammes et al., 2016; Wammes et al., submitted). In addition, these findings indicate for the first time that drawing leads to better memory than tracing, a non-elaborative form of drawing. Overall, as evidenced by the linear trend and the general pattern of the contrasts between trial types, the results indicate that as more of the proposed components are added to the encoding task, memory accuracy increased. This is in line with the predictions derived from my 'integrated components' model. The baseline was pictorial memory in the view condition. Adding an elaborative aspect ('imagine') increased memory slightly, though not significantly, and adding a motor aspect ('trace') increased memory significantly. Over and above these two trial types, 'draw' improved accuracy even more, ostensibly as a result of adding the remaining third component. A secondary finding was that drawing led to better memory accuracy than the three components combined, suggesting that there may be some additional benefit to drawing resultant from the seamless integration of these components.

There are some limitations to the design here though, especially as it pertains to making conclusions about the importance of the number of components engaged by the encoding task. In particular, using 'view' as the baseline may have been an issue. Because I chose to use 'view', a proxy for the pictorial component, as the baseline, pictures were presented in every trial type. While this is not damaging in the trace trial type, it may have undermined the supposed elaborative aspects of the imagine trial type. By providing the participant with an extant visual representation of the item they were to imagine, I likely constrained – or worse - prevented entirely any elaborative thought. This would also explain why ‘imagine’ did not lead to performance that was significantly higher than ‘view’. The second issue is that there was only one trial type which engaged a single component, which precludes attaining a more general sense of how memory is improved from having just one extra source of contextual information. In contrast, there were two tasks designed to engage two of the components each. I also did not include any trial types proposed to have none of these components. The foregoing issues render it more difficult to conclude that the proposed components add linearly to memory performance – an issue I remedied in Experiment 8.

Experiment 8

In the current experiment, it was my aim to not only conceptually replicate the findings from Experiment 7, but also to design a more comprehensive set of trial types. Specifically, I aimed to use ‘write’, which has been my most oft-used baseline, as the control condition. Writing was important to include because it arguably engages none of the proposed components. From here, ‘imagine’ and ‘view’ were employed as trial types which had one component (elaborative; pictorial). In addition, I could also use ‘trace’ and a new condition, ‘blind-draw’, as trial types which had two components (motoric and pictorial; elaborative and motoric). These would again all be compared with the ‘draw’ trial type (See Table 3).

Method

Participants. Participants were 39 undergraduate students (27 female) at the University of Waterloo who completed the study in two parts, in exchange for partial course credit. One participant did not provide their age, but the remainder of the sample ranged in age from 17 to 28 (M = 20.50, SD = 2.38).

Materials. The list from Experiment 1 was expanded to include 228 words (See Appendix). For each participant, the words were completely randomly assigned to be lures (75 words), and to each of the 6 trial types (20 words each). Rather than re-instantiating the laborious process of creating preassembled pads, I switched again to an entirely computer-based paradigm. Stimulus presentation and response recording were controlled using Python and displayed via a Toshiba Portege M750 touchscreen laptop/tablet with 12.5-inch monitor and a stylus. The program was designed such that 'draw' and 'write' trial types displayed a blank white screen which participants could draw on in black lines using the stylus. For 'trace' trials, a picture of the item was displayed in a faded grey, and participants were able to draw over it in the same fashion as in the preceding two trial types. For 'blind' trials, the screen was blank, and participants could again draw, but instead of placing lines on the screen where one had drawn, the program simply collected the coordinates of the stylus at each refresh of the screen. This was done as a manipulation check so the images participants created could be regenerated later from the collected coordinates. For 'view' trials, a picture of the item was presented, and for 'imagine' trials, a blank screen was displayed. In both of these trial types, running the stylus across the screen would have no effect.

Procedure. As in Experiment 7, participants completed the experiment in two separate sessions separated by two days.

Encoding. The laptop was placed in tablet mode during the encoding phase, such that the screen was flat on the desk, and analogous to the pad of paper in Experiment 7. Instructions were presented on the tablet screen, and participants were given the same instructions as in Experiment 7. Instructions were also read aloud by the experimenter, and the participant touched the screen twice to trigger the next instruction screen when told to do so. There were now six tasks/trial types. Four of these were the same as in Experiment 7, and the instructions were simply modified to reflect that the participant could now use the stylus to create their drawings directly on the tablet screen. The instructions associated with the two new trial types were as follows:

For ‘blind’ trials:

> If the prompt is 'blind' we again ask that you use the stylus to draw a picture illustrating the word on the screen. In this case you will not see a picture of the object OR the lines you are creating with the stylus.

For ‘write’ trials:

> If the prompt is 'write' we ask that you use the stylus to clearly and carefully write the word multiple times. In other words, just continue to rewrite the word until time runs out.

Beyond this, the timing and implementation of the encoding phase was the same as in Experiment 7, except that the idle time previously allowed to flip the page of the pad was omitted, as the next prompt would simply appear on the screen at the end of a given trial.

Retrieval. The retrieval test was the same as in Experiment 7.

Results and Discussion

Overall recognition. Repeated-measures ANOVA indicated a main effect of trial type, F (5, 190) = 48.61, MSE = .18, p < .001, η_p^2 =.56, and a massive linear effect, F (1, 38) = 148.34, MSE = .32, p < .001, η_p^2 =.79. In what follows, I unpack the structure of the trial types, and the source of this main effect.

First, the main effect was driven by draw being a more potent manipulation than most other trial types. Specifically, d' was higher in draw, relative to trace, t (38) = 7.33, SE = .07, d = 1.17, p < .001, imagine, t (38) = 9.38, SE = .10, d = 1.50, p < .001, view, t (38) = 9.65, SE = .10, d = 1.55, p < .001, and write, t (38) = 10.44, SE = .12, d = 1.67, p < .001, trial types. Draw also led to higher d' than blind, though this comparison failed to reach the Bonferroni-corrected significance level of 0.0033 (0.05/15), t (38) = 2.72, SE = .09, d = 0.44, p = .01.

Next, the effect seems to be driven by 'two-component' manipulations leading to greater memory performance than 'one-component' manipulations. The blind trial type led to significantly higher performance than imagine, t (38) = 6.94, SE = .09, d = 1.11, p < .001, and view, t (38) = 6.75, SE = .11, d = 1.09, p < .001. Mirroring this pattern, performance in the trace trial type was significantly higher than in both the imagine, t (38) = 4.39, SE = .08, d = 0.70, p < .001, and the view trial types, t (38) = 5.30, SE = .09, d = 0.85, p < .001. The blind trial type did also lead to higher performance than trace, though not significantly based on corrected thresholds, t (38) = 3.05, SE = .09, d = 0.49, p = .004.

There was also some indication that 'one-component' manipulations led to greater memory performance than the control or 'zero-component' manipulation. That is, both view, t (38) = 2.37, SE = .10, d = 0.38, p = .02, and imagine, t (38) = 3.32, SE = .10, d = 0.53, p < .003, showed higher d’ than write, though the former failed to reach corrected thresholds for significance. The imagine and view trial types did not differ from one another, t (38) = 1.02, SE = .10, d = 0.16, p = .32.

Perhaps not surprisingly, the ‘two-component manipulations’ were also significantly higher than write in terms of d’, as blind, t (38) = 10.07, SE = .10, d = 1.62, p < .001, and trace, t (38) = 8.13, SE = .09, d = 1.30, p < .001, were both significantly higher than write. Hit rate and accuracy analyses followed the same pattern.

Table 5

Mean (Standard Error in Parentheses) Hit Rate, False Alarm Rate, Sensitivity, and Response Time (RT) for Each Item Type in Experiment 8

	Hits	False alarms	Accuracy	Sensitivity (d’)	Correct RT
Trial Type					
Draw	0.91 (0.01)	0.13 (0.01)	0.78 (0.02)	2.55 (0.10)	1315.85 (51.03)
Blind	0.84 (0.02)		0.72 (0.02)	2.29 (0.11)	1367.49 (67.29)
Trace	0.78 (0.02)		0.65 (0.02)	2.01 (0.10)	1443.26 (54.77)
Imagine	0.65 (0.03)		0.52 (0.03)	1.65 (0.10)	1558.39 (70.78)
View	0.61 (0.04)		0.48 (0.03)	1.55 (0.10)	1764.39 (136.26)
Write	0.53 (0.03)		0.41 (0.03)	1.31 (0.10)	1704.96 (89.60)

Response time. Repeated-measures ANOVA indicated a main effect of trial type, F (2.27, 86.23) = 8.77, MSE = 325322.27, p < .001, η_p^2 =.19, as well as a significant linear pattern, F (1, 38) = 49.02, MSE = 120196.76, p < .001, η_p^2 =.56. Drawn items were responded to more quickly than traced, t (38) = 3.30, SE = 38.60, d = 0.53, p < .003, imagined, t (38) = 3.79, SE = 64.12, d = 0.63, p < .003, viewed, t (38) = 3.85, SE = 116.67, d = 0.80, p < .001, and written items, t (38) = 5.03, SE = 77.30, d = 0.89, p < .001. However, drawn items were not responded to more quickly than blind-draw items, t (38) = 1.02, SE = 50.53, d = 0.17, p = .31. Blind-drawn items were responded to more quickly than imagined, *t* (38) = 2.98, *SE* = 64.30, *d* = 0.48, *p* = .005, and viewed items, *t* (38) = 4.31, *SE* = 92.11, *d* = 0.98, *p* < .001, though the former failed to surpass Bonferroni-corrected significance thresholds (0.05/. Responses to traced items were not significantly faster than either blind-drawn, *t* (38) = 1.33, *SE* = 57.19, *d* = 0.22, *p* = .19, imagined, *t* (38) = 1.95, *SE* = 59.27, *d* = 0.32, *p* = .058, or viewed, *t* (38) = 2.63, *SE* = 122.25, *d* = 0.51, *p* = .012, items, though this latter contrast between trace and view was marginal. Responses to the blind-draw, *t* (38) = 4.04, *SE* = 83.49, *d* = 0.66, *p* < .001, and trace trial types, *t* (38) = 3.99, *SE* = 65.64, *d* = 0.73, *p* < .001, were faster than to the write trial type, while responses to the view, *t* (38) = 0.43, *SE* = 139.81, *d* = 0.07, *p* = 0.67, and imagine trial type, *t* (38) = 1.99, *SE* = 73.22, *d* = 0.33, *p* = .053, were not. Imagine was also not significantly different from view, *t* (38) = 1.66, *SE* = 123.78, *d* = 0.30, *p* = .105.

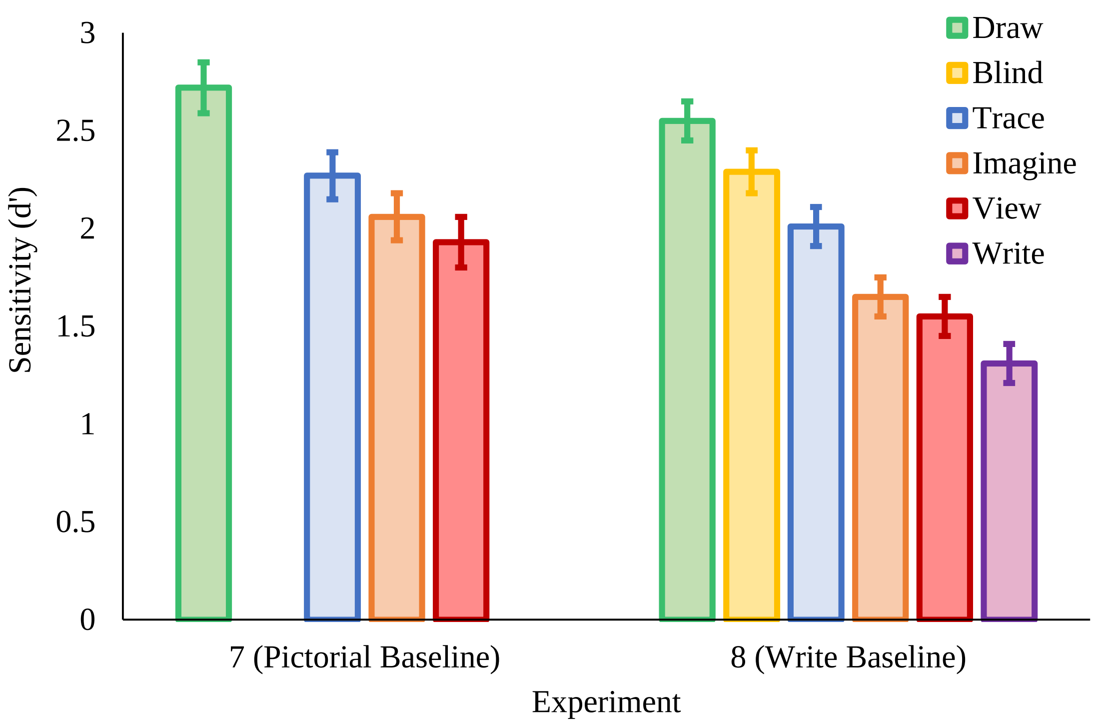

Figure 8. Sensitivity (d') in recognition memory decisions in Experiments 6 and 7, for the Draw, Blind, Trace, Imagine, View and Write trial types. Error bars represent standard error of the mean.

Component estimates. Similar to Experiment 7, the trial types here were designed with the intention of estimating the beneficial effects of the elaborative, motor and pictorial components of drawing. Our baseline was the write condition, which had none of the proposed components. The imagine trial type engaged the elaborative component, while the view trial type engaged the pictorial component. Similarly, the blind trial type involved both elaborative and motor information, while the trace trial type involved both motor and pictorial information. The draw trial type, as before, required all three components in concert (See Table 3). Again, I conducted subtractions to attain a rough estimate of the relative contribution to memory of each component. The contributions of the elaborative (imagine – write), motor (trace – view) and pictorial components (view – write) were estimated, and compared to the contribution of drawing (draw – write). A paired samples t-test revealed that d' was marginally higher for the benefit of the draw trial type (M = 1.23, SE = 0.12) than for the additive components (M = 1.03, SE = 0.17), t (38) = 1.91, SE = 0.11, d = 0.34, p = .063. It is worth noting that the motor component might also be computed by subtracting imagine from blind (M = 1.21, SE = 0.16), which leads to a non-significant difference from draw, t (38) = 0.19, SE = 0.11, d = 0.03, p = .85. In both cases, including two-component trial types in the subtraction is not ideal, as there would likely already be some over-additivity in the higher-accuracy condition, over-estimating the contribution of the motor component. However, I believe that the more intuitive subtraction is trace minus view. In this case, I know for sure that the participant has been exposed to the exact same pictorial information in either case, as the picture that a given participant traced is identical to the picture that another viewed. In the case of comparing blind and imagine, it is less clear that the elaborative component required would be identical, and it is certainly not verifiable, as these are internal processes that likely differ at least in part across participants.

The findings from this experiment replicate those of Experiment 7, suggesting that the more multimodal components that are engaged at encoding, the better memory accuracy will be. This was manifest in the linear pattern, such that memory accuracy was best in the draw trial type, followed by trial types with two components (blind and trace), then trial types with one component (view and imagine), and lastly the baseline condition of write, with no components. Notably, the accuracy estimates for the trial types were relatively consistent across experiments as well (See Table 5, Figure 8). While the estimates are uniformly slightly lower, these existing differences can be explained by the removal of the images that were presented alongside the words in Experiment 7. It is not surprising that performance was slightly lower in these trial types in this experiment, as participants in this experiment had less exposure to the pictorial representation. The secondary findings, with respect to the component estimates, are less clear in this experiment, but still point toward the idea that drawing might lead to memory accuracy that is greater than the sum of its component parts. Together the findings from Experiment 7 and 8 are in line with the 'integrated components' model of the beneficial effects of drawing.

CHAPTER V: General Discussion

Across this dissertation, my results showed consistently and unequivocally that drawing pictures of presented words at encoding led to better memory than writing out words. This finding was consistent in free recall, regardless of whether the participants were instructed to add detail to their drawings and write repeatedly (Experiment 1A), or to add detail to their writing and draw repeatedly (Experiment 1B). It did not seem to matter whether encoding was incidental (Experiment 1A, 1B, 2A, 2B and 2C, 5A, 5B and 6), or intentional (Experiment 3, 4, 7 and 8), as the effect still occurred in either case. There was also relatively little influence of list length or shorter encoding times, as when the available encoding time was reduced to just four seconds (Experiments 3 and 4), the beneficial effects of drawing on free recall were still quite evident. It also did not appear to matter whether memory was tested using free recall (Experiments 1 through 4) or recognition (Experiments 5 through 8), as there was a reliable 'drawing effect' in all experiments. My results across all experiments demonstrate that drawing is a robust encoding manipulation that can, and does improve memory performance dramatically.

Given that something comparable to this effect was uncovered long ago (Paivio & Csapo, 1973), but since received little attention in the cognitive domain, I attempted to explore potential mechanisms which might be driving the benefit. One such mechanism is 'lazy writing' or loafing in the control condition. However, because drawing led to a substantial benefit over writing even when encoding time was reduced to 4 s, this does not seem plausible. Accordingly, in this dissertation I also explored whether distinctiveness, a mechanism which was been put forward as an explanation for many noteworthy encoding effects (e.g. Kinoshita, 1989; MacLeod et al., 2010; McDaniel & Bugg, 2008; Ozubko & MacLeod, 2010; Slamecka & Graf 1978) could account for the beneficial effects of drawing. The influential role of distinctiveness is evident because memory benefits are often absent when these encoding effects are manipulated between-participants (McDaniel & Bugg, 2008). However, my between-participants data suggest that while distinctiveness may contribute to the benefit, it cannot fully explain the drawing effect (Experiment 4). I also tested and ruled out the most obvious potential explanations as exhaustive explanatory mechanisms. That is, basic picture superiority, visual imagery and deep LoP failed to lead to performance that was comparable to the memory performance precipitated by drawing (Experiment 2).

Context and the Integrated Components Model

Having ruled out the foregoing candidate mechanisms, I proposed that drawing led to beneficial effects in memory as a result of the integration of three multimodal components – elaborative, motoric and pictorial – which can later be relied upon as contextual cues at retrieval. This 'integrated components' model is reminiscent of Paivio's (1971) dual coding theory, and Engelkamp's (2001) multimodal account of the enactment effect, which proposes that memory is improved by bringing online additional codes that would not otherwise be present during encoding. To reiterate, I define the components as follows (see Figure 9):

> *Elaborative:* The generative processes one must engage in upon viewing a word to bring to mind an internal representation of what that item looks like. This would require deep semantic processing of the item, in concert with visual imagery.
>
> *Motoric:* The manual motor program implemented when translating one's internal image from mind to paper. This motor movement must be directly associated with creating a representation of the meaning of the item, and thus is item-specific.
>
> *Pictorial*: Visual processing of the picture that one has created through drawing.

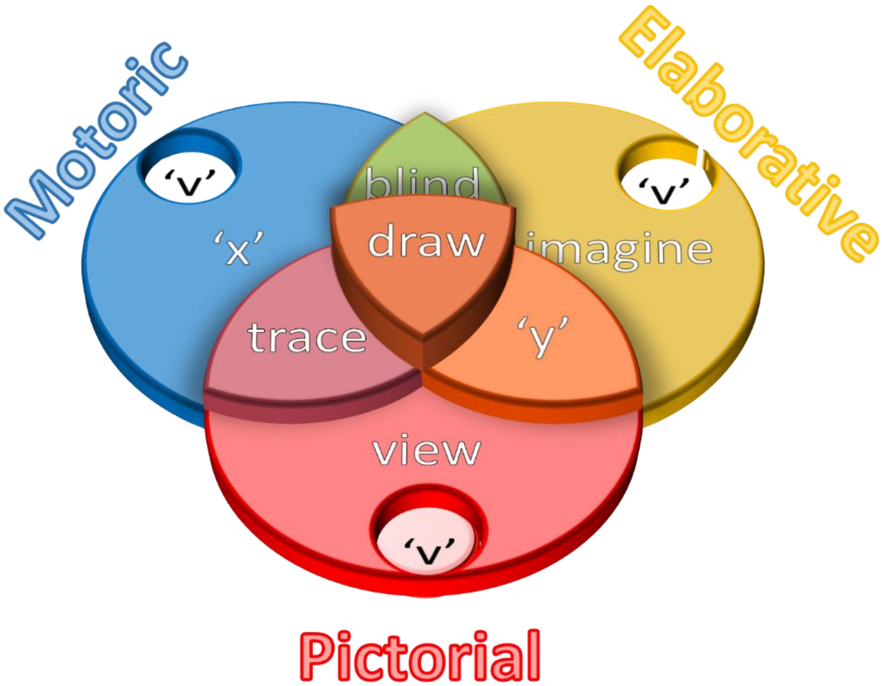

Figure 9. The integrated components model of the drawing effect, with Experiment 8's trial types overlaid. In this model, the beneficial effects of drawing, over and above basic verbal memory ('v') are driven by the integrated contributions of elaborative, motoric, and pictorial information. The 'draw' trial type lies at the intersection of all three components, as it engages all three. The 'trace' trial type, for example, lies at the intersection of the motoric and pictorial components, as it does not require elaborative thought. A purely motoric task ('x'), as well as a task that involves only elaborative and pictorial information ('y') without undermining the elaborative process is difficult, if not impossible to design.

I submit that each of these components serves as a cue, or source of contextual information from the encoding experience, which could later be used to aid retrieval. Of course, a prerequisite for this model would be that drawing allows one to retrieve contextual information more effectively than does writing. To test whether this was the case, the remember/know/new paradigm (Experiments 5A and 5B) was employed, as well as speeded recognition decisions (Experiment 6). Results from these experiments indicated that drawing words, relative to writing them out, led to a memory trace that was dependent primarily on recollection. This was evident in increased proportion of 'Remember' responses (Experiments 5A and 5B), as well as a blunted effect when access to the recollective process was offset (Experiment 6). These findings indicate that much like many other active encoding tasks (Engelkamp & Dehn, 1997; Gardiner, 1988; Gardiner et al., 1996; Lövdén et al., 2002; Manzi & Nigro, 2008; Rajaram, 1993), the retrieval benefit resultant from drawing is driven by recollection.

Having established this, I proposed and systematically tested the importance of each of the hypothesized components in driving the benefit that drawing affords memory (Experiments 7 and 8). This was accomplished through the development of trial types which did or did not require each of the 'integrated components'. The names of these trial types are overlaid on Figure 9. Results across the two experiments revealed a logical structure, with a few subtle exceptions. Namely, drawing, which engaged all three of the components yielded the best performance, followed by 'two-component' manipulations (blind-drawing and tracing), 'one-component' manipulations (viewing and imagining), and finally the 'zero-component' baseline (writing). In addition to these raw differences between the designed trial types, I also explored whether the contributions of each component could be derived using logical subtractions. It is important to note that there are well-documented issues with cognitive subtractions, including the problematic assumption that the components of a measured process are independent (See Gottsdanker & Shragg, 1985; Friston et al., 1996). While these arguments are typically levelled against experiments in which the duration of a given process is estimated through response time subtractions (e.g. Donders, 1868/1969), similar issues may occur here. However, in this case, I do not assume that the components are totally independent, and am simply attempting to test whether anything is gained in memory accuracy through the addition of a given component. While the component estimates I derived were necessarily noisy and the findings they yielded were not perfect, these experiments overall provide preliminary evidence for the importance of each of the components in driving the beneficial effects of memory.

Relation to Skilled Drawing

The 'integrated components' model of memory accuracy for drawn items proposed here has some fascinating and unintended parallels with the literature on the components or skills thought to subserve accurate drawing (Cohen & Bennett, 1997). In particular, the components I proposed here map onto the skilled drawing components with near-perfect correspondence. There are understandably subtle differences across fields, as the model of drawing skill was proposed to explain the act of drawing a representation of something that the artist was viewing and drawing concurrently. In contrast, in my 'draw' trial type, participants were creating their own generative drawing and not sketching a tangible object. Nevertheless, Cohen and Bennett (1997) proposed that drawing skill is driven by (1) perceiving the object one is drawing, (2) deciding which features of the object are important, (3) using motor coordination to translate these features to paper, and (4) viewing the outcome to gauge one's accuracy and actively correct.

As mentioned, drawings in my dissertation were not created to represent an object that was tangible or visible in the environment. Accordingly, my proposed elaborative component is rendered necessary in the absence of perception of the to-be-drawn object. Thus, to draw, participants were likely elaboratively using visual imagery to bring to mind the features they deemed important to draw, analogous to a combination of Cohen and Bennett's (1997) first and second ability. Furthermore, the theorized motoric and pictorial components in the current work are near identical to Cohen and Bennett's third and fourth abilities respectively. It is the active correction process mentioned here that I believe yields a deeply integrated representation, as drawing necessitates constant cross-talk between the internal representation, the requisite hand movements, and the developing drawing.

Interestingly, recent work has suggested that the most impactful of the foregoing skills thought to underlie effective drawing was the initial selection stage, where the picture to be drawn is decided upon (Kozbelt, Seidel, El Bassiouny, Mark & Owen, 2010; Ostrofsky, Kozbelt & Seidel, 2012). If this skill is most critical to effective *drawing*, it is possible that the generative component proposed in this dissertation is of particular importance to *memory for drawings*, and should serve as a target for future work in this line. It seems plausible that since drawing requires the internal generation of an appropriate picture to draw for a particular item, it requires honing in on the 'sailboat-ness' of a sailboat for instance, thereby strengthening the representation of that item. Perhaps then, the generative component is of particular importance in improving memory. There is some indirect evidence reflecting the necessity of the generative component, as the blind-draw trial type led to numerically superior performance relative to trace, and imagine to superior performance than view, though both non-significantly.

Drawing Inferences Beyond Memory

The impact of drawing goes beyond simply being an effective mnemonic as well. There has been some suggestion that drawing, both through years of training (Carson & Allard, 2013; Carson, Quehl, Aliev, & Danckert, 2013; Chamberlain, McManus, Riley, Rankin & Brunswick, 2013; Perdreau & Cavanagh, 2013; 2014; 2015), and in short-term repetitive tasks (Fan, Yamins, & Turk-Browne, in press), might measurably change the way that visually presented items are represented, a finding which extends to enhancing scientific thinking more generally (Fan, 2015). In one example, participants viewed possible or impossible figures, with a gaze-contingent display revealing only a circular section of the figure that was centered on the fovea and varied in radius. Researchers discovered that better artists performed the task successfully with smaller viewable windows than their less artistically-inclined counterparts (Perdreau & Cavanagh, 2013b). Interestingly, this does not seem to have anything to do with broader perceptual ability, as artists performed no better than novices on this task when the entire figure was displayed (Perdreau & Cavanagh, 2013b). Furthermore, artists are comparable to novices in shape constancy tasks (McManus, Loo, Chamberlain, Riley & Brunswick, 2011) and are equally susceptible to distracting visual information in perceptual tasks (Perdreau & Cavanagh, 2013a).

The foregoing suggests that the ability to draw is associated with the ability to maintain a coherent internal representation, in line with the proposition that the generative component might be most important. The relation between drawing and the strength of internal representations has also been corroborated by neuroimaging evidence, suggesting that the volume of structures associated with visual imagery is associated with drawing ability (Chamberlain, McManus, Brunswick, Rankin, Riley & Kanai, 2014). Interestingly, this same work also found associations between drawing ability/experience, and changes in structures associated with fine motor control and procedural memory. It is possible then, that drawing generally increases the coherence of one's internal representation, and also is driven by motor control. These neuroimaging findings align nicely with my proposed integrated components.

There is also evidence to suggest that drawing might serve to differentiate between items that are highly visually similar (Fan, Yamins & Turk-Browne, in press). In this work, a convolutional neural network model, designed to reflect veridical characteristics of the ventral visual stream, was trained to perform an object recognition task at near human-level performance. Participants were then brought in to complete a drawing training task, where they repeatedly practiced drawing sets of objects that were frequently confused with one another by the network, and thus visually similar in some capacity. The task, in essence, involved playing Pictionary with the computer. Participants repeatedly drew, for example, a cactus and a lobster. Then the neural network 'guessed' what they were trying to draw. Over time, the repeated drawings became more distinguishable from one another, and non-repeated drawings that were also highly similar to the repeated drawings (e.g. crab, palm tree) became *more* similar to one another. While more work will be needed to determine whether these changes are instantiated neutrally, and to determine the longevity of this modification of representational space, it seems clear that drawing, especially of similar items, can serve to differentiate them and make them more distinguishable from one another. It is possible that in the current work, drawing accomplished just this, differentiating similar concepts from one another and thus reducing inter-item interference at retrieval.

Limitations and Future Directions

While I did demonstrate a between-participants drawing effect in free recall, which provides evidence against distinctiveness as an explanatory mechanism (e.g. McDaniel & Bugg, 2008), the effect *was* quite a bit smaller in this experiment than in its within-participants comparator. This suggests that distinctiveness is contributing, at least in part, to the benefit. Moreover, despite the fact that within-experiment distinctiveness could not have been the primary driving force of the drawing effect, it is possible that distinctiveness (Dodson & Schacter, 2001) may also be influential as a more general heuristic. That is, memory may have improved because drawing is a rarer task than writing across one's lifetime, despite the fact that drawing and writing events were equally common in most of my experiments and despite the fact that participants encoded exclusively drawn items in one experiment.

To clarify, Icht, Mama and Algom (2014) have differentiated between two 'species'' of distinctiveness that could theoretically impact memory. The first – 'relative distinctiveness' – is what researchers are most frequently referring to when discussing distinctiveness as an explanatory mechanism in mixed-lists. This is the idea that produced (or enacted, or drawn) items are more distinct than their respective baselines because they contain some extra information that can be used at retrieval. This was largely ruled out in the drawing effect due to my demonstration of a between-participants effect. The second – 'statistical distinctiveness' - refers to the idea that particular items or tasks may stand out because they simply occur less frequently than others. While the authors refer to this in the context of trial distribution *within* an experiment (Icht et al., 2014), it could also apply across one's lifespan. That is, it could be the case that drawn items are distinct from written, even when presented in pure lists, simply because many people do not often draw. Future work could explore this possibility by including a drawing 'wear-out' phase, where participants draw ad nauseam before and after the to-be-remembered information is presented in the encoding phase, thus making the drawing manipulation less statistically distinct, and therefore less diagnostic. Alternatively, sketch artists could be tested to determine whether drawing, despite being relatively commonplace for them, still improves memory.

One unexplored mechanistic explanation comes from item-order accounts (McDaniel & Bugg, 2008), which hold that several common encoding manipulations affecting free recall performance are due to shifts in whether item or relational information is enhanced. This account has been invoked to explain production effects (e.g. Jonkeret al., 2014; Jonker & MacLeod, 2015), generation effects (e.g. Serra & Nairne, 1991), enactment effects (Engelkamp & Dehn, 2000), and several other memory phenomena (McDaniel & Bugg, 2008). To clarify, various rich encoding tasks tend to greatly enhance item information, at the expense of relational information. Thus, when compared in mixed lists, the more distinctive condition benefits from enhanced item processing, while the baseline condition does not. The enhanced item encoding actively disrupts relational information, or information about the order in which items were presented. Now, when these two same conditions are compared between-participants, the more distinctive condition still benefits from item information as it did in the within-participants implementation, while the baseline does not. However, the baseline condition now benefits from relational information about order, because these temporal contingencies have not been disrupted by stronger items and their enhanced item information. In future work, whether or not the item-order account can be extended to the drawing effect should be tested.

Our data (Experiments 7 and 8) and research on the requirements for skilled drawing (Kozbelt et al., 2010; Ostrofsky et al., 2012) have highlighted the potential importance of generative information (or the elaborative component) in producing the observed memorial benefit (or accurate drawing). While I have demonstrated that each component likely contributes to the benefit of drawing, there is still work to be done to adjudicate whether any one component is more important than the others. For practical reasons, it will also be important to determine how many (or how few) of the component parts of the act of drawing is enough to boost memory. To explore this, future work could stage a situation wherein participants must instruct another person in what to draw rather than creating a drawing themselves. This would create conditions wherein the person is not implementing the motor program, nor observing the final pictorial product, but still employs the elaborative process of deciding on something to draw. This alone could be enough to produce an appreciable increase in memory performance. Another approach might be to give participants a preparatory time period after viewing a word, but prior to creating a drawing. If, in some cases, participants were not allowed to actually create the drawing, this would also isolate the elaborative process, and provide an estimate as to the benefit it affords.

Lastly, while I have demonstrated that performance scales up as more components are added, the ‘integrated’ part of the integrated components model remains untested at this time. It has been suggested that for contextual information to be bound in memory, a necessary precondition is that the distinct features of an episode are encoded as a whole during the initial experience (Uncapher, Otten & Rugg, 2006). This is undoubtedly the case with the three disparate sources of contextual information I propose to be playing a role in the benefit of drawing. That is, elaboration, hand movement, and pictorial processing are all engaged in one act, and thus are very likely to be bound within one experience. With this, it is possible that the integration aspect of the model is an obvious byproduct of the task. Interestingly, other work has shown that memory for one contextual feature of an encoding event is highly correlated with memory for a totally distinct contextual feature, suggesting that context is often bound into one seamless representation, or at the very least, that recalling one is likely to precipitate recall of the other (Marsh, Hicks & Cook, 2004; Starn & Hicks, 2005). With this, future work could test whether cuing a participant using information from one component (e.g. using a 'connect-the-dots' task to reinstantiate the motor context), leads to enhanced memory for that item, and/or its other associated contextual information.

Conclusions

The experiments reported in this dissertation explored the efficacy of drawing as an encoding technique which might profitably be used to increase the success of later retrieval. Drawing was compared directly to writing, and my findings indicated that drawing improves memory in free recall as well as recognition. Moreover, this effect was robust to subtle changes in emphasis made through instructions to add detail to one trial type or the other, and was superior to other commonly touted encoding manipulations. Unlike many other seemingly stable encoding enhancers, drawing was also robust to the change from within-participants to between-participants designs. This latter finding rules out relative distinctiveness as an explanatory mechanism that can entirely account for the effect. Therefore, drawing is in rare company as a relatively easily implemented memory improvement task that is strikingly consistent across experimental designs.

My findings also show that drawing improves memory primarily through increasing recollection, while familiarity stays relatively stable. In other words, drawing leads one to be more likely to remember the context and specifics of the encoding event, instead of just having a general sense of knowing that one has seen an item before. This was evident in the finding that participants were more likely to indicate that they remembered contextual details from the encoding experience following drawing, relative to writing. I found converging evidence for the critical role of recollection by limiting access to this process through enforcing strict time constraints on retrieval decisions. When time pressure was included, the drawing effect was substantially reduced, likely due to reduced access to contextual information.

Lastly, I sought to break the act of drawing down into its component parts, and demonstrated that performance increased linearly as more multimodal components were added, building up to drawing. In my integrated components model, I proposed that drawing required an elaborative component to decide what to draw, a motoric component to translate that internal representation to the page, and a pictorial component from viewing the drawing that has been created. Data showed that drawing led to the greatest success in later retrieval, while trial types that engaged only two of these components were inferior to drawing, and so on. The experiments reported herein provide preliminary support for the beneficial effects of integrated components, and highlight the importance of including or emphasizing information from multiple distinct codes, modalities or sources at encoding, as opposed to simply manipulating memory strength more broadly. This dissertation also clearly highlights that drawing is a consistently strong encoding strategy, whose benefits appear to be strikingly task-invariant.

Footnote

[1] The analysis detailed in Erlebacher (1977) requires that sample sizes in all three groups be equal. The within subjects design had 28 participants, and the between subjects groups had 23 and 24 participants respectively. In order to complete the analysis, 23 participants were randomly selected from the larger groups, and the analysis was conducted on this truncated data set.

[2] It is important to note that the inclusion of these participants does not alter the pattern of data in any substantive way. All significance tests follow precisely the same pattern, whether they are included in the analysis or not.

References

Addante, R. J., Ranganath, C., & Yonelinas, A. P. (2012). Examining ERP correlates of recognition memory: Evidence of accurate source recognition without recollection. *Neuroimage*, *62*, 439-450.

Agerbeck, B. (2012). *The graphic facilitator's guide: How to use your listening, thinking & drawing skills to make meaning*. Loosetooth.com Library.

Ainsworth, S., & Loizou, A. T. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science, 27*, 669-681.

Alesandrini, K. (1981). Pictorial-verbal and analytic-holistic learning strategies in science learning. *Journal of Educational Psychology*, *73*, 358–368.

Arnold, T. C., & Dwyer, F. M. (1975). Realism in visualized instruction. *Perceptual and Motor Skills*, *40*, 369-370.

Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy of word recognition. *Attention and Performance IV*, 583-612.

Balota, D. A., & Black, S. (1997). Semantic satiation in healthy young and older adults. *Memory & Cognition*, *25*, 190-202.

Begg, I., & Roe, H. (1988). On the inhibition of reading by generating. *Canadian Journal of Psychology*, *42*, 325-336.

Benjamin, A. S., & Craik, F. I. (2001). Parallel effects of aging and time pressure on memory for source: Evidence from the spacing effect. *Memory & Cognition*, *29,* 691-697.

Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007). The generation effect: A meta-analytic review. *Memory & Cognition*, *35*, 201-210.

Besson, G., Ceccaldi, M., Didic, M., & Barbeau, E. J. (2012). The speed of visual recognition memory. *Visual Cognition*, *20*, 1131-1152.

Besson, G., Ceccaldi, M., Tramoni, E., Felician, O., Didic, M., & Barbeau, E. J. (2015). Fast, but not slow, familiarity is preserved in patients with amnestic mild cognitive impairment. *Cortex*, *65*, 36-49.

Bodner, G. E., & MacLeod, C. M. (2016). The benefits of studying by production... and of studying production: Introduction to the special issue on the production effect in memory. *Canadian Journal of Experimental Psychology, 70*, 89–92.

Bodner, G. E., & Taikh, A. (2012). Reassessing the basis of the production effect in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1711-1719.

Bodner, G. E., Taikh, A., & Fawcett, J. M. (2014). Assessing the costs and benefits of production in recognition. *Psychonomic Bulletin & Review*, *21*, 149-154.

Boldini, A., Russo, R., & Avons, S. E. (2004). One process is not enough! A speed-accuracy tradeoff study of recognition memory. *Psychonomic Bulletin & Review*, *11*, 353-361.

Boldini, A., Russo, R., Punia, S., & Avons, S. E. (2007). Reversing the picture superiority effect: A speed—accuracy trade-off study of recognition memory. *Memory & Cognition*, *35*, 113-123.

Bowles, B., Crupi, C., Mirsattari, S. M., Pigott, S. E., Parrent, A. G., Pruessner, J. C., Yonelinas, A. P., & Köhler, S. (2007). Impaired familiarity with preserved recollection after anterior temporal-lobe resection that spares the hippocampus. *Proceedings of the National Academy of Sciences*, *104*, 16382-16387.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, *105*), 14325-14329.

Brown, S. (2014). *The doodle revolution: Unlock the power to think differently*. New York, NY: Penguin Group.

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: what are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience*, *2*, 51-61.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, *14*, 5-26.

Carson, L. C., & Allard, F. (2013). Angle-drawing accuracy as an objective performance-based measure of drawing expertise. *Psychology of Aesthetics, Creativity, and the Arts*, *7*, 119-129.

Carson, L., Quehl, N., Aliev, I., & Danckert, J. (2014). Angle-based drawing accuracy analysis and mental models of three-dimensional space. *Art & Perception*, *2*, 183-212.

Chamberlain, R., McManus, I. C., Brunswick, N., Rankin, Q., Riley, H., & Kanai, R. (2014). Drawing on the right side of the brain: a voxel-based morphometry analysis of observational drawing. *NeuroImage, 96*, 167-173.

Chamberlain, R., McManus, I. C., Riley, H., Rankin, Q., & Brunswick, N. (2013). Local processing enhancements associated with superior observational drawing are due to enhanced perceptual functioning, not weak central coherence. *The Quarterly Journal of Experimental Psychology, 66*, 1448-1466.

Charles, S. T., Mather, M., & Carstensen, L. L. (2003). Aging and emotional memory: The forgettable nature of negative images for older adults. *Journal of Experimental Psychology: General*, *132,* 310-324.

Cohen, D. J., & Bennett, S. (1997). Why can't most people draw what they see? *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 609.

Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. *Journal of Memory and Language*, *26*, 341-361.

Cooper, E. H., & Pantle, A. J. (1967). The total-time hypothesis in verbal learning. *Psychological Bulletin*, *68*, 221-234.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, *11*, 671-684.

Curran, T., & Doyle, J. (2011). Picture superiority doubly dissociates the ERP correlates of recollection and familiarity. *Journal of Cognitive Neuroscience*, *23*, 1247-1262.

Dean, R., & Kulhavy, R. (1981). Influence of spatial organization in prose learning. *Journal of Educational Psychology*, *73,* 57–64.

Defeyter, M. A., Russo, R., & McPartlin, P. L. (2009). The picture superiority effect in recognition memory: A developmental study using the response signal procedure. *Cognitive Development*, *24*, 265-273.

Dewhurst, S. A., & Conway, M. A. (1994). Pictures, images, and recollective experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20,* 1088-1098.

Dodson, C. S., & Schacter, D. L. (2001). "If I had said it I would have remembered it: Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin & Review*, *8*, 155-161.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica*, *30*, 412-431.

Duzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H. J., & Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Sciences, 94*, 5973-5978.

Ebbinghaus, H. (1885). *Über das Gedächtnis. Untersuchungen zur Experimentellen Psychologie [About memory. Investgations in experimental psychology*]. Leipzig: Duncker & Humblot (Reprint 1966, Amsterdam: E.J. Bonset).

Elliott, L. (1973). Imagery versus repetition encoding in short-and long-term memory. *Journal of Experimental Psychology*, *100*, 270-276.

Engelkamp J. (2001). Action memory: A system-oriented approach. In H. D. Zimmer, R. L. Cohen, M. J. Guynn, J. Engelkamp, R. Kormi-Nouri, & M. A. Foley (Eds.) *Memory for action: A distinct form of episodic memory?* 46–96. New York, NY: Oxford University.

Engelkamp, J., & Dehn, D. M. (2000). Item and order information in subject-performed tasks and experimenter-performed tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 671-682.

Engelkamp, J., & Zimmer, H. D. (1997). Sensory factors in memory for subject-performed tasks. *Acta Psychologica*, *96*, 43-60.

Engelkamp, J., Zimmer, H. D., & Biegelmann, U. E. (1993). Bizarreness effects in verbal tasks and subject-performed tasks. *European Journal of Cognitive Psychology*, *5*, 393-415.

Erlebacher, A. (1977). Design and analysis of experiments contrasting the within-and between-subjects manipulation of the independent variable. *Psychological Bulletin*, *84*, 212-219.

Fan, J. E. (2015). Drawing to learn: How producing graphical representations enhances scientific thinking. *Translational Issues in Psychological Science*, *1*, 170.

Fan, J., Yamins, D., & Turk-Browne, N. (in press). Common object representations for visual recognition and production. In *Proceedings of the 37th Annual Conference of the Cognitive Science Society*.

Forrin, N. D., Jonker, T. R., & MacLeod, C. M. (2014). Production improves memory equivalently following elaborative vs non-elaborative processing. *Memory*, *22*, 470-480.

Forrin, N. D., MacLeod, C. M., & Ozubko, J. D. (2012). Widening the boundaries of the production effect. *Memory & Cognition*, *40*, 1046-1055.

Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S. J., & Dolan, R. J. (1996). The trouble with cognitive subtraction. *Neuroimage*, *4*, 97-104.

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, *16*, 309-313.

Gardiner, J. M. (2001). Episodic memory and autonoetic consciousness: A first–person approach. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *356*, 1351-1361.

Gardiner, J. M., Java, R. I., & Richardson-Klavehn, A. (1996). How level of processing really influences awareness in recognition memory. *Canadian Journal of Experimental Psychology*, *50*, 114-122.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (1999). Response deadline and subjective awareness in recognition memory. *Consciousness and Cognition*, *8*, 484-496.

Gardiner, J. M., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition*, *18*, 579-583.

Gathercole, S. E., & Conway, M. A. (1988). Exploring long-term modality effects: Vocalization leads to best retention. *Memory & Cognition*, *16*, 110-119.

Gottsdanker, R., & Shragg, G. P. (1985). Verification of Donders' subtraction method. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 765-776.

Graf, P., & Mandler, G. (1984). Activation makes words more accessible, but not necessarily more retrievable. *Journal of Verbal Learning and Verbal Behavior*, *23*, 553-568.

Greene, T. (1989). Children's understanding of class inclusion hierarchies: The relationship between external representation and task performance. *Journal of Experimental Child Psychology*, *48*, 62-89

Guttentag, R. E., & Hunt, R. R. (1988). Adult age differences in memory for imagined and performed actions. *Journal of Gerontology*, *43*, 107-108.

Hall, V. C., Bailey, J, & Tillman, C. (1997). Can student-generated illustrations be worth ten thousand words? *Journal of Educational Psychology, 89,* 677-681.

Düzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H. J., & Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Sciences*, *94*, 5973-5978.

Hockley, W. E. (2008). The effects of environmental context on recognition memory and claims of remembering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1412-1429.

Hopkins, R. H., & Edwards, R. E. (1972). Pronunciation effects in recognition memory. *Journal of Verbal Learning and Verbal Behavior*, *11*, 534-537.

Huff, M. J., Bodner, G. E., & Fawcett, J. M. (2015). Effects of distinctive encoding on correct and false memory: A meta-analytic review of costs and benefits and their origins in the DRM paradigm. *Psychonomic Bulletin & Review*, *22*, 349-365.

Hunt, R. R., & Elliot, J. M. (1980). The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. *Journal of Experimental Psychology: General*, *109*, 49-74.

Hunt, R. R., & Mitchell, D. B. (1982). Independent effects of semantic and nonsemantic distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(1), 81-87.

Icht, M., Mama, Y., & Algom, D. (2014). The production effect in memory: Multiple species of distinctiveness. *Frontiers in psychology*, 1-7.

Jacoby, L., & Craik, F. (1979). Effects of elaboration of processing at encoding and retrieval: trace distinctiveness and recovery of initial context. In L. Cermak and F. Craiks (Eds.) *Levels of processing in human memory*, 1-21. Hillsdale, NJ: Lawrence Erlbaum Associates.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*, 306-340.

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. M. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. D. Cohen, & J. W. Schooler (Eds.), *Scientific approaches to consciousness*, 13-47. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Jones, G. V. (1987). Independence and exclusivity among psychological processes: Implications for the structure of recall. *Psychological Review*, *94*, 229-235.

Jones, A. C., & Pyc, M. A. (2014). The production effect: Costs and benefits in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 300-305.

Jonker, T. R., & MacLeod, C. M. (2015). Disruption of relational processing underlies poor memory for order. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 831.

Jonker, T. R., Levene, M., & MacLeod, C. M. (2014). Testing the item-order account of design effects using the production effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 441-448.

Jonker, T. R., & MacLeod, C. M. (2015). Disruption of relational processing underlies poor memory for order. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 831-840.

Joordens, S., & Merikle, P. M. (1993). Independence or redundancy? Two models of conscious and unconscious influences. *Journal of Experimental Psychology: General*, *122*, 462-467.

Juola, J. F., Fischler, I., Wood, C. T., & Atkinson, R. C. (1971). Recognition time for information stored in long-term memory. *Perception & Psychophysics*, *10*, 8-14.

Karpicke, J. D., & Zaromb, F. M. (2010). Retrieval mode distinguishes the testing effect from the generation effect. *Journal of Memory and Language*, *62*, 227-239.

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*, 1169-1180.

Kieras, D. (1978). Beyond pictures and words: Alternative information-processing models for imagery effect in verbal memory. *Psychological Bulletin*, *85*, 532-554.

Kinjo, H., & Snodgrass, J. G. (2000). Does the generation effect occur for pictures? *The American Journal of Psychology*, *113*, 95-121.

Kinoshita, S. (1989). Generation enhances semantic processing? The role of distinctiveness in the generation effect. *Memory & Cognition*, *17*, 563-571.

Kleinsmith, L. J., & Kaplan, S. (1964). Interaction of arousal and recall interval in nonsense syllable paired-associate learning. *Journal of Experimental Psychology*, *67*, 124-126.

Knopf, M. (1991). Having shaved a kiwi fruit: Memory of unfamiliar subject-performed actions. *Psychological Research, 53*, 203-211.

Knopf, M., & Neidhardt, E. (1989). Aging and memory for action events: The role of familiarity. *Developmental Psychology, 25*, 780-786.

Knowlton, B. J., & Squire, L. R. (1995). Remembering and knowing: two different expressions of declarative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 699-710.

Koen, J. D., & Yonelinas, A. P. (2010). Memory variability is due to the contribution of recollection and familiarity, not to encoding variability. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1536-1542.

Kormi-Nouri, R., & Nilsson, L. G. (2001). The Motor Component is Not Crucial! In H. D. Zimmer, R. L. Cohen, M. J. Guynn, J. Engelkamp, R. Kormi-Nouri, & M. A. Foley (Eds.) *Memory for action: A distinct form of episodic memory?* 97-111. New York, NY: Oxford University.

Kozbelt, A., Seidel, A., El Bassiouny, A., Mark, Y., & Owen, D. R. (2010). Visual selection contributes to artists' advantages in realistic drawing. *Psychology of Aesthetics, Creativity, and the Arts, 4*, 93.

Kulhavy, R., Lee, J., & Caterino, L. (1985). Conjoint retention of maps and related discourse. *Contemporary Educational Psychology*, *10*, 28-37.

Lesgold, A. M., De Good, H. & Levin, J. R. (1977). Pictures and young children's prose learning: a supplementary report. *Journal of Reading Behavior,* 9, 353-360.

Lesgold, A. M., Levin, J. R., Shimron, J, & Guttman, J. (1975). Pictures and young children's learning from oral prose. *Journal of Educational Psychology*, 67, 636-642.

Leutner, D., Leopold, C., & Sumfleth, E. (2009). Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content. *Computers in Human Behavior, 25,* 284-289.

Lövdén, M., Rönnlund, M., & Nilsson, L. G. (2002). Remembering and knowing in adulthood: effects of enacted encoding and relations to processing speed. *Aging, Neuropsychology, and Cognition*, *9*, 184-200.

Lupiani, D. A. (1977). The facilitative effects of imagery instructions and stimulus characteristics on immediate and long term free recall and clustering. *The Journal of General Psychology, 97*, 73-87.

Luo, L., Hendriks, T., & Craik, F. I. (2007). Age differences in recollection: Three patterns of enhanced encoding. *Psychology and Aging*, *22*, 269-280.

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 671-685.

Maisto, A. A., & Queen, D. E. (1992). Memory for pictorial information and the picture superiority effect. *Educational Gerontology: An International Quarterly*, *18*, 213-223.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*, 252-271.

Manzi, A., & Nigro, G. (2008). Long-term memory for performed and observed actions: Retrieval awareness and source monitoring. *Memory, 16*, 595-603.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British journal of Psychology, 64*, 17-24.

Marsh, R., Hicks, J., & Cook, G. (2004). Focused attention on one contextual attribute does not reduce source memory for a different attribute. *Memory, 12*, 183-192.

Marschark, M., Richman, C. L., Yuille, J. C., & Hunt, R. R. (1987). The role of imagery in memory: On shared and distinctive information. *Psychological Bulletin, 102*, 28-41.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*, 43-52.

Mazzoni, D., & Dannenberg, R. (2000). *Audacity [software]*. Pittsburg, PA: Carnegie Mellon University.

McBride, D. M., & Dosher, B. A. (2002). A comparison of conscious and automatic memory processes for picture and word stimuli: A process dissociation analysis. *Consciousness and Cognition, 11*, 423-460.

McDaniel, M. A., & Bugg, J. M. (2008). Instability in memory phenomena: A common puzzle and a unifying explanation. *Psychonomic Bulletin & Review*, *15*, 237-255.

McDaniel, M. A., & Einstein, G. O. (1986). Bizarre imagery as an effective memory aid: The importance of distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 54-65.

McManus, I. C., Loo, P. W., Chamberlain, R., Riley, H., & Brunswick, N. (2011). Does shape constancy relate to drawing ability? Two failures to replicate. *Empirical Studies of the Arts*, *29*(2), 191-208.

Meade, M. E., Wammes, J. D., & Fernandes, M. A. (submitted). Drawing as an encoding tool: Memorial benefits in younger and older adults. *Experimental Aging Research.*

Mecklinger, A. (2000). Interfacing mind and brain: A neurocognitive model of recognition memory. *Psychophysiology*, *37*, 565-582.

Meng, Y., Ye, X., & Gonsalves, B. D. (2014). Neural processing of recollection, familiarity and priming at encoding: Evidence from a forced-choice recognition paradigm. *Brain Research*, *1585*, 72-82.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of verbal learning and verbal behavior*, *16*, 519-533.

Mulligan, N. (2002). The emergence of item-specific encoding effects in between-subjects designs: perceptual interference and multiple recall tests. *Psychonomic Bulletin & Review*, *9*, 375–382.

Mulligan, N., & Hirshman, E. (1995). Speed-accuracy trade-offs and the dual process model of recognition memory. *Journal of Memory and Language, 34*, 1-18.

Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 702-709.

Nairne, J. S., Thompson, S. R., & Pandeirada, J. N. (2007). Adaptive memory: survival processing enhances retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 263.

Nelson, D. L., Schreiber, T. A., & McEvoy, C. L. (1992). Processing implicit and explicit representations. *Psychological Review*, *99*, 322-348.

Ostrofsky, J., Kozbelt, A., & Seidel, A. (2012). Perceptual constancies and visual selection as predictors of realistic drawing skill. *Psychology of Aesthetics, Creativity, and the Arts*, *6*, 124-136.

Ozubko, J. D., & MacLeod, C. M. (2010). The production effect in memory: Evidence that distinctiveness underlies the benefit. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1543-1547.

Ozubko, J. D., Major, J., & MacLeod, C. M. (2014). Remembered study mode: Support for the distinctiveness account of the production effect. *Memory*, *22*, 509-524.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, *76*, 241.

Paivio, A (1971). Imagery and verbal processes. New York: Holt, Rinehart, and Winston.

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding?. *Cognitive Psychology*, *5*, 176-206.

Paivio, A., & Foth, D. (1970). Imaginal and verbal mediators and noun concreteness in paired-associate learning: The elusive interaction. *Journal of Verbal Learning and Verbal Behavior*, *9*, 384-390.

Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science, 11*, 137-138.

Perdreau, F., & Cavanagh, P. (2013a). Is artists' perception more veridical?. *Frontiers in Neuroscience, 7*, 1-11.

Perdreau, F., & Cavanagh, P. (2013b). The artist's advantage: Better integration of object information across eye movements. *i-Perception, 4*, 380-395.

Perdreau, F., & Cavanagh, P. (2014). Drawing skill is related to the efficiency of encoding object structure. *i-Perception*, *5*(2), 101-119.

Perdreau, F., & Cavanagh, P. (2015). Drawing experts have better visual memory while drawing. *Journal of vision, 15*, 1-10.

Perfect, T. J. (1996). Does context discriminate recollection from familiarity in recognition memory? *The Quarterly Journal of Experimental Psychology*, *49A*, 797-813.

Peterson, D. J., & Mulligan, N. W. (2013). The negative testing effect and multifactor account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1287-1293.

Peynircioğlu, Z. F. (1989). The generation effect with pictures and nonsense figures. *Acta Psychologica, 70*, 153-160.

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition*, *21*, 89-102.

Rajaram, S. (1996). Perceptual effects on remembering: Recollective processes in picture recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 365-377.

Rasco, R., Tennyson, R., & Boutwell, R. (1975). Imagery instructions and drawings in learning prose. *Journal of Educational Psychology, 67*, 188–192.

Reid, D. J. (1984). The picture superiority effect and biological education. *Journal of Biological Education*, *18*, 29-36.

Rigney, J. W., & Lutz, K. A. (1976). Effect of graphic analogies of concepts in chemistry on learning and attitudes. *Journal of Educational Psychology,68,* 305-311.

Roam, D. (2013). *The back of the napkin: Solving problems and selling ideas with pictures*. New York, NY: Penguin Group.

Roediger, H. L. (1974). Inhibiting effects of recall. *Memory & Cognition*, *2*, 261-269.

Rotello, C. M., & Zeng, M. (2008). Analysis of RT distributions in the remember—know paradigm. *Psychonomic Bulletin & Review*, *15*, 825-832.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*, 251-257.

Sauvage, M. M., Beer, Z., & Eichenbaum, H. (2010). Recognition memory: Adding a response deadline eliminates recollection but spares familiarity. *Learning & Memory*, *17*, 104-108.

Scaife, M., & Rogers, Y. (1996). External cognition: how do graphical representations work?. *International Journal of Human-Computer Studies*, *45*, 185-213.

Schwamborn, A., Mayer, R. E., Thillmann, H., Leopold, C., & Leutner, D. (2010). Drawing as a generative activity and drawing as a prognostic activity. *Journal of Educational Psychology, 102*, 872-879.

Senkfor, A. J. (2008). Memory for pantomimed actions versus actions with real objects. *Cortex*, *44*, 820-833.

Senkfor, A. J., & Van Petten, C. (1998). Who said what? An event-related potential investigation of source and item memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1005-1025.

Serra, M., & Nairne, J. S. (1993). Design controversies and the generation effect: Support for an item-order hypothesis. *Memory & Cognition*, *21*, 34-40.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, *14*, 47-69.

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

Slamecka, N. J., & Katsaiti, L. T. (1987). The generation effect as an artifact of selective displaced rehearsal. *Journal of Memory and Language, 26*, 589-607.

Snodgrass, J. G., & McClure, P. (1975). Storage and retrieval properties of dual codes for pictures and words in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, *1*, 521-529.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174-215.

Snowman, J, & Cunningham, D. J. (1975). A comparison of pictorial and written adjunct aids in learning from text. *Journal of Educational Psychology, 67*, 307-311.

Staresina, B. P., Fell, J., Dunn, J. C., Axmacher, N., & Henson, R. N. (2013). Using state-trace analysis to dissociate the functions of the human hippocampus and perirhinal cortex in recognition memory. *Proceedings of the National Academy of Sciences, 110*, 3119-3124.

Starns, J. J., & Ratcliff, R. (2008). Two dimensions are not better than one: STREAK and the univariate signal detection model of remember/know performance. *Journal of Memory and Language, 59*, 169-182.

Stein, B. S., Littlefield, J., Bransford, J. D., & Persampieri, M. (1984). Elaboration and knowledge acquisition. *Memory & Cognition*, *12*, 522-529.

Toth, J. P. (1996). Conceptual automaticity in recognition memory: Levels-of-processing effects on familiarity. *Canadian Journal of Experimental Psychology*, *50*, 123-138.

Tulving, E. (1983). *Elements of episodic memory*. Oxford, UK: Clarendon.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*, 1-12.

Uncapher, M. R., Otten, L. J., & Rugg, M. D. (2006). Episodic encoding is more than the sum of its parts: an fMRI investigation of multifeatural contextual encoding. *Neuron*, *52*, 547-556.

Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology*, *93*, 129-140.

Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review*, *17*, 285-325.

Waddill, P. J., & McDaniel, M. A. (1998). Distinctiveness effects in recall. *Memory & Cognition*, *26*, 108-120.

Wagner, A. D., Gabrieli, J. D., & Verfaellie, M. (1997). Dissociations between familiarity processes in explicit recognition and implicit perceptual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 305-323.

Wais, P. E., Mickes, L., & Wixted, J. T. (2008). Remember/know judgments probe degrees of recollection. *Journal of Cognitive Neuroscience*, *20,* 400-405.

Walsh, D. A., & Jenkins, J. J. (1973). Effects of orienting tasks on free recall in incidental learning: "Difficulty," "effort," and "process" explanations. *Journal of Verbal Learning and Verbal Behavior*, *12*, 481-488.

Wammes, J. D., Meade, M. E., & Fernandes, M. A. (2016). The drawing effect: Evidence for reliable and robust memory benefits in free recall. *The Quarterly Journal of Experimental Psychology, 69*, 1752-1776.

Wammes, J. D., Meade, M. E., & Fernandes, M. A. (in press). Creating a recollection-based memory through drawing. *Journal of Experimental Psychology: Learning, Memory and Cognition*.

Weldon, M. S., & Coyote, K. C. (1996). Failure to find the picture superiority effect in implicit conceptual memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 670-686.

Wheeler, M. E., & Buckner, R. L. (2003). Functional dissociation among components of remembering: control, perceived oldness, and content. *Journal of Neuroscience*, *23*, 3869-3880.

Wheeler, M. E., & Buckner, R. L. (2004). Functional-anatomic correlates of remembering and knowing. *Neuroimage*, *21*, 1337-1349.

Whitehouse, A.J.O., Mayberry, M.T., & Durkin, K. (2006). Inner speech impairments in autism. *Journal of Child Psychology and Psychiatry, 47*, 857-866.

Winnick, W. A., & Brody, N. (1984). Auditory and visual imagery in free recall. *The Journal of psychology*, *118*, 17-29.

Wilding, E. L., & Ranganath, C. (2011). Electrophysiological correlates of episodic memory processes. In S.J. Luck, & E. S. Kappenman (Eds.), *The Oxford handbook of ERP components*, 373-396. New York, NY: Oxford University Press.

Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research, 1100*, 125-135.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1341-1354.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441-517.

Yonelinas, A. P., & Jacoby, L. L. (1994). Dissociations of processes in recognition memory: effects of interference and of response speed. *Canadian Journal of Experimental Psychology*, *48*, 516-535.

Yonelinas, A. P., & Jacoby, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory and Language*, *34*, 622-643.

Yonelinas, A. P., Kroll, N. E., Dobbins, I., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: Convergence of remember-know, process dissociation, and receiver operating characteristic data. *Neuropsychology*, *12*, 323-339.

Appendix

accordion	cannon	Key	spoon	comb	Puzzle
acorn	carrot	Kite	stapler	cookie	Radio
airplane	cat	Knife	stool	crackers	Rainbow
alligator	caterpillar	Ladder	stove	curtains	Rocket
ambulance	cherry	Lamp	strawberry	diaper	rope
anchor	clock	Lemon	sweater	doctor	rose
ant	coat	Lion	tent	drill	scarf
apple	corn	lips	toaster	egg	shell
apron	couch	monkey	truck	fan	shovel
asparagus	cow	mushroom	trumpet	fire	skateboard
axe	crab	octopus	turtle	firetruck	slide
baby	crown	owl	violin	fishtank	snail
balloon	desk	panda	wagon	genie	spaghetti
banana	dice	pants	wheelchair	ghost	squirrel
barn	dog	peanut	whistle	globe	statue
basket	doll	pear	wrench	gorilla	stroller
beaver	dolphin	penguin	***extended***	hamburger	submarine
bed	door	pepper	antlers	helmet	tank
bee	drum	pig	backpack	hose	teapot
beetle	duck	pineapple	bag	knot	teepee
bell	ear	pizza	bandaid	ladybug	telescope
bike	elephant	porcupine	banjo	leaf	tie
blouse	feather	pumpkin	bathtub	lighthouse	tiger
bone	fish	rabbit	beard	lizard	toilet
book	flamingo	rake	belt	llama	tractor
boot	flute	robot	bench	magnet	trash
bottle	football	rooster	box	mailbox	turkey
brain	fork	ruler	bride	moose	tweezers
bridge	frog	sailboat	bucket	mouse	unicycle
broccoli	giraffe	satellite	butter	nest	vacuum
broom	glove	saturn	cane	nurse	volcano
burger	grapes	scissors	canoe	parrot	waffle
butterfly	guitar	screwdriver	car	peas	whale
cactus	hammer	shark	castle	pillar	wheat
cake	harp	sheep	chain	pillow	wheelbarrow
camel	horse	shoe	cheese	popcorn	windmill
camera	kangaroo	skirt	chest	popsicle	wine
candle	kettle	spider	church	purse	wolf