THE ASSESSMENT OF VERBAL AND IMAGINAL ENCODING PROCESSES IN THE BIZARRENESS EFFECT

Miri Besken

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology.

Chapel Hill
2011

Approved by:

Neil W. Mulligan

Kelly S. Giovanello

Peter C. Gordon

Joseph Hopfinger

Peter A. Ornstein

ABSTRACT

MIRI BESKEN: The assessment of verbal and imaginal encoding processes in the bizarreness effect
(Under the direction of Neil W. Mulligan)

The bizarreness effect refers to the finding that sentences that are contrary to the expectations or general world knowledge (e.g. *The dog rode the bicycle down the street)* produce superior memory as compared to mundane and schema-consistent sentences (e.g. *The dog chased the bicycle down the street*). There are two major accounts that try to explain the role of the encoding processes in the emergence of the bizarreness effect. The imaginal encoding accounts contend that increased visual processing for the bizarre sentences is responsible for the superior memory. The verbal rehearsal accounts, on the other hand, argue for the role of increased verbal elaboration processes for the bizarre items in the emergence of the bizarreness effect. This project aimed to discourage one of the processes during sentence encoding through the use of a concurrent working memory (WM) task known to selectively impair either verbal or visual WM, to investigate whether the size of the bizarreness effect decreases in any of the distraction conditions. Using tasks such as visual dynamic noise (Experiment 1), spatial tapping (Experiment 2) and visual patterns task (Experiment 3 & 6) that are well known to disrupt visual processing selectively, there was no decrement in the size of the bizarreness effect. Similarly, using tasks such as irrelevant speech (Experiment 1), articulatory suppression

(Experiment 2) and letter span task (Experiment 4 & 7) that are well known to disrupt verbal processing selectively, there was no decrement in the size of the bizarreness effect. The results strongly argue against the role of visual and verbal WM processes in the emergence of the bizarreness effect. The results are discussed in terms of Baddeley's working memory model and the attentional accounts of the bizarreness effect.

DEDICATION

This thesis is dedicated to my mother, Sandra Besken, for all the unconditional love and support she provided me with throughout my whole life.

ACKNOWLEDGEMENTS

To an outsider, a dissertation thesis might seem to be solitary work. However, there are individuals, whose contributions to my work are undeniable. Without their assistance, support and inspiration, this thesis would not have been possible.

First, I would like to express my deepest gratitude towards my advisor, Dr. Neil Mulligan, for his continuous assistance during my graduate school life. He always welcomed any questions or concerns that I have, helped me think more thoroughly with his careful consideration of my research, and provided me with continuous moral support and motivational feedback throughout the writing process. Moreover, his love of Cognitive Psychology, his infinite knowledge of the subject matter, his excitement for research and his willingness to help have always been a source of inspiration for me during the course of my graduate career.

I would also like to thank my committee members, Dr. Kelly Giovanello, Dr. Peter Gordon, Dr. Joseph Hopfinger and Dr. Peter Ornstein for their valuable intellectual contribution and constructive feedback. My interactions with them in and outside of class and dissertation committee have always been enjoyable, and their positive attitude throughout the dissertation process facilitated the quality of this project.

My academic career has also been shaped by some other mentors. Dr. Umur Talaslı was the first person to implant the love of Cognitive Psychology with his amazing lectures during my undergraduate years. Dr. Sami Gülgöz helped me gain the courage to conduct my own research and become an independent researcher who believes in her product. I would like to thank them both for making it seem so easy and so exciting.

I would also like to thank all my friends, who were there to provide me with moral support whenever I needed it through endless conversations and celebrations of all kinds. They made graduate school an amusing and enjoyable period in my life.

Last of all, I would like to thank my family for their constant support. They always believed in me and encouraged me at all times, even when I doubted myself. They believed that I should pursue my dreams, and they were prepared to miss me for extended periods of time as I left home for UNC. I would not have been the independent person that I am today, if it were not for their confidence in me and willingness to let me be who I want to be.

TABLE OF CONTENTS

LIST OF TABLES

Table

CHAPTER 1

INTRODUCTION

Bizarre Imagery and Memory Enhancement in Every-day Life

The idea that bizarre images and events facilitate memory has been around since ancient times. Ad Herennium, the oldest surviving Latin book dating back to 90 B.C., which focuses on the uses of rhetoric and persuasion, describes methods to use imagery to memorize speeches. The author of the book (attributed to Cicero, but author unknown) emphasized the use of ridiculous or weird images in order to maintain memories by describing how bizarre events stand out in our memory in everyday life:

> *"We ought, then, to set up images of a kind that can adhere longest in the memory. And we shall do so if we establish likenesses as striking as possible; if we set up images that are not many or vague, but doing something; if we assign to them exceptional beauty or singular ugliness; if we dress some of them with crowns or purple cloaks, for example, so that the likeness may be more distinct to us; or if we somehow disfigure them, as by introducing one stained with blood or soiled with mud or smeared with red paint, so that its form is more striking, or by assigning certain comic effects to our images, for that, too, will ensure our remembering them more readily. The things we easily remember when they are real we likewise remember*

> *without difficulty when they are figments, if they have been carefully delineated."*
>
> *(Ad Herennium, cited in Yates, pp.9-10).*

The author also suggested that people who are practicing the art of memory should generate active, distinctive, surprising and uncommon images, with exceptional beauty or ugliness rather than common, traditional, unsurprising ones in order to remember information *(Ad Herennium, cited in Yates, pp.9-10).*

Similarly, modern self-help books for memory enhancement have also advised that it is important to form uncommon associations between nouns in order to remember them. In their book on maximizing memory power, Lorayne and Lucas (1974), for example, emphasized the importance of absurdity in encoding the information.

> *"Making the pictures ridiculous is what enables you to really see them; a logical picture is usually too vague. Once you see the ridiculous picture, it registers in your mind (Lucas & Lorayne, p.33)… When something assaults our senses in an unusual, great, unbelievable or ridiculous way, it stirs the mind. It is usually retained without effort. It is the ordinary, everyday things that we have trouble remembering. Forming ridiculous pictures helps to make them outstanding, novel or marvelous. (p. 34)"*

If one had to remember the words *envelope* and *tree*, for example, Lorayne and Lucas suggest generating a ridiculous image of millions of envelopes growing on a tree. This, they argued, would enhance memory for those two items more than creating a common or mundane image, such as an envelope positioned next to a tree. Similar recommendations may be found in many self-help books on memory (e.g. Buzan, 1991; Lorayne, 2010; Trudeau, 1997).

In addition to the ancient texts and self-help books about memory enhancement, many people seem to believe that forming bizarre images enhances memory. Even though this idea has been around for thousands of years, surprisingly little experimental research has been conducted until the last four decades. Before going into the details, it is important to define bizarreness as the term is used in research, and explain how bizarre imagery has been operationalized and measured experimentally.

Operational Definition and Manipulation of Bizarreness in Modern Research

Bizarreness generally refers to items, sentences, pictures or mental images that are contrary to the expectations or general world knowledge of participants. Worthen (2006) defines bizarreness as an extreme form of distinctiveness, whereby the stimulus is in proportional minority relative to all previously stored knowledge. This definition of bizarreness also corresponds to Schmidt's (1991) definition of secondary distinctiveness. Borrowing terms from William James for primary and secondary memory, Schmidt differentiated between primary and secondary distinctiveness by emphasizing an item's standing in relation to the either currently activated memory schema, or long term store. Primary distinctiveness, in this sense, refers to an item which is only incongruent to the immediately activated structures (e.g., a word in red print in the context of a study list in which all other words are presented in black print). Secondary distinctiveness, on the other hand, refers to any item that stands out in terms of its incongruence to the general knowledge of participants. Bizarreness, in this sense, refers to manipulations where the target item is unusual in relation to the general world knowledge of the participants. The bizarre items are distinctive because the participants have rarely or never encountered an item, sentence, picture like that. For example, a bizarre sentence (e.g. *The dog rode the bicycle down the street)* is distinctive with respect to long term memory because it does not match the general experience or knowledge of the participants, and it has secondary distinctiveness regardless of whether it is presented in a list of common sentences (e.g. *The dog chased the bicycle down the street*) or other bizarre sentences.

There have been three main methods in how researchers manipulate bizarreness at encoding: image generation, image presentation and sentence presentation. In the image generation method, participants are given examples of common and bizarre images prior to the encoding phase. During encoding, participants are presented with single nouns or noun pairs, along with a prompt specifying the type of image (common or bizarre) to be generated. Once the image is formed, the participant is asked to describe the image with a sentence while trying to maintain the image (e.g. Cornoldi, Cavedon, De Beni & Pra Baldi, 1988; Weir & Richman, 1996).

Bizarreness has also been manipulated through image presentation which, as the name implies, is implemented by presenting participants with images that depict common or bizarre scenes (Nicholas & Worthen, 2009; Worthen, 1997). In this type of manipulation, participants are asked to associate the items in the picture for a pending memory test. For example, two words (e.g. *cigar, piano*) can be associated with a common image depicting a cigar balanced on a piano, or with a bizarre image where the piano is smoking the cigar (Wollen, Weber, & Lowry, 1972).

The sentence presentation method involves presenting participants with nouns (e.g. dog, bicycle, street) embedded in sentences that depict common images (e.g. *the dog* *chased the bicycle down the street*) or bizarre images (e.g. *the dog rode the bicycle down the street*) and instructing the participants to form images that correspond to these sentences (McDaniel & Einstein, 1986). The bizarreness of sentences has been manipulated in different ways by various researchers. In particular, the sentences in the bizarre condition vary across studies on a continuum from unusual (or implausible) but possible, to conceivable (if extremely unlikely or impossible) to anomalous and incomprehensible. For example, McDaniel and colleagues have usually used bizarre sentences that are unusual but possible, such as *The maid licked the ammonia off the table* to conceivable but extremely unlikely or impossible, such as *The dog rode the bicycle down the street*. In other research, bizarreness is defined as a type of anomaly, in which an image may be difficult or impossible to imagine. Collyer, Jonides and Bevan (1979) employed bizarre sentences such as *tablespoon blast lemonade*. Similarly, Graf (1980) gave his participants either common (e.g. *the tiny mouse frightened the cook*) or anomalous sentences (e.g. *the cheerful carpet exchanged the mouse*). As we shall see, using different types of bizarre sentences has implications for when the bizarreness effect is found.

In addition to the use of different bizarreness manipulations during encoding, the effects of bizarreness on memory are assessed with various measures of memory during the retrieval phase. The effect of bizarre imagery is often assessed with cued recall or free recall tests. In cued- recall, participants are typically presented with one of the words from the study sentence, image, or pair, and asked to recall other information from the study stimulus, such as the other word in a pair (e.g. Nappe and Wollen, 1973) or the entire sentence (Hirshman, Whelley, & Palij, 1989). Another way to test performance is through the use of free recall. In this method, participants are asked to recall the images, the sentences or the nouns in any order that they can remember them.

A short history in initial findings of the bizarreness effect

Even though bizarre images have long been recommended as a way to enhance memory, surprisingly, experimental research on bizarreness and memory only started in the 1960s. Earlier studies on the use of imagery in paired-associates learning were too unsystematic to be considered manipulations of bizarreness (e.g., Wallace, Turner and Perkins, 1957). Delin (1968, 1969) conducted the initial studies which actually examined the effects of bizarre imagery on memory performance relative to a rote rehearsal condition. In a list learning experiment, Delin (1969) instructed the participants in the experimental group to connect the current item to the previous item by thinking up as *fantastic a connecting image as they can* between the two words. For example, if the words pig and piano were presented in sequence, the participant might create an image of a pig playing the piano. The participants in the control group, on the other hand, were not given a specific encoding strategy and were simply instructed to memorize the presented words in the allocated time. The results indicated that participants who used the fantastic images showed superior memory than the control group in a free recall task. However, this study failed to assess the bizarreness of images systematically, as participants were not asked about the images that they formed for each pair of words. Consequently, the positive effects found could be simply due to the use of imagery rather than bizarre imagery, per se.

Another initial technique to investigate the effects of bizarre imagery was the use of a peg-word system. In this procedure, participants first memorized a list of rhyming peg-words that correspond to each number (e.g. one is BUN, two is SHOE, three is TREE, four is DOOR, five is HIVE, etc.). Then, they were presented with the words in the target list. As each word in the to-be-learned list was presented, participants were instructed to form a bizarre image between the peg word and the target word in the new list. For example, if the first word was *pencil*, participants would imagine a pencil in a bun. Using the peg-word method, Bugelski, Kidd and Segmen (1968) asked participants in the experimental group to associate the peg-words to the list words in 2, 4 or 8 seconds for each word in the target list in a between subjects design. The control group, on the other hand, was not informed about the peg-word mnemonic even though they also previously learned the number-peg word associations. The results indicated that participants who were given 4 or 8 seconds in the experimental group performed better than the control group. Still, the type of images that participants generated could not be assessed in terms of how unusual they were, because participants were again not asked about the types of images that they created.

This method was criticized by subsequent studies on two different grounds: First, there was no direct comparison between common and bizarre images, because the participants were instructed to generate bizarre images in all imagery conditions. Second, there was no way to know whether the images were bizarre or not, because participants were never asked to describe the images. Subsequent studies overcame these limitations by examining the effects of bizarre images on learning lists of nouns. Nappe and Wollen (1973) gave participants instructions to create either common or bizarre mental images for each word pair. Before the experiment began, the participants were given operational definitions and examples for both common and bizarre mental images. Then, participants were presented with 48 word pairs along with the type of image to be formed presented next to each word pair. Each participant was asked to press a button as soon as they formed the images and to give a verbal description of the mental image, for which the level of bizarreness was judged later by four different judges. At test, participants were given a cued-recall test, in which one noun from a pair was presented as a cue to recall the other noun. Surprisingly, the results indicated no difference in cued recall for common and bizarre images.

The foregoing manipulation of bizarre imagery has also been criticized by subsequent studies on the ground that the when the participants generate the mental image, the experimenter control is low and the level of bizarreness cannot be directly manipulated. In order to avoid this, Collyer, Jonides and Bevan (1972) conducted a study in which participants were asked to construct images for noun-verb-noun triplets. In a between subjects design, half of the participants received common triplets such as *tablespoon-measure-lemonade, shotgun-blast-building.* The rest of the participants received bizarre triplets formed by swapping the verbs of the sentences (e.g. *tablespoon-blast-lemonade, shotgun-measure- building).* The participants were told that they should construct images in which the first noun performs the action on the second noun. At retrieval, participants were asked to write down all the nouns that they could recall in no particular order. The results revealed superior recall in the common imagery condition than the bizarre condition, a result in conflict with the common belief that bizarre imagery helps memory performance more than the common imagery. Cox and Wollen (1981) reported a similar advantage for common imagery over bizarre imagery, also using pure lists at encoding.

However, other research showed superiority for bizarre sentences. Merry and Graham (1978), for example, gave their participants mixed lists of bizarre and common sentences with either intentional or incidental memory instructions. As in previous studies, participants created an image of each sentence, followed by a rating of image bizarreness. The results indicated that the incidental condition led to the superior memory for bizarre items in free recall compared to common items. The superiority of the bizarre sentences was diminished in the intentional condition, but was still present.

Given all the conflicting findings, an early review of the literature contended that the inconsistent findings put in question the common belief that bizarre imagery leads to better recall (Yarmey, 1984). Yarmey argued that the effects of bizarre imagery might be born out of the distinctiveness and elaboration of information. When participants have to process images that are bizarre, they attend to them more, and as a consequence they elaborate them more. Thus, it is not necessarily the bizarreness of the images that produce the effect, but it is the level of effort put into the sentence, the image or the word pair in order to associate and understand it.

The new era in bizarre imagery effect and imagery debate

Yarmey's view was partially challenged by a series of experiments conducted by McDaniel and Einstein (1986). In a classical experiment, McDaniel and Einstein (1986) presented participants with either a pure list of common sentences (e.g. *The maid spilled the ammonia on the table*), a pure list of bizarre sentences (e.g. *The maid licked the ammonia off the table*), or a mixed list of common and bizarre sentences. During the study phase, participants were instructed to form an interactive image of the sentence and rate the image for its vividness. This was followed by a surprise free recall test. For pure lists, recall was similar for bizarre and common sentences. For mixed lists, on the other hand, the recall of bizarre sentences was significantly higher than common sentences (Experiment 1). In this same series of experiments, they also investigated the effects of imagery instructions during the study phase. Participants were randomly assigned either to imaginal processing or semantic elaboration groups. In the imaginal processing group, participants were instructed to create a mental image of the sentences presented to them and then judge the vividness of the images. The participants in the semantic elaboration group, on the other hand, were asked to judge the degree to which the relation among the three underlined words in each sentence was unusual and then rate it on a 5-point scale. At test, participants were asked to recall all the underlined nouns that they could remember. The analysis of the total noun recall revealed an interaction between the sentence type and type of processing such that participants assigned to the imaginal processing group recalled more bizarre items than common items, whereas participants assigned to the semantic elaboration group recalled approximately equal numbers of bizarre and common items (Experiment 2).

In a review of the previous literature, Einstein and McDaniel (1987) outlined the necessary conditions for the bizarre imagery effect to be displayed. They contended that bizarre imagery effect is consistently obtained in mixed lists with free recall and imagery instructions. The finding that mixed lists followed by free recall produce the bizarreness effect has been replicated in many subsequent studies (e.g. Hirshman, Whelley & Palij, 1989; Waddill & McDaniel, 1998; McDaniel, Einstein, DeLosh, May & Brady, 1995). However, the necessity of the imagery instructions has been controversial.

For example, Kroll and Tu (1988) questioned the necessity of imagery instructions by giving participants mixed lists of bizarre and common sentences with different types of ratings designed to increase or decrease imaginal processing at encoding. In one experiment, one group of participants was given imagery instructions, where they were asked to generate images for the presented sentences and rate them for the vividness. The other group was given semantic processing instructions, where they were asked to judge the degree of bizarreness between the nouns in the sentence (bizarreness ratings). The results revealed a main effect for the bizarreness, with bizarre nouns recalled more than common nouns. Yet, neither the main effect of instruction nor its interaction with the bizarreness was significant, indicating that the superior memory performance for bizarre items is obtained even with instructions that emphasize semantic elaboration. Moreover, the vividness ratings for bizarre sentences were always lower than the ratings for common sentences. Thus, unlike McDaniel and Einstein (1986), Kroll and Tu (1988) argued that it is not necessarily the imaginal encoding that leads to increased recall on two grounds: First, that bizarreness effect was produced even when a semantic orienting task such as bizarreness rating is employed, even though bizarreness rating does not necessarily entail imaginal processing. Second, if the visual distinctiveness of the bizarre images is the reason for memory performance facilitation in the bizarre condition, one would expect the vividness of bizarre images to be rated higher than common images. Yet, research show that the vividness ratings for bizarre sentences are almost always lower than the ratings for the common sentences. Thus, Kroll and Tu (1988) argued that the bizarre imagery effect is a misnomer, and the superior performance for bizarre items should be called the *bizarre context effect*.

The use of sentences for studying the bizarreness effect raises additional issues because participants first have to understand the sentences and then imagine them, which will require some verbal processes. One way that researchers tried to overcome this problem is to ask participants to generate bizarre and common images for nouns, without providing sentences. Cornoldi, Cavedon, De Beni and Pra Baldi (1988) used this strategy, manipulating encoding instructions. In the imaginal processing condition, participants were told to generate an image for each noun, describe the image and use the rest of time to maintain the image. In the verbal processing condition, participants were asked to generate a sentence for each noun and rehearse it in the allotted time. The type of image or the sentence to be generated was manipulated within subjects as common and bizarre. Before the encoding phase, participants were given operational definitions and examples for bizarre and common items. At retrieval, participants were asked to recall as many nouns as they could. Free recall performance revealed greater recall in the bizarre than common condition. This effect was qualified by a marginally significant interaction such that the superior memory for the bizarre items was only significant in the imaginal processing group. Cornoldi et al (1988) argued that the use of imagery is especially important in obtaining the bizarreness effect, because the superiority for the bizarre images is only obtained in the imagery condition.

This study was criticized by Weir and Richman (1996) on several grounds. Weir and Richman argued that the interaction between instruction type and type of sentence only reached a lenient level for rejection of the null hypothesis ($p<.10$), and that this might only be due to sampling error. In addition to this, Cornoldi et al (1988) only used two groups in their study. One of the groups generated an image and described it with a sentence. The other group just generated a sentence. Thus, there was no group that only imagined the bizarre event without being asked to describe the image, resulting in the confounding factor of using both imaginal and verbal processing in imaginal elaboration instructions. However, the necessity of imagery can only be supported through the use of a group that actually imagines the event without describing it. Consequently, Weir and Richman (1996) gave their participants three kinds of instructions to generate common and bizarre sentences from single nouns or noun pairs: The imagery group received instructions to create an image using the nouns and focus on the image. The verbal rehearsal group was instructed to write a sentence using the materials presented and to reread the sentence until the presentation of the next word pair. The third group was instructed to both create an image and write a sentence describing the image and then focus on the image in the remaining time. The bizarreness effect occurred in all groups regardless of the instructions given to the participants. Weir and Richman argued that this result might be a consequence of both verbal and imaginal processes, and that imaginal encoding is not critical in obtaining the bizarreness effect.

The debate about the role of imagery in the bizarreness effect has been investigated in other ways as well. Anderson and Buyer (1994) reasoned that if imagining the sentences is the fundamental factor in obtaining the bizarreness effect, then participants with better imaging skills should show an increased bizarreness effect, whereas people with poor imaging ability should show comparable recall levels for bizarre and common items. They presented participants with mixed lists of common and bizarre sentences, followed by a recall test, and a battery of imagery ability tests. They found no correlation between imagery ability and recall. Further, those who rated images as less vivid showed a larger bizarreness effect. Thus, Anderson and Buyer (1994) concluded that it is not the imaginal processing per se that produces the phenomenon, but rather it is the type of instruction that causes participants to engage in deep rather than shallow processing.

Another way researchers have investigated the effects of imagery are with incidental vs. intentional memory instructions. Most of the studies reviewed so far employed incidental memory instructions, in which participants are told that the study investigates individual differences in imaging ability. Hirshman et al (1989) argued that intentional learning instructions encourage less imaginal processing and more semantic elaboration than overt imaginal processing instructions, because participants try to repeat the verbal sequence and elaborate it rather than focus on the image. If the bizarre imagery is due to semantic elaboration and not imaginal processing, the bizarreness effect should still occur with the intentional memory instructions. Thus, Hirshman et al. (1989) presented participants with a mixed list of bizarre and common sentences, and told them that they would be tested on their memory for these sentences. The participants were not asked to create mental images or perform vividness ratings. The free recall test yielded a main effect for bizarreness, with better free recall performance for bizarre than common sentences. Thus, Hirshman et al concluded that the bizarreness effect is independent of the imaginal processing instructions. Yet, one problem with the use of Hirshman et al (1989) argument is that even if the participants are not given any explicit visual orienting tasks, they can still spontaneously come up with a strategy to visualize the sentences, in which case they will do some additional processing for each sentence, helping their performance.

Hirshman's finding was challenged by Burns (1996). In a 2 (sentence type: bizarre vs. common) x 2 (instructions: intentional vs. incidental) design, participants were informed that they were participating in an experiment about imagery ability. The task was to form a mental image of the sentences that they read, followed by vividness ratings. The participants in the incidental group were told that the relationship between imagery and problem-solving would be measured, whereas the participants in the intentional learning conditions were told that their imagery ability would be related to their ability to remember the underlined words in the sentences. The results revealed an interaction between for the two learning conditions and the sentence type such that the recall in the incidental condition revealed the bizarreness effect whereas recall for common and bizarre items was equivalent in the intentional learning condition. Burns (1996) argued that the intentional learning instructions diminish the effect even if it does not totally eliminate it. He also contended that this finding might be a consequence of verbal elaboration at encoding. To explain, intentional learning instructions might foster elaborative rehearsal of the items, which leads to use of other cues than distinctiveness, decreasing the discrepancy between bizarre and common sentences. However, the findings of Burns were not replicated in other research. Merry and Graham (1978) obtained a robust bizarreness effect with intentional memory instructions, as did Worthen and Roark (2002). Thus, research using intentional memory instructions do not appear to resolve the imagery-verbal rehearsal debate for two reasons. First, intentional memory instructions yield inconsistent results. Second, the assumption that intentional instructions enhance verbal processing and diminish imaginal processing has not been tested directly.

The intentional memory instructions may induce imaginal processes rather than verbal rehearsal if the participants believe that this will enhance their memory.

Worthen (1997) argued that in the classical bizarreness paradigm, the requirement to read and understand sentences entails verbal processing, and the requirement to rate vividness entails imaginal processing. Thus, it is difficult to separate the effects of imaginal and verbal processing on the bizarreness effect. He reasoned that if the bizarreness effect is obtained even when the formation of mental images is controlled and imaginal instructions are withheld, then the bizarreness effect does not depend on the imaginal capacity. In order to test this idea, he presented participants bizarre and common sentences in four different conditions. In the image condition, participants were provided with black and white drawings of bizarre and common items, and asked to form a mental image of the drawing after it disappeared and rate the image for its vividness. In the sentence condition, participants were presented with common and bizarre sentences and were asked to rate the sentences for pleasantness. In the sentence plus image generation condition, participants were given sentences, asked to form images and rate them for vividness. Last of all, in the sentence plus image provision condition, participants were presented with the picture while reading the sentence. Once the picture disappeared, participants were asked to create a mental image of the picture and rate it for its vividness. The results yielded a main effect for bizarreness, regardless of the instructions given to the participants. Unlike the results by McDaniel and Einstein (1986), the pleasantness rating, hypothesized to induce high degrees of semantic processing, lead to increased memory for the bizarre items. Moreover, even when participants make up their own images or are provided with the images, the findings essentially failed to lead to any differences between the conditions. Thus, Worthen (1997) concluded that the effects are not limited to a specific set of images or to situations where images are unconstrained or self-generated and that the bizarreness effect is independent of imaginal and verbal processes.

Some other evidence for the verbal elaboration view comes from research investigating the effects of humor on the bizarreness effect. Worthen and Deschamps (2008) asked their participants to generate bizarre and common images for word triplets and say aloud a sentence that described each image. The participants came back 15 weeks later. In the meanwhile, all the sentences were judged by an independent group for humor and bizarreness. The free and cued-recall results showed that the original participants remembered more bizarre than common sentences only if the sentences were humorous. Worthen and Deschamps (2008) concluded that humor is an important element in obtaining the bizarreness effect, because it induces an exaggerated tendency to process the differences between the context and the item, leading to increased verbal elaboration for the bizarre events.

Another line of research that has been interpreted as supporting the verbal rehearsal account is the moderating effect of study time in producing the bizarreness effect. One of the factors that influences the bizarreness effect is the type of stimulus used. As described earlier, the bizarre sentences vary across studies on a continuum from atypical and possible to anomalous and incomprehensible. The bizarreness effect is consistently obtained in research using mixed lists of sentences that are bizarre and imaginable (Hirshman, Whelley &Palij, 1989; Waddill & McDaniel, 1998; McDaniel, Einstein, DeLosh, May & Brady, 1995). On the other hand, when bizarreness is defined as a type of anomaly, the bizarreness effect is often not obtained (e.g. Collyer, Jonides & Bevan, 1972). Yet, even anomalous sentences can produce the classical bizarreness effect if participants are given sufficient time at encoding. For example, Imai and Richman (1991) manipulated the degree of plausibility of sentences as common (e.g. *The maid spilled the ammonia on the table),* bizarre (e.g., *The maid licked the ammonia off the table)* or anomalous (*The table dropped the maid out of the ammonia)* At a processing time of 7 seconds, bizarre sentences were better remembered than common sentences, whereas the memory performance for anomalous sentences was equivalent to common sentences. In contrast, when participants were given 35 seconds for each sentence at encoding, the bizarreness effect was obtained for both bizarre and anomalous sentences. Imai and Richman (1991) argued that the moderation of time on plausibility might be a sentence reorganization effect rather than an imagery effect. In this sense, common sentences do not violate the expectations of the participants; thus, they are not reorganized verbally. Bizarre sentences, in contrast, are contrary to the world knowledge of the participants and violate their expectations. Thus, participants try to understand the sentences by relating the bizarre sentences into common sentences, which produces a richer, more detailed memory trace, leading to increased recall. Since anomalous sentences are harder to understand and re-organize verbally, longer duration is necessary for the bizarreness effect to appear for anomalous sentences as compared to bizarre and common sentences. Thus, it is the spontaneous re-organization of the sentence into a more logical form that leads to better memory.

However, the same finding may be explained just as well with a hypothesis that depends on imaginal processes. Wollen and Margres (1987) explained the bizarreness effect in terms of the imagery-multiprocess model. According to this model, when a participant processes a sentence, the participant first begins with one or more schematic images from long term memory. For common items, the process is easy: For example, if participants are given the sentence *The dog chased the bicycle down the street,* participants have to choose a schematic image of a dog and a bike and imagine the dog chasing the bike without making too many transformations on each visual element of the image. Bizarre images, on the other hand, require more transformations for each element, because the schematic images of the elements have to be altered. For example, when the participants receive the sentence *The dog rode the bicycle down the street,* the schematic dog image has to be transformed so that the hind legs of the dog reach the pedals, and the paws reach the handle bars, leading to increased visual elaborations for the image. According to this view, it is the nature of the elaborations and the number of changes required to convert a schematic image into a bizarreness image that leads to increased memory performance. In some cases, though, the images will not be easy to form, especially for bizarre sentences, which allow this account to explain the effects of study time on obtaining the bizarreness effect with anomalous sentences. Therefore, the effect of increased encoding time in producing the effect with anomalous sentences might be due to either verbal rehearsal processes, whereby the sentence is reorganized, or to imagery processes, whereby more visual transformations lead to increased memory performance.

Increased encoding duration has a similar effect on complex bizarre sentences as it does on the plausibility of the stimulus: For complex sentences, longer study durations are required to obtain the bizarreness effect, as occurs with anomalous stimuli. Much of the research on the bizarreness effect employed sentences which are short, concrete and simple like "The *maid* licked the *ammonia* off the *table*", and participants were asked to remember the nouns *maid, ammonia* and *table.* The simple short sentences produce the typical bizarreness effect in mixed-list designs (e.g. Hirshman, Whelley & Palij, 1989; McDaniel, Einstein, DeLosh, May & Brady, 1995; Waddill & McDaniel, 1998). However, the employment of more complex sentences yields inconsistent results. Kroll, Schepeler and Angin (1986) operationalized complexity with the number of modifiers and actions in a sentence. By this definition, a simple common sentence employs only one verb with no modifiers for the nouns (e.g., *The ant goes around a comb*). A complex sentence, on the other hand, includes modifiers for the nouns and more descriptive information, such as *A large, black ant crawls in and out of the comb* for the common version, and *A large black ant carefully fixes its hair with a plastic comb* as an example of a complex bizarre sentence. With these stimuli (and a relatively brief study duration of self-paced study), Kroll et al (1986) failed to obtain any differences between complex common and bizarre sentences.

Using only the complex sentences, Richman (1994) manipulated the encoding time per sentence (15, 30 or 60 seconds) and used free recall as the memory test. Free recall was superior for the bizarre than common sentences for the 30 and 60 second study times, but there was no bizarreness advantage in the 15 second condition. Richman argued that this effect was a consequence of verbal comprehension and reorganization processes. To explain, when participants are given sufficient time, they spontaneously reorganize bizarre sentences into common and meaningful ones, thus encoding the bizarre sentences in both bizarre and common versions, leading to more retrieval routes later than for common sentences, which are only encoded in one way. Because complex sentences are harder to understand than simple sentences, it takes longer to reorganize complex bizarre sentences; thus, this reorganization can only produce an effect for longer study durations. Even though Richman (1994) argued that the effect is a consequence of verbal rehearsal, this finding might again be a result of imaginal processes. In line with Wollen and Margres's imagery multiprocess hypothesis (1987), when participants are given extra adjectives such as *tired* and *crabby*, it might take longer for the participants to integrate the adjectives into the image, leading to the bizarreness effect at longer encoding times.

To sum up, the encoding processes employed in obtaining the bizarreness effect have been a source of controversy. Some researchers have proposed that visual imagery forms the basis of the classical superiority for bizarre items for two reasons. First, the bizarreness effect is not always obtained with instructions that emphasize semantic elaboration. Moreover, the intentional memory instruction, which is hypothesized to rely heavily on semantic elaboration instructions rather than imagery generation, leads to decrements in the effect size for the bizarreness manipulation (at least in some studies). Theoretically, two variants of the imagery hypothesis exist. McDaniel and Einstein (1986) claimed that bizarre images are more visually distinctive than common images, leading to increased memory performance. Many other studies have tried to assess the importance of the visual distinctiveness and how it increases recall of bizarre items (McDaniel, Dornburg & Guynn, 2005; McDaniel, Einstein, DeLosh, May & Brady, 1995; Waddill & McDaniel, 1998). Wollen and Margres (1987), on the other hand, claimed that participants put more effort into generating the bizarre images, because bizarre images cannot be obtained by using schematic pictures of the elements and have to be transformed more than common elements, a form of visual elaboration leading to superior memory.

In contrast, other researchers argue that imagery is not necessary for obtaining the bizarreness effect for several reasons. First, the effect is obtained even in studies that do not employ imagery instructions (Weir & Richman, 1996; Wollen, 1997). Second, there is no correlation between participants' self-reported use of imagery and the recall rates (Anderson & Buyer, 1994). Moreover, the participants with higher imagery capabilities do not show an increased superiority for bizarre items. Third, intentional memory instructions, which assume an increased reliance on semantic elaboration, yield contrary results. The proponents of verbal rehearsal processes explain the superior memory for bizarre items through the use of sentence re-organization or other verbal elaboration accounts. Bizarre sentences require greater modification or semantic elaboration for complete comprehension than do common sentences, leading to better memory.

Disentangling the role of verbal and imaginal processes in obtaining the bizarreness effect has been challenging, because most extant research employed sentences. Even when participants are asked to generate common and bizarre images from nouns or noun pairs, participants are typically asked to describe the image to the experimenter as a manipulation check, leading to the use of sentences and potential verbal elaboration. A consideration of a study by Worthen (1997) highlights the typical ambiguity. Worthen reasoned that if the bizarreness effect is obtained even when the formation of mental images is controlled and imaginal instructions are withheld, the bizarreness effect must not depend on the imaginal capacity. Even though this interpretation is reasonable, his manipulations of imagery and verbal processes were not necessarily unambiguous. For example, in one condition, the participants were shown a picture and after it disappeared, they were asked to regenerate the image and rate it for vividness. This manipulation aimed to focus on imaginal processing (and minimize the use of verbal elaboration). In another condition, participants were presented with bizarre and common sentences, and given semantic elaboration instructions. The bizarreness effect was found in both conditions, leading Worthen to conclude that bizarreness effect can occur due to either verbal elaboration or imagery processes. Even though this is an interesting methodology, the manipulations do not fully rule out the employment of the verbal or imagery processes. To explain, participants may spontaneously generate an image of the sentence in the semantic elaboration condition, giving them increased routes for retrieval. For example, a participant in the semantic elaboration condition, reading the sentence *The dog rode the bicycle down the street,* may spontaneously generate the image of a dog on a bicycle. Similarly, participants who are given images with imagery instructions may try to verbalize the event and try to assess its unusualness in order to understand the picture. For example, if the participant sees the drawing of a dog on a bicycle, they might verbally elaborate on the situation by thinking how counterfactual this picture is, relative to all their world knowledge, leading to increased verbal elaboration.

Perhaps a better way to disentangle the contribution of imaginal and verbal processes to the bizarreness effect would be to selectively impair either verbal or imaginal processes separately in order to see if the bizarreness effect still persists when one of the processes is impaired. Thus, if a situation could be engineered to disrupt either verbal or the imaginal processes during the encoding of the sentences, we might be able to better determine the basis of the effect. For example, if the bizarreness effect were still obtained while the participants' imaginal processes were impaired, this would indicate that imaginal processes are not responsible for the effect. On the other hand, if the bizarreness effect were obtained when verbal processes were disrupted, we could infer that verbal processes are not the basis of the effect. One way to achieve this is to use secondary working memory tasks, which employ verbal or visual working memory during encoding. In order to explain how this would work, it is necessary to briefly review the mechanics of working memory.

Working Memory

One of the most prominent models of working memory was devised by Baddeley and his colleagues (Baddeley, 2003; Baddeley & Hitch, 1974). According to this model, working memory is composed of multiple limited-capacity subsystems, consisting of a central executive with two slave systems: phonological loop and visuo-spatial sketchpad. The phonological loop is involved in processing verbal information. It comprises a phonological store, which can hold sound based memory traces for a few seconds before they fade, and an articulatory process that is analogous to subvocal speech. It is also referred to as verbal working memory. Visuo-spatial sketchpad is involved in processing visual and spatial stimuli, such as color, shape and spatial location. Both of these systems have limited capacity, and can only process, maintain and store a certain amount of information within a certain time. The visuo-spatial sketchpad is also referred to as visual working memory. The central executive, an attentional system, controls and coordinates the functioning of the two slave systems (Baddeley, 2003). The phonological loop is hypothesized to facilitate language acquisition and processing verbal information, where visuo-spatial sketchpad is implicated in nonverbal intelligence. The central executive is hypothesized to achieve more complex tasks such as reasoning and comprehension, with the help of the slave systems. [1]

There is compelling evidence that these two slave systems work independently of each other. For example, the functioning of the phonological loop remains largely unaffected when paired with a secondary task requiring visual or spatial processing. In a classical experiment, Baddeley, Logie, Bressi, Della Sala and Spinler (1986) investigated performance on two demanding tasks. The first task was a digit span task, in which participants heard digits and immediately recalled them in their original order. This task requires the use of phonological loop as the participant has to process and retain verbal information. The second task was a perceptuo-motor tracking task, in which participants were required to follow a randomly moving light patch on the screen with a light sensitive pen. This task demands the use of visual sketchpad, as tracking can only be achieved through processing visual and spatial information. Each task was calibrated in isolation. The digit span was adjusted so that each participant was given as many numbers at each trial as they could retain. Similarly, the speed of the light patch was adjusted so that the participants were only able to track it 40-60% of the time. Baddeley et al (1986) hypothesized that if the working memory consists of one flexible system, the performance in the verbal task should drop when it is performed concurrently with the visual task. On the other hand, if the working memory consists of two different systems,

[1] Baddeley (2003) has also introduced a third slave system called episodic buffer which is hypothesized to link information across domains to form integrated units of visual, spatial, and verbal information. However, since it is not directly relevant to the current research, it will not be explained any further.

verbal and visual, working independently of each other, the performance on one task should not be diminished by the performance on the other. The results showed that performance on the digit task did not change when coupled with the light patch task, providing evidence that the phonological loop is not affected by the presence of the secondary task when the secondary task involves visual and/or spatial processing.

There is also evidence that performance on immediate visual-memory tasks is unaffected by secondary tasks that employ the phonological loop, but is severely impaired by secondary tasks that use the visual sketchpad. Measures of immediate verbal memory show the opposite pattern. In an experiment by Logie, Zucco and Baddeley (1990), participants were given one of two short-term span tasks. In the visual span task, a matrix of randomly colored black and white squares was presented, then the matrix disappeared for 2 seconds, then the matrix reappeared with a color change for one of the squares. Participants were asked to point out which square had changed in color. The other task was a letter span task. For this task, each letter appeared sequentially at the center of a computer screen. Then, the sequence disappeared for two seconds, and then the same sequence was presented to the participant again, with one of the letters changed. Participants were asked to identify the changed letter. All participants were exposed to these two tasks in counterbalanced order. Concurrently, they also were presented with a secondary task. Half of the group received an arithmetic task, where the experimenter asked participants to add numbers given to them and say the obtained number aloud. The other half of the group generated an image of a three by five matrix of squares, for which the experimenter read aloud a sequence of instructions for the squares to be filled and left blank. When arithmetic was used as the secondary task, letter span performance was significantly lower than the visual span task. In contrast, when the secondary number matrix task was used, visual span performance dropped substantially more than the letter span task. The selective decline in the primary task performance when the secondary task employs the same slave system confirms further that visual sketchpad and phonological loop are qualitatively different from each other, and can be selectively disrupted with an appropriate secondary task.

The detrimental effects of employing the same system for primary and secondary tasks are observed not only with immediate recall, but also over brief delays. Cocchini, Logie, Della Sala and McPherson (2002) tested the effects of a few different tasks for an analysis of the detrimental effects of the secondary tasks on the primary task on a longer retention period of 15 seconds. In experiment 1, participants were given three different tasks: perceptuo-motor tracking, in which they followed a ladybug on the screen with a light sensitive pen; serial digit recall, in which they recalled a series of numbers in the order presented; and visual pattern recall, in which they were presented with a matrix of randomly filled squares and then were asked to reconstruct the pattern in an empty matrix. During the experiment, participants first carried out all the tasks individually. Then, participants carried out the tasks concurrently in pairs of two. Digit recall was not impaired by either the tracking or the visual pattern recall task. On the other hand, performance dropped significantly in the visual pattern recall when carried out concurrently with the tracking task, but not with the serial digit recall task. In Experiment 2, the perceptuo-motor tracking task was replaced with the articulatory suppression task, a manipulation that has been shown to disrupt the contents of verbal working memory. In this task, participants are asked to repeat a word aloud at a certain rate (e.g. saying *the* at a rate of three times per second). Experiment 2 interpolated this task with serial digit recall and visual pattern recall to see whether this task would lead differentially decrease performance. The results complemented Experiment 1: visual pattern recall was disrupted by neither articulatory suppression nor the digit serial recall. In contrast, digit serial recall performance dropped significantly when combined with the other tasks. More importantly, digit span recall was significantly worse when combined with articulatory suppression than with visual pattern recall. This study confirms that using two tasks that employ the same working-memory subsystem impairs both tasks significantly, whereas when two tasks employ different subsystems, minimal impairment occurs for either task.

In the current study, using two different concurrent tasks that impair one of the slave systems but leave the other one intact can help disentangle the effects of verbal rehearsal and visual imagery on the bizarreness effect. If the bizarreness effect is diminished or eliminated by a distractor task that employs verbal working memory (i.e., the phonological loop), this would indicate that verbal rehearsal processes are responsible for the effect. On the other hand, if the effect is impaired when a visual distractor task is used, it would be consistent with a visual-imagery account of the bizarreness effect.

Use of Visual Distractors in Working Memory

One of the tasks that have been known to interfere with visual working memory is the irrelevant pictures task (Logie, 1986). In this task, participants are presented with irrelevant pictures of living and non-living objects that change constantly, with the instruction that they should ignore the pictures without looking away from the screen. Logie (1986) presented his participants with irrelevant pictures as a secondary task as they were trying to memorize short lists of words presented on the headphones, with two different memory instructions: rote rehearsal and pegword mnemonics. Participants in the rote rehearsal condition were told to commit the words to memory by adding each word to the one heard previously and to repeat them subvocally in the order that they were presented. Participants in the visual mnemonic condition first memorized a group of peg words (one-BUN, two-SHOE, three-TREE, etc.) to be associated to the target list. Then, as the word in the target list was presented, they were asked to generate an image that integrated the target word with the pegword. The free recall performance indicated that rote rehearsal condition produced similar levels of memory both in the presence and the absence of the irrelevant pictures. In contrast, the pegword mnemonic instructions produced better memory in the absence of irrelevant pictures than their presence. This implies that disruption to visual working memory can disrupt memory effects reliant on imagistic encoding, a conclusion consistent with subsequent, more refined research.

Quinn and McConnell (1996) criticized Logie's (1986) demonstration, arguing that the irrelevant picture task might employ the central executive system as well as the visuo-spatial sketchpad, because meaningful pictures might attract more attention than random displays of meaningless stimuli. They devised a technique called dynamic visual noise (DVN). DVN is a rectangular display, consisting of small black and white squares on a computer screen. The squares change randomly and continuously between black and white over the presentation period, very similar to the experience of viewing an out-of-tune TV monitor. In their study, Quinn and McConnell used this rectangular dynamic visual noise display, changing randomly at a continuous on/off rate of 291 dots per second in order to see whether dynamic visual noise selectively disrupted visual working memory, without invoking the central executive. As in Logie (1986), they assigned participants to two different memory instructions: rote rehearsal and visual mnemonic rehearsal conditions. Moreover, participants in each condition were subjected to three different distractor conditions. In no distraction condition, participants looked at the blank screen while they were trying to learn the words. In irrelevant pictures condition, participants received a series of line drawings that changed every 5 seconds while they were trying to memorize the words. In the dynamic visual noise condition, the participants received the constantly changing dynamic visual noise display. In all encoding conditions, participants were told to look at the screen and to ignore the displayed stimuli without closing their eyes. At test, they were asked to recall the words in the order that the list was presented, with the place number (e.g. one) serving as the cue word. The researchers hypothesized that any type of visual distraction should diminish performance for participants assigned to visual mnemonic instructions more than participants assigned to rote rehearsal condition. Moreover, the dynamic visual noise should only impair performance in the visual mnemonic condition as it only interferes with the visuo-spatial sketchpad, whereas the irrelevant pictures should impair performance in both visual mnemonic and rote rehearsal conditions, as the changes in irrelevant pictures may draw on the central executive.

The results yielded the expected results. First, there was a main effect of visual distraction, with no distractor leading to better memory performance than either distractor condition. Second and more critically, there was an interaction between memory instructions and the type of distractor. In the rote condition, drawings interfered with rote rehearsal but visual noise did not, whereas in the visual mnemonic condition, both visual noise and line drawings interfered with memory performance and did so to an equal degree (experiment 1). One counter argument that could be made against the findings is that visual mnemonic is more difficult than rote rehearsal; thus, any kind of secondary task will disrupt the visual mnemonic condition more than it does rote rehearsal, not because of the selective visual interference of the DVN, but because of the difficulty of the visual mnemonic. Thus, Quinn and McConnell also manipulated the modality of the distraction as follows: 1) visual presentation of word lists with irrelevant speech presented through the headphones 2) visual presentation of the lists with no irrelevant speech 3) verbal presentation of the lists while watching the dynamic visual noise on the monitor 4) verbal presentation of the lists while watching the blank screen (experiment 3). They reasoned that if it is the difficulty of the visual mnemonic that leads to the decreased performance, all secondary tasks should lead to more impairment with the visual mnemonic condition than rote rehearsal. In contrast, if it is the selective visual disruption of DVN that leads to decreased memory performance, irrelevant speech, a verbal task which impairs verbal memory but leaves visual memory intact, should lead to more impairment in rote rehearsal condition than visual mnemonic condition. As in Experiment 1, half of the participants received the visual mnemonic instructions, whereas the other half received rote rehearsal instructions. The results indicated that with visual mnemonic instructions, memory performance was only diminished in the dynamic visual noise condition compared to the three other conditions. In contrast, with rote rehearsal instructions, performance only decreased in the irrelevant speech condition. This provides further evidence that dynamic visual noise selectively interferes with the visual working memory, and not with the verbal working memory (see also Andrade, Kemps, Werniers, May and Szmalec, 2002).

The effectiveness of the DVN paradigm has also been shown with experiments concerning the long-term retention of words. Parker and Dagnall (2009) examined the effect of DVN on the concreteness effect. The concreteness effect refers to the finding of superior memory for words that signify concrete objects such as *desk*, compared to words that signify abstract concepts such as *justice*. It has been hypothesized that the concreteness effects arise because the concrete objects are encoded in both verbal and imaginal systems and thus possess multiple retrieval routes. In contrast, abstract words can only be encoded verbally, leading to fewer retrieval routes (Paivio, 1971; 1991). Parker and Dagnall reasoned that if concrete words entail the processing of both visual and verbal information, then a distracting visual task should selectively interfere with the former processing and selectively reduce memory for concrete words. In the study phase of the experiment, participants heard abstract and concrete words while passively watching either a static visual noise display or a dynamic visual noise display. Static visual noise comprised of a large rectangular display with fixed random black and white dots on it. For the DVN, the display changed constantly with the condition that the proportion of the back to white dots within the display was maintained. The type of interference was manipulated between subjects, whereas the type of word was manipulated within. At retrieval, participants received a free recall test. The results yielded a significant interaction between word type and interference condition such that when participants were presented with the static noise field, the classical concreteness effect was obtained, with concrete words recalled more than abstract words. In contrast, the effect was reversed in the dynamic visual noise condition, with abstract words being recalled more often than concrete words. Moreover, the type of interference did not affect recall for abstract words, whereas it significantly decreased recall for concrete words (experiment 1). This finding was also replicated with a recognition test (Experiment 2). This finding is important for three reasons. First of all, this study shows that DVN can be successfully used in studies of long-term memory, especially when an encoding condition requires imaginal processes. Secondly, the maintenance of the memory performance for abstract words and the decline in the performance for concrete words in the dynamic visual noise condition indicates that DVN was able to selectively interfere with the imaginal codes, providing further support for the use of this distractor condition. Last of all, in the previous studies, participants were actually trained to use the visual mnemonics in learning the words. In this study, participants were not given any explicit instructions to use the imaginal processes. However, DVN was able to successfully reduce performance for imaginal processes even when participants were not given any explicit instructions to use them.

Another task that selectively interferes with the visual working memory is the spatial tapping task, where the participants are asked to repeatedly tap a certain sequence of keys in the shape of a figure. Spatial tapping has been repeatedly shown to disrupt visual working memory, but not verbal memory. For example, in a study Salway and Logie (1995) presented participants with primary tasks, which either employed verbal or visual processes. In the visual primary task, participants were given instructions to put certain letters to certain locations in a matrix by the sentence format, such as *in the starting square put an A. In the next square to the right, put a B.* In the verbal task, the instructions were altered from directions to adjectives so that they did not indicate spatial locations so that participants could only memorize the verbal directions with no reference to spatial locations. An example for this is as follows: *in the starting square put an A. In the next square to the slow, put a B.* At test, participants were asked to recall the sentences verbally. For the secondary task, participants either had to perform spatial tapping or articulatory suppression. Recall in the visual primary task decreased with spatial tapping but not with articulatory suppression. Similarly, for the verbal primary task, performance decreased with articulatory suppression but not with spatial tapping. Moreover, research also shows that spatial tapping task is a more demanding task than DVN. Andrade et al (2002) presented participants with static matrix patterns filled randomly with black and white squares to retain in short term memory. This pattern was either followed by DVN or spatial tapping in the retention period. The experiment showed that participants' performance drops more in spatial tapping compared to DVN. Thus, spatial tapping differs from DVN in that it reduces performance more than DVN and it requires repeated responses from the participants.

Use of Auditory Distractors in Working Memory

One of the tasks that has been frequently used as a distractor for verbal working memory is irrelevant speech. In a first demonstration, Salame and Baddeley (1982) showed that presentation of irrelevant speech impairs serial recall of digits presented visually. In this experiment, participants tried to memorize sets of nine digits on each trial. They were exposed to three different distractor tasks: meaningful speech, white noise and silence. In the meaningful speech condition, participants listened to a set of prerecorded three letter words for each digit that they saw. In the white noise condition, the distractor consisted of noise bursts during the presentation of the digits. The silence condition served as the control. In all conditions, participants were instructed to recall the digits in the original order immediately after the presentation of the last digit, while trying to ignore anything that they hear over the headphones. The irrelevant meaningful speech reduced serial recall performance relative to the white noise and silence conditions, whereas the latter two conditions did not differ (LeCompte, 1994). This effect has also been shown with auditory materials (e.g. Hanley & Broadbent, 1987). Using a dichotic listening paradigm, Hanley and Broadbent (1987) presented participants with 9 single digits through one headphone for immediate serial recall, while irrelevant speech or silence was presented in the other headphone. Serial recall was lower in the irrelevant speech condition. In a comparison of the of the effects of irrelevant speech on visual and auditory target lists, Jones, Macken and Nicholls (2004) asked participants to memorize lists of 7 digits that were either seen or heard. As the lists were presented, participants were exposed to a repeated token of syllables, changing token of syllables, or to silence. The experiment yielded two results: First, there was no recall difference between the auditory and visual presentation of the target lists. Second, the changing irrelevant syllables impaired recall compared to repeated syllables and silence, but the last two conditions did not differ from each other (see Hanley& Bakapoulou, 2003 for a similar result)

The meaningfulness of the irrelevant speech affects recall performance as well. Oswald, Tremblay and Jones (2000) showed that the encoding of the visual verbal material was impaired more when participants heard meaningful irrelevant speech (e.g. an excerpt from the evening news) compared to meaningless irrelevant speech (e.g.

excerpt from evening news played backwards), which in turn impaired working memory compared to a quiet condition, a finding replicated in other studies (e.g. Bell, Buchner & Mund, 2008; Van Gerven, Meijer, Vermeeren, Vuurman & Jolles, 2007). This finding is in some ways similar to the findings in visual working memory. Just as irrelevant meaningful pictures impair recall more than DVN condition for tasks requiring the use of visual processes in working memory (Quinn & McConnell, 1996), meaningful speech impairs recall more than meaningless speech for tasks requiring the use of verbal processes in working memory. This might be indicative of greater central executive involvement in ignoring meaningful speech.

Another critical issue to the current project is that even though irrelevant speech impairs verbal rehearsal processes, it does not impair basic comprehension of sentences. In a study about the effects of IS on reading comprehension, Boyle and Coltheart (1996) presented their participants with acceptable or unacceptable sentences that varied in complexity. Simple sentences contained a right-branching relative clause such as *The applause pleased the woman that gave the speech*, whereas the syntactically complex sentences included a center-embedded relative clause *The hay that the farmer stored fed the hungry animal.* The acceptability of the sentences was manipulated by using either homophones (e.g. *The teacher taut the child that played the clarinet)* or orthographic control words (e.g. *The teacher tight the class that painted the poster)*. Participants' task was to make a decision about whether the sentence was grammatically correct or not. While they were reading the sentences, participants heard background sound with a silent control condition and four noise conditions: (1) instrumental music, (2) singing with instrumental accompaniment, (3) unaccompanied singing, and (4) speech. The results indicated that even though the complexity of sentences increased reading time and decreased decision accuracy, the background noise had no effect on comprehension of the sentences, because the reaction time to decision making was statistically insignificant, and the errors made in accepting or rejecting unacceptable sentences did not change across noise conditions.

Thus, meaningless irrelevant speech appears to be a good candidate as a verbal distractor for a few reasons. First, the evidence indicates that irrelevant speech disrupts rehearsal of other verbal information. Baddeley (2003) argued that this is because speech has direct and automatic access to the phonological loop. Moreover, even though irrelevant speech disrupts the recall of verbal information in working memory, it leaves the comprehension processes intact (Boyle & Coltheart, 1996). Furthermore, irrelevant speech has been confirmed as an effective verbal distractor task in previous studies, employing imagery. Both Logie (1986) and Quinn and McConnell (1996) showed that the recall with peg-word mnemonic instructions, which requires the use of visual aids, was unaffected by irrelevant speech, even though recall in rote rehearsal conditions was impaired by it. Last of all, theoretically it seems to be an equivalent verbal analogue for the dynamic visual noise display.

Another candidate for a verbal distractor task is articulatory suppression (AS), the repeated articulation of a specific word or sound. A classical experiment by Richardson and Baddeley (1975) examined the effect of articulatory suppression on free recall. Participants were presented with lists of 16 words. For half of the lists, participants were asked to say “hi-ya” repeatedly as the words were presented on the computer screen. For the other half of the lists, participants remained silent. The results indicated that the free recall of the lists was significantly impaired in the articulatory suppression condition, as compared to the silent condition. The study by Cocchini et al (2002), described earlier, provides further evidence that articulatory suppression impairs serial digit recall, a task known to employ verbal rehearsal, whereas it leaves visual pattern recall intact.

One other important aspect of articulatory suppression is that even though it prevents verbal rehearsal, it does not impair comprehension. In a study, Calvo and Castillo (1995) investigated the effects of anxiety on reading comprehension when participants have to perform a concurrent working memory task. Relevant to the current study, in the low anxiety condition, participants were presented with a text in four different conditions: no interference, articulatory suppression, meaningful irrelevant speech and meaningless (reversed) speech. At test, participants had to answer true-false questions assessing comprehension of the text. Articulatory suppression and meaningless speech produced levels of comprehension equal to the no-interference condition (only meaningful speech decreased comprehension). This confirms that participants have no trouble comprehending text in the presence of meaningless speech or articulatory suppression.

The detrimental effect of articulatory suppression has also been shown with stories in long term memory retrieval. Christoffels (2006) found that participants listening to stories over the headphones while completing an AS task had lower delayed recall performance for the story than the participants who did not have to complete any concurrent tasks. Thus, articulatory suppression prevented the participants from rehearsing the story, leading to decreased recall performance, in a delayed free recall test. This finding shows that articulatory suppression can also be a good candidate as a task that disrupts the verbal working memory performance for a few reasons. First, AS leads to decreased recall performance for verbal material relative to the silence condition. Secondly, even though it disrupts the memory for verbal material, it leaves the visual memory intact, as participants are still able to maintain their performance in tasks requiring visual memory, such as perceptuo-motor tracking tasks. Last of all, the impairment of articulatory suppression on free recall memory performance has also been shown with long-term retention.

Current Study

Given the debate about the basis of the bizarreness effect, the goal of the current study is to determine whether the critical factor is imaginal processes or verbal rehearsal. The critical step is to find a method to disentangle the role of each process in obtaining the bizarreness effect (Worthen, 1997). Earlier studies have tried to achieve this by encouraging one type of processing more than the other. This has been done in different ways. Some researchers have used different instructions such as incidental vs. intentional instructions (Burns, 1996; Merry & Graham, 1978; Worthen & Roark, 2002), or imaginal vs. semantic elaboration instructions (Kroll & Tu, 1988; McDaniel & Einstein, 1986; Worthen, 1997). Other researchers have manipulated materials that depended either more or less on imaginal encoding, such as sentence presentation (McDaniel & Einstein, 1986), sentence generation from nouns (Cornoldi et al, 1988), or image presentation (Nappe & Wollen, 1972; Nicholas & Worthen, 2009; Worthen, 1997). Other research tried to find answers by looking for relationships between the bizarreness effect and other factors, such as imaging ability (Anderson & Buyer, 1994), plausibility (Imai & Richman, 1991) or complexity (Richman, 1994). Although a main goal in the previous research was to parse out the effect of each element by encouraging one type of processing, the studies do not produce unambiguous results; as reviewed earlier, the studies do not necessarily rule out the alternative hypothesis, leading to an unresolved controversy.

Perhaps a better way to disentangle the contribution of imaginal and verbal processes would be to impair one of the processes through the use of a secondary task and see if the bizarreness effect persists. There has been only one study that investigated the bizarreness effect with the use of a secondary task. Worthen, Garcia-Rivas, Green and Vidos (2000) investigated how the reaction time to a secondary task changes during the comprehension of common and bizarre sentences. However, this study is not particularly relevant because the secondary task was not designed to selectively disrupt one type of processing. Thus, the goal of the project is to identify which process is critical by impairing the verbal or imaginal components of working memory selectively, and seeing under what conditions the bizarreness effect is obtained.

CHAPTER 2

EXPERIMENT 1

In the study phase of this experiment, participants were presented with bizarre and common sentences in a mixed-list design. Two experimental groups also carried out a secondary task designed to occupy one of the slave systems of working memory, either a secondary task that occupies verbal working memory or a task that employs visual working memory. The control groups did not carry out a secondary task.

The secondary tasks to be used in this experiment needed to fulfill several requirements in order for the tasks in different modalities to be equivalent. To begin with, both of the secondary tasks should impair only one of the slave systems, leaving the other system intact. Thus, the tasks chosen were shown by prior research to disrupt verbal working memory but not visual working memory, or vice versa. In addition, the secondary tasks should have similar qualities in terms of the cognitive load, the meaningfulness and the continuity of the stimulus. For example, if the task in the visual domain is continuous and meaningless to the participants throughout the encoding of the sentences, the task in the verbal domain should also be continuous and meaningless. Last of all, both tasks should have similar instructions, and thus require similar types of responses from the participants. For example, if the secondary task for visual working memory requires the participants to look at the stimuli while trying to ignore it, the secondary task for verbal working memory should similarly require the participant to hear the stimuli while trying to ignore it.

As explained earlier, dynamic visual noise (DVN) is an ideal candidate as a secondary task to disrupt visual working memory for a few reasons. First, it impairs the visual working memory selectively, leaving verbal working memory intact. Moreover, because the dynamic visual display is meaningless, no response is required, and verbalization is minimal, the task requires minimal use of the central executive or the phonological loop. In addition, it has been successfully used in long-term memory experiments and has proven successful in disrupting imaginal processes for complex stimuli.

For the distraction of verbal working memory, irrelevant speech (IS) is analogous to dynamic visual noise in many respects. Just like DVN, IS impairs the verbal working memory selectively and leaves the visual memory intact. IS has been successfully used to disrupt verbal elaboration and rehearsal. IS requires no response from the participant; they are instructed to passively listen to the stimuli and ignore it. Moreover, IS can be modified to be even more similar to DVN in other important respects, such as continuity and meaningfulness. Thus, if participants are presented with continuous speech in a language that they do not know (such as Turkish), the speech will still have the characteristics of speech, such as changing tokens of sound, intonation, articulation, etc., but will be meaningless to the participants; thus, the cognitive load of IS on the central executive and the visual sketchpad will be minimal, analogous to the visual distraction condition. Therefore, for the first experiment, DVN and IS served as the secondary tasks.

During the study phase, participants were presented with the common and the bizarre sentences for a fixed period of time and asked to create a mental image for the sentence and then rate each image for its vividness. The vividness rating paired with incidental-learning instructions has been found to maximize the bizarreness effect. In the control condition of Experiment 1, we wanted to obtain the largest possible effect size to maximize our ability to detect interactions.

Participants in the visual distraction condition listened to the bizarre and the common sentences over headphones while simultaneously watching DVN on the computer screen. They were told to look at the screen at all times. Similarly, participants in the verbal distraction condition were presented with the common and the bizarre sentences on the computer screen while they listened to the IS over the headphones. In both cases, the study trial was followed by the vividness rating. Two separate control groups were presented with the sentences in either visual modality or auditory modality: The control group for the visual distraction group listened to bizarre and common sentences over the headphones, without exposure to DVN. The control group for the verbal distraction group saw the sentences on the computer screen without exposure to IS. Based on prior research, there is no reason to think that the size of the bizarreness effect in the control condition should be impacted by modality of presentation. However, two control groups were necessitated by the need to present the sentences in different modalities in the two experimental conditions.

Method

Participants

Forty-eight undergraduates from the University of North Carolina participated in exchange for course credit.

Design and Materials

The design was a 2 (sentence type: bizarre and common) x 4 (group: visual distraction, verbal distraction, visual distraction control, verbal distraction control) mixed design, with sentence type manipulated within subjects and group manipulated between subjects. The stimulus set consisted of 16 nouns triplets, each used to create one bizarre and one common sentence, taken from McDaniel and Einstein (1986), and producing a total of 32 sentences. The bizarreness of the material was manipulated by presenting sentences that depict scenes which are either possible and unsurprising or impossible and surprising. For example, for the word triplet *banker, newspaper, puddle,* the common sentence was *The banker dropped the newspaper in the puddle,* whereas the bizarre sentence was *The banker floated across the puddle on a newspaper.* The thirty-two sentences, consisting of 16 bizarre and 16 common sentences, were further randomly divided into two mixed lists of 8 common and 8 bizarre sentences. These two lists were counterbalanced across subjects so that each noun triplet was presented in bizarre and common contexts equally often. For each list, two common sentences from the original McDaniel and Einstein (1986) were used as practice sentences (See Appendix. A for a complete list of materials used). For auditory presentation, each sentence was recorded to an audio-file in a neutral voice. For visual presentation, each sentence was shown on the computer screen.

The irrelevant speech consisted of 10-s segments of a voice reading sentences from a story in Turkish. The dynamic visual noise was based on the technique used in Quinn and McConnell (1996), and consisted of a large rectangular display covering 80% of the computer screen and consisting of small black and white squares. The squares randomly either remained the same color or changed color (from black to white or vice versa) every .25 seconds with a probability of .5, giving the sense of a flickering rectangular display, with the ratio of the black and white squares maintained equal throughout the presentation.

Procedure

Participants were told that the experiment was about individual differences in visual imagery in the face of distraction. They were told that they would read or listen to sentences while they tried to ignore a secondary stimulus that might be presented to them. Furthermore, they were instructed to create visual images of the sentences and rate the vividness of each image on a scale of 1 (not at all vivid) to 5 (very vivid). During the study phase, participants were presented with the common and the bizarre sentences in a random order, with the restriction that no more than two sentences of the same kind were presented consecutively. Keeping in line with McDaniel and Einstein (1986), participants were not informed about the nature of the sentences. Participants were given two practice trials with common sentences.

In the visual distraction condition, the sentences were presented over the headphones. Each trial started with a short beep for half a second, immediately followed by the DVN display, which was presented for a total of 10 seconds. Two seconds after the DVN was initiated, the sentence was presented over the headphones. Participants were given 8 seconds to encode each sentence and create a mental image, followed by a screen asking them to rate the vividness of the image within 4 seconds. Once they rated the sentence, the program proceeded onto the next trial. The whole trial lasted up to14.5 seconds. In this condition, participants were instructed to watch the dynamic visual noise presented on the screen for the duration of encoding. Participants were told to look at the screen, without looking in any other direction. They were also told that the noise display is only presented to try to distract them from the sentence task. The experimenter monitored the participants' gaze to ensure compliance with the instructions.

In the verbal distraction condition, the sentences were presented visually. Each trial started with a half-second beep, immediately followed by the irrelevant speech presented over the headphones. Two seconds after the initiation of IS, the sentence was presented on the screen for 8 seconds. Participants were then prompted by the screen to enter their vividness rating within 4 seconds, after which the next trial began. As in the visual distraction condition, each trial lasted for up to 14.5 seconds. In this condition, participants were told that the speech was only presented to try to distract them from the sentence task.

In the control conditions, the trial started with a beep and a blank screen. Two seconds later, the sentence was presented either visually (the verbal distraction control group) or aurally (the visual distraction control group). Eight seconds later, participants entered the vividness rating and then proceeded onto the next sentence. As in other conditions, each trial lasted up to 14.5 seconds.

After the study phase, participants were asked to solve math problems for 3 minutes in order to minimize recency effects (e.g. 67 + 46 =____). The distractor task was followed by the testing phase. Participants were asked to recall and write down as many sentences as they could remember from the encoding phase. They were also told that they should write any fragments, nouns or verbs that they might recall, even if they did not remember the whole sentence.

Results

The use of concurrent WM tasks

Previous research has used a few different measures of recall, although the measures generally yield very similar results. We used the two most common measures: the total noun recall and the sentences access measure. The sentence access measure counts a sentence as accessed if at least one noun from the sentence has been recalled. The total noun recall refers to the proportion of nouns recalled. Both measures were computed separately for bizarre and common sentences (Table 1).

The primary analyses compared performance in the visual or verbal distraction groups, compared to their respective control groups. First, the verbal distraction group and its control group were compared. For this analysis, the dependent variables of total noun recall and sentence access were separately submitted to an ANOVA, with the item type (bizarre vs. common) as the within subjects variable and the distraction condition (verbal distraction vs. verbal distraction control) as the between subjects factor. First, the analyses of the access measure showed a significant effect of sentence type, with bizarre sentences accessed more often than common sentences ($F(1,22) = 24.26$, $MS_e = .026$, $p<.01$). There was no significant effect of experimental condition ($F(1,22) = 2.07$, $MS_e = .023$, $p>.10$, nor interaction ($F<1$). The analysis of total noun recall produced identical results: a main effect of sentence type, $F(1,22) = 14.81$, $MS_e = .022$, $p<.01$, and no effect of group or interaction ($Fs<1$).

Before conducting the primary analyses for the visual distraction group, it was necessary to investigate whether the bizarreness effect emerged when the sentences were presented aurally, instead of visually as in the vast majority of bizarreness research. An analysis of the visual-distraction control group revealed no significant different between the bizarre and common condition for either the access measure (*t*(11) = .92, *ns*) or total noun recall (*t* (11) = .34, *ns*). Thus, the bizarreness effect was not found in the control condition using auditory presentation, a point to which I return in the discussion section.

Even though the bizarreness effect was not obtained in the visual-distraction control group condition, the second a-priori ANOVA, in which the visual distraction group was compared to its control group, was conducted. The access measure revealed a main effect for the sentence type ($F(1,22) = 17.76$, $MS_e = .021$, $p<.01$), with bizarre sentences ($M = .52$, $SD = .18$) accessed more often than common sentences ($M = .34$, $SD = .16$). There was no significant effect for experimental condition ($F<1$), but the interaction was significant ($F(1,22) = 7.43$, $MS_e = .021$, $p<.05$), with bizarre items accessed more than common items in the visual distraction condition, and no difference between common and bizarre items in the control group. The proportion of total noun recall followed the identical pattern, with a significant main effect for sentence type ($F(1,22) = 11.63$, $MS_e = .017$, $p<.05$), no effect of experimental condition ($F<1$, *ns*), and a significant interaction ($F(1,22) = 8.23$, $MS_e = .017$, $p<.01$), with more nouns recalled from the bizarre sentences in the visual distraction condition but not in the control condition.

Vividness ratings

Vividness ratings were submitted to a 2 x 4 ANOVA, with sentence type as a within-subjects factor, and the experimental condition as the between subjects factor. The results yielded a main effect for the bizarreness such that the vividness ratings for common sentences ($M = 3.96$, $SD = .83$) were significantly higher than for bizarre sentences (*M*= 3.01 *SD* =.89) ($F(3,44) = 24.69$, $MS_e = .87$, $p<.01$). Neither the effect of experimental group ($F(322) = 1.54$, $MS_e = .61$, $p>.05$) nor the interaction reached significance ($F<1$).

Discussion

First, the bizarreness effect was replicated in the visual presentation mode. The participants who were presented with the sentences on the screen remembered more bizarre sentences than common sentences, regardless of whether they were in the verbal distraction group or the control group. Moreover, the size of the bizarreness effect did not differ across groups, indicating that the verbal distractor task did not modify the bizarreness effect. Irrelevant speech is an effective task in terms of selectively disrupting the phonological loop. Thus, the finding that the size of the bizarreness effect was not reduced by irrelevant speech shows that the bizarreness effect might not be associated with the verbal slave system.

Secondly, the bizarreness effect was not obtained with auditory presentation in the control group. Even though there was a slight advantage for bizarre sentences over common sentences, the finding did not reach significance. Still, the comparison of the visual distraction group to its control group warrants consideration in terms of the original goal of the study. I had hypothesized that if the bizarreness effect depends on imaginal processes, there should be an interaction between these two factors such that the advantage for the bizarre sentences decreases or is eliminated in the visual distraction condition. When the DVN condition is compared to its control group, the findings do not show a reduction in the size of the bizarreness effect. On the contrary, the concurrent presentation of DVN along with the sentences produced a robust bizarreness effect, significantly larger than in the control condition (and numerically larger than the other two groups). Thus, this experiment indicates that visual distraction does not decrease the bizarreness effect. Thus, neither of the concurrent tasks, which are well known to disrupt one of the slave systems, selectively impaired the recall of bizarre sentences.

The finding that the bizarreness effect is not obtained in auditory modality is interesting. Most bizarreness studies have employed visual modality, and obtained the effect with an average of 12-16 participants in each condition. To my knowledge, there is only one published study that used auditory presentation of the sentences, and this study demonstrated a significant bizarre effect (Anderson & Buyer, 1994). However, Anderson and Buyer's (1994) study used a group consisting of 80 participants. Thus, the bizarreness effect that they obtained in the auditory modality might be associated with the number of participants that they used. The fact that there are no other published studies of bizarreness conducted in the auditory modality might be indicative of the possibility that this effect is not generally obtained in the auditory modality especially with smaller sample sizes. It is unclear why this might happen, but since the effect was not obtained in the control condition, the aural presentation of sentences was not used in the subsequent experiments.

As an additional goal to the study, I investigated how the vividness ratings relate to sentence type and experimental condition. One of the most consistent findings in the literature is that the vividness ratings are always lower for the bizarre items as compared to the common items (e.g. Burns, 1996; McDaniel & Einstein, 1986; Robinson-Riegler & McDaniel, 1994). This has been explained in different ways: For example, Baddeley (2000) suggested that vividness depends on the amount of detail that can be represented in working memory. Since the bizarre images are not very usual, the richness of representation in working memory is low, leading to lower ratings for the bizarre items. Similarly, Wollen and Margres (1987) suggested that imagination of bizarre sentences requires more transformations for the image as compared to common sentences, and the difficulty to transform the images also results in subjectively less vivid ratings of the images. Thus, one would expect that participants in the visual distraction condition should have more difficulty in general in terms of visualizing the image, especially if the bizarreness effect is moderated by the visual system. I investigated whether the vividness of the images is quantified differently for common and bizarre sentences at encoding with respect to experimental condition. The results suggested that the vividness ratings for the bizarre images were lower than the common sentences, as expected from previous research. However, there was no main effect for the experimental group. One reason for this might be that since the experimental condition is a between subjects factor, participants have no chance to compare the distraction condition to no distraction condition; thus, they only take into account the ease of processing, relative to the other within subjects factor (sentence type). Even though it would be quite interesting to compare the ratings with the experimental condition as a within subjects variable, this idea is beyond the scope of this project. Another possibility is that the bizarreness effect is not moderated by the visual slave system. Therefore, the experimental condition does not affect the feeling of vividness for the images.

Some other analyses that investigate the relationship between vividness, memory accuracy, sentence type and experimental condition are also reported in the Appendix B.[2]

[2] Since the results are quite inconsistent across experiments, they are reported in the Appendix and discussed in the general discussion section.

CHAPTER 3

EXPERIMENT 2

Experiment 2 further investigates the bizarreness effect using alternate secondary tasks to disrupt visuo-spatial processing, on the one hand, and verbal rehearsal processes, on the other. Although the secondary tasks of Experiment 1 were defensible starting points for the present investigation, the failure to obtain the bizarreness effect in the control condition for the visual distraction condition requires that the sentences be presented in the visual modality. Moreover, the failure to obtain any kind of impairment in the verbal task suggests that we should use tasks requiring more active involvement from the participants, instead of the passive listening of IS or the passive viewing of DVN.

As well as increasing the workload on participants, there are some additional questions to be answered by the current experiment. In terms of visuo-spatial working memory, it is important to investigate the effect of spatial processes in the bizarreness effect. In the original working memory model, the visuo-spatial sketchpad supports both visual and spatial processing. However, there has been some debate about their separability. For example, Baddeley (2003) made a distinction between visual and spatial processes that take place in the visuo-spatial sketchpad. Alternatively, Salway and Logie (1995) found that performance on an imagery task was decreased to the same degree by a secondary task requiring spatial processing as by a visual secondary task. Thus, this study used an alternate secondary task, spatial tapping, which is hypothesized to impair the spatial aspects of working memory.

With regard to the verbal secondary task, an issue that needs further attention is the verbal storage versus verbal rehearsal of the sentences. To explain, verbal-WM is hypothesized to have two components: the phonological store, which retains phonological representations of verbal material in a short-lived phonological store, and the articulatory rehearsal process, which serves to reactivate and maintain phonological representations (Baddeley, 2003). IS is hypothesized to disrupt the storage of the verbal material, whereas articulatory suppression disrupts active rehearsal. Since IS did not impact the bizarreness effect in Experiment 1, it is necessary to determine if active rehearsal processes rather than passive storage of verbal information moderates the bizarreness effect. Hence, an articulatory suppression (AS) task is used in Experiment 2.

As described in the introduction, AS is a good task for disrupting verbal rehearsal selectively, without affecting visual working memory. Similarly, spatial tapping is a good task for selectively disrupting spatial processing. Spatial tapping can be considered analogous to articulatory suppression in important ways. First, both tasks selectively impair the verbal or visual domain. Second, both tasks require active, repetitive responses from the participants. Moreover, by using tasks that require responses, participants' cognitive load is increased which might produce more challenges than tasks that do not require responses, such as IS or DVN.

As in experiment 1, participants in all groups were presented with common and bizarre sentences. Only visual presentation of the sentences was employed, as auditory presentation in the control condition of Experiment 1 did not produce the bizarreness effect. Participants in the verbal-distraction condition were asked to count from 1 to 4 in a repetitive fashion during the presentation of the sentences. Participants in the visuospatial distraction condition were asked to tap pegs on a tapping board repetitively. As in Experiment 1, it was hypothesized that if the bizarreness effect is a consequence of spatial-imaginal processes, the superior performance for bizarre sentences should be reduced or eliminated in the visuo-spatial distraction condition and stay intact in AS condition. On the contrary, if the bizarreness effect is a consequence of verbal rehearsal processes, the superior performance for bizarre sentences should be eliminated or diminished by the AS task and remain intact for the spatial tapping task.

Method

Participants

Ninety participants from the University of North Carolina participated in the experiment for exchange for course credit or $10.

Design and Materials

Sentence type (bizarre vs. common) was manipulated within subjects and secondary task (articulatory suppression, spatial tapping, and control) was manipulated between subjects. The sentence materials of Experiment 1 were used. The spatial tapping board consisted of four wooden cubes (2.5 cm on a side), arranged 14 cm apart in a diamond shape on a square wooden board. This was modeled on the procedure of Smyth, Pearson, and Pendleton (1988). In order to make sure that the participants tapped the keys or counted rhythmically, audiofiles with the metronome beat of 120 times a minute (2 per second) were recorded.

Procedure

The study phase was the same as Experiment 1 with the exception of the change in the secondary tasks. The visuospatial distraction condition made use of spatial tapping. Each trial started with a beep for half a second, followed by the metronome beat over the headphones. The participants were instructed to tap the four pegs on the board repetitively, following a diamond shape pattern, with each tap corresponding to a single beat of the metronome. In order to ensure that the participants did not look at the board, a separate board was placed between the computer screen and the tapping board to block the view of the tapping board. The participant could reach around the occluding board and comfortably carry out the tapping task. Two seconds after the beep sound, the sentence was presented on the screen for the duration of 8 seconds. As in Experiment 1, the participant was instructed to read the sentence and create a mental image within this time duration. Next, the vividness rating was displayed on the screen, requiring a response within 4 seconds, yielding a trial lasting up to 14.5 seconds.

The verbal distraction condition made use of a standard articulatory suppression task in which participants were instructed to count aloud from 1 to 4 repetitively (ONE-TWO-THREE-FOUR-ONE-TWO…) as they read the sentences (Macken & Jones, 1995, Toppino & Pisegna, 2005). Each trial started with a beep followed by the metronome beat over the headphones, signaling the participants to start articulating. Two seconds later, the sentence was presented on the computer screen for 8 seconds while the participant continued to count aloud to the metronome beat. Then, the participants provided the vividness rating within 4 seconds. For both distraction conditions, the participants practiced the distractor task by itself prior to the presentation of the main study trials for

30 seconds, and were given feedback. Then, they had two practice trials with two common sentences to assure that the participant understood the procedure fully.

The control condition was identical to the control condition of Experiment 1, except for the beat of the metronome over the headphones. The distractor and testing phases were the same as Experiment 1.

Results

The use of concurrent WM tasks

The sentence access and the proportion of total noun recall were separately calculated for common and bizarre sentences for both the experimental groups and the control group (Table 2).

Table 2

The access and total noun recall were separately submitted to a 2 x 3 ANOVA, with sentence type as a within subjects factor and the experimental condition as a between subjects factor. The access measure revealed a main effect for the sentence type ($F(1,87) = 51.60$, $MS_e = .021$, $p<.01$), with the bizarre sentences accessed ($M = .45$, $SD = .18$) more often than common sentences ($M = .30$, $SD = .16$). Neither the main effect of experimental condition nor the interaction reached significance (Fs < 2). The identical pattern was obtained for total nouns recalled. The proportion recalled for the bizarre items ($M = .36$, $SD = .17$) was significantly higher compared to common items ($M = .25$, $SD = .14$), ($F(1,87) = 39.25$, $MS_e = .015$, $p < .01$). Neither the experimental condition nor its interaction with the sentence type was significant (Fs < 2).

Vividness ratings

Vividness ratings were submitted to a 2 (sentence type) x 3 (experimental group) ANOVA. The results yielded a main effect of sentence type, with common sentences rated more vividly than bizarre sentences ($F(1,87) = 118.06$, $MS_e = .33$, $p < .01$). There was no main effect of experimental condition ($F(1,87) = 2.109$, $MS_e = .1.054$, $p > .10$), but the interaction was significant ($F(2,87) = 3.34$, $MS_e = .33$, $p < .05$). Post-hoc analyses revealed no difference in vividness ratings of the common items across experimental condition ($F<1$). However, participants in the control condition rated the bizarre items as less vivid than the other two experimental groups ($F(2,87) = 4.22$, $MS_e = .77$, $p < .05$).

Discussion

As in the previous experiment, there was a significant advantage for the bizarre sentences compared to common sentences. Thus, the bizarreness effect was replicated in the control condition. There was no main effect of experimental group and no interaction between sentence type and experimental group. This indicates that neither articulatory suppression nor irrelevant speech impaired the encoding of the sentences. Articulatory suppression is a task well known to selectively impair verbal working memory and articulatory rehearsal processes (e.g. Cocchini et al, 2002; Richardson and Baddeley, 1975) and has successfully produced long term memory impairment as well (Christoffels, 2006). For this reason, the fact that the AS did not reduce the size of the bizarreness effect puts into question the role of phonological loop and verbal rehearsal processes in the emergence of the bizarreness effect. Similarly, spatial tapping has been documented to selectively disrupt visuospatial processing and cause decrements in imaginal processing (e.g. Salway & Logie, 1995). Considering that there was no decrement in the size of the superiority for the bizarre sentences, the role of visuo-spatial processes in the emergence of the bizarreness effect should also be questioned. Taken together, Experiments 1 and 2 have shown that concurrent working memory tasks which are well known to selectively impair one of the slave systems have no impact on the bizarreness effect.

The analyses pertaining to participants' vividness ratings replicated Experiment 1, revealing lower vividness ratings for the bizarre sentences as compared to common sentences. In addition that even though there was no statistically significant difference in the memory performance for the control group and the working memory groups for the bizarre sentences, participants in the control group rated the bizarre items as less vivid than the other working memory conditions. Thus, the presence of concurrent memory tasks may actually have caused the participants to experience the vividness of the sentences differentially. However, this difference in the subjective vividness rating is not necessarily predictive of the recall, at least for the bizarre sentences, given that the size of the bizarreness effect is not affected by the experimental condition.

CHAPTER 4

EXPERIMENT 3

The first two experiments investigated the effect of secondary working memory tasks, specifically designed to impair one of the slave systems and leave the other one intact. These experiments made use of standard distractor tasks known to selectively impair either visual-spatial WM (the DVN and spatial tapping tasks) or verbal WM (the IS and AS tasks). However, not only did these distractors not impact the bizarreness effect, none produced any significant impairment to memory performance at all. The previous experiments may not have impaired the encoding of the sentences because the secondary tasks were either passive or very repetitive. Perhaps the bizarreness effect only reduces in size with concurrent tasks which have an active maintenance component throughout the processing of the sentences, which change at each trial. For this reason, experiment 3 and 4 used concurrent working memory tasks that require active maintenance of either verbal or visuospatial information.

Experiment 3 further assessed the imaginal account of bizarreness effect. The experimental group carried out a secondary task designed to occupy visual working memory, which required the maintenance of a visual pattern while generating images for bizarre and common sentences. Participants were presented with an image which they had to maintain while they encoded the sentence, followed by another image, which they judged to be the same or different from the first image. The participants in the control group were exposed to the same images but were not asked to maintain or compare the images. This design was modeled after Della Sala, Gray, Baddeley, Allamano and Wilson's (1999) visual span test. In this test, participants were provided with a checkerboard pattern for 3 seconds, consisting of randomly-placed black and white squares in a grid (which varied in size from 2 x 2 to 5 x 6). A blank grid of the same dimensions was then presented and participants tried to reproduce the pattern by filling in the black squares. This task was modified for the current experiment in a few ways. First, the memory for the grids was tested through a recognition test rather than a recall test, since filling in squares in a recall test would take more time than desired and cause disruptions in the flow of the procedures. Moreover, the visual span test was given after an 8-second delay (rather than immediately), with a sentence inserted between encoding of the visual grid and recognition. In this experiment, only 5x6 grids were used to maximize the employment of visual working memory capacity.

This design differed from Experiment 1 and 2 in a few ways. To begin with, this experiment had an active maintenance component, which changed on each trial. Thus, the responses that the participants provided could not be automatized throughout the study. Moreover, participants were tested on how well they maintained the image in the distractor task. This provides two advantages. First, participants had to make a response in the visual distraction condition after they rated the sentences, which assured that there was little extra time to think about the sentences after the ratings. Secondly, since participants were tested on each trial for their maintenance of the visual pattern, it was possible to determine whether the successful maintenance of the visual image actually reduced the memory performance for bizarre and common sentences on a trial-by-trial basis.

Method

Participants

Thirty-two participants participated in exchange for course credit or monetary compensation.

Materials and design:

Sentence type (bizarre vs. common) was manipulated within subjects and task (visual vs. control) was manipulated between subjects. The sentence stimuli were the same as in the previous experiments. The material for the visual part of the test consisted of checkerboard patterns which are difficult to code verbally, adapted from Della Sala et al. (1999). A black and white visual pattern was created by randomly filling in half of the squares in a grid of 5x6. Eighteen patterns were prepared (2 for practice trials and 16 for experimental trials), along with a modified version of each in which one of the black and white squares were switched, making the shape slightly different than the original one.

Procedure

Each trial began with a half-second beep followed by a visual matrix for four seconds. This was followed by a sentence for eight seconds, the vividness rating for up to 4 seconds, and another visual pattern for 4 seconds. The second pattern was either the same as the original or the altered version. Half of the patterns were the same and half different. Each trial lasted up to 20.5 seconds. Participants in the visual distraction group were told to keep the matrix in mind while reading the sentence and creating a mental image of the sentence. They were told that they would see another visual matrix after the vividness rating task and this matrix might be the exactly the same as the first matrix or it might be slightly different. Participants were told to press “s” for same or “d” for different.

The participants in the control group were told that they would see a visual matrix, which they should look at carefully, followed by the sentence and vividness rating task. In addition, they were told that they would see another matrix right after the rating task and that they should look at it carefully as well. Participants were not told anything about the similarity or difference between the patterns presented before and after the sentences. In both the control and experimental groups, participants were given two practice trials to get them used to the procedure. The distractor and test phases were the same as Experiments 1 and 2.

Results

Concurrent WM tasks

The descriptive statistics for the access measure and the total noun recall for Experiment 3 are presented in Table 3.

The dependent variables of access and total noun recall were separately submitted to 2 (sentence type: bizarre vs. common) x 2 (experimental condition: visual distraction vs. control) ANOVAs. The analysis for the access measure indicated a main effect of sentence type ($F(1,30) = 17.73$, $MS_e = .020$, $p<.01$), with the bizarre sentences ($M = .38$, $SD = .19$) accessed more often than common sentences ($M = .24$, $SD = .14$). Furthermore, there was a main effect for the experimental condition ($F(1,30) = 5.62$, $MS_e = .029$, $p<.05$), with more sentences accessed in the control condition ($M = .36$, $SD = .17$) than in the visual distraction condition ($M = .26$, $SD = .14$). The interaction was not significant ($F<1$). The same pattern was replicated for the total noun recall: A significant main effect for sentence type ($F(1,30) = 10.88$, $MS_e = .017$, $p<.01$), with higher total noun recall in bizarre sentences (*M* =.30 , *SD*=.19) than the common sentences (*M* =.19 , *SD*=.14); a significant main effect for the experimental condition (($F(1,30) = 8.48$, $MS_e = .028$, $p<.01$), with lower total noun recall in the visual distraction condition (*M* =.30 , *SD*=.18), than control condition (*M* =.18 , *SD*=.12), and no interaction ($F<1$).

The test of the visual distraction requires that participants are performing the distractor task as they are processing the sentences. The mean correct completion rate for the distractor task was .75, significantly above the chance level of .50 ($t(15)= 8.11$, $d = .25$, $p<.01$). Moreover, a paired samples t-test showed that participants' successful completion of the distractor task was not different for bizarre (*M* =.76 , *SD*=.18) and common (*M* =.74 , *SD*=.14) sentences ($t(15)= .43$, $d = .02$, $p>.5$). A conditionalized analysis of the recall data was also conducted in which only the trials with correct answers on the visual working memory task were analyzed (see Table 3 for descriptives). Each sentence on the recall sheet was evaluated to determine if the participant completed the visual working memory task for the specified trial accurately. Sentences from inaccurate trials were excluded. Thus, the conditionalized analysis is limited to trials for which we have evidence that the participant is doing the concurrent task. Moreover, because the participants are only doing the task in the distraction condition, the control condition scores for the conditionalized analysis remain unchanged. The conditional results showed the exact same pattern as the non-conditional scores. For the conditional access measure, the effect of sentence type was significant ($F(1,30) = 10.98$, $MS_e = .027$, $p<.01$), with greater access for bizarre than common sentences, as was the effect of experimental condition ($F(1,30) = 10.03$, $MS_e = .025$, $p<.01$), with lower access in the visual experimental condition. The interaction was not significant ($F<1$). The conditional total noun recall revealed the identical pattern: A significant main effect for the sentence type ($F(1,30) = 6.76$, $MS_e = .020$, $p<.01$), a significant effect for the experimental condition ($F(1,30) = 10.53$, $MS_e = .026$, $p<.01$), and no interaction ($F<2$).

Vividness ratings

The vividness ratings were submitted to a two-way ANOVA with the sentence type and the experimental condition as factors. The data yielded a main effect for the sentence type ($F(1,29) = 24.36$, $MS_e = .59$, $p<.01$), with common sentences ($M = 3.97$, $SD=.80$) producing higher vividness ratings than bizarre sentences ($M = 3.00$, $SD= .98$). All the other effects were non-significant (*F*s<1).

Discussion

Experiment 3 investigated the effect of a visual concurrent task, which contained an active maintenance component, on the bizarreness effect. The experiment yielded higher memory performance for bizarre sentences, as compared to common sentences. Moreover, experimental condition had an effect on memory performance, with lower performance under visual distraction. Thus, the encoding of the sentences was impaired in general with the visual patterns test as a concurrent task. However, the interaction was not significant, indicating that the size of the bizarreness effect was not moderated by a visual task. The visual patterns task reliably occupies visual working memory, and correlates with many other measures of visual working memory (Della Sala et al, 1999). Even though this visual distractor task with active maintenance component disrupted overall memory performance, it did so proportionally for common and bizarre items.

Along with the results of Experiments 1 and 2, the present results provide no evidence that the visuospatial sketchpad is involved in the bizarreness effect.

As an additional motivation, the vividness ratings were analyzed. Thus, even though the vividness ratings for the bizarre items were lower than the common items, the rating difference was similar on both the visual distraction condition and the control condition. This analysis replicates Experiment 1. The finding that vividness ratings do not change in experimental and control conditions might be a consequence of the between subjects design for the experimental condition: Participants only depend on the contrast of the bizarre and common items as a basis, and not on the visual cues. The vividness ratings are similar in both experimental groups, even though the performance in experimental condition is lower. This provides further proof that vividness ratings are unrelated to memory performance.

CHAPTER 5

EXPERIMENT 4

The previous experiment assessed the visual encoding account of the bizarreness effect, with a concurrent task that required participants to maintain a visual matrix in mind while processing common and bizarre sentences. Even though the visual task led to a general decrement in memory performance, the size of the bizarreness effect was not moderated by the concurrent visual task, providing further evidence against a role for the visual sketchpad as a mechanism in the emergence of the bizarreness effect. The goal of the current experiment was to see whether a concurrent verbal task, which contained an active maintenance component eliminated or reduced the size of the bizarreness effect in order to more fully test the verbal rehearsal account. The design of the experiment was similar to experiment 3, with the exception that the concurrent task was designed to predominantly employ verbal rather than visual working memory. This is an important because the previous verbal tasks had no active maintenance component. Irrelevant speech required no responses or maintenance from the participants. Articulatory suppression required repetitive responses from the participants but no maintenance. The responses in this task are quite routinized and may not tax verbal rehearsal processes sufficiently. In contrast, this experiment required the participants to actively keep in mind an array of verbal items which change on each trial. During the experiment, participants listened to a string of five letters over the headphones, followed by the sentence and vividness rating, followed by the presentation of a single letter. Participants in the verbal distraction group were told to rehearse the letters while they read the sentences, and then to indicate whether the single letter was in the set, whereas control participants simply ignored the extraneous letters.

Method

Participants

24 participants, from the University of North Carolina, participated in the experiment in exchange for course credit.

Materials and design

The design was a 2 (sentence type: bizarre vs. common) x 2 (distraction: verbal vs. control), with sentence type manipulated within subjects and the verbal working memory task manipulated between subjects. The sentence material was identical to the previous experiments. For the verbal distraction task, 18 five-consonant strings were assembled. Each letter was used at most five times across strings, and letters never repeated within a string. For half of the trials, one of the letters in the string was randomly selected as the prompt letter; on the other half of the trials, the prompt was not a letter from the string. Each letter was presented as a prompt no more than once. Moreover, the type of response for the letter prompt (yes v. no) was counterbalanced across subjects. The consonants were recorded individually and presented over headphones.

Procedure

Each trial began with a beep for half a second, followed by 5 letters presented over the headphones, each for 800 milliseconds, summing to a duration of 4 seconds. The presentation of letters over the headphones was followed by the presentation of a bizarre or a common sentence on the computer screen for 8 seconds. This was followed by the vividness ratings, up to 4 seconds. After entering the rating, the screen went blank and participants heard a letter over the headphones. Each trial lasted up to a total of 20.5 seconds, as in experiment 3.

Participants in the verbal distraction group were told to rehearse the letters silently to themselves as they read the sentence, create a mental image for it, and provide the vividness rating. They were further told to determine whether the final letter was one of the letters that preceded the sentence, pressing “y” for yes, and “n” for no. Participants in the control condition were simply told to listen to the letters, but make no response. Otherwise, the participants carried out the trials as in the verbal distraction group. Participants in both groups were given 2 practice trials, followed by 16 experimental trials.

The distractor and test phases were the same as earlier experiments.

Results

Concurrent WM task

The descriptive statistics for Experiment 4 are presented in Table 4.

First, the access measure was submitted to a 2 (sentence type: bizarre and common) x 2 (experimental condition: verbal vs. control) ANOVA. The bizarreness effect emerged in this experiment ($F(1,22) = 11.41$, $MSe = .024$, $p<.01$), with bizarre sentences (*M* = .41, *SD* = .13) accessed more often than common sentences (*M* = .25, *SD* = .16). The main effect of the experimental condition was also significant ($F(1,22) = 5.64$, $MSe = .017$, $p<.05$), with greater access in the control condition (*M* = .38, *SD* = .15) than the verbal distraction condition (*M* = .29, *SD* = .14). The interaction was not significant (*F*<1). The total noun recall produced the same results as the access measure: a main effect for sentence type (*F*(1,22) = 9.88 , *MSe* = .018, *p* <.01), with more recall of nouns from bizarre (*M* = .32, *SD* = .14) than common sentences (*M* = .20, *SD* = .14); a main effect of experimental condition, (*F*(1,22) = 9.83 , *MSe* = .017, *p* <.01), with reduced recall in the verbal distraction (*M* = .20, *SD* = .11) compared to the control condition (*M* = .32, *SD* = .15); and no interaction (*F*<1).

The mean correct completion proportion for the verbal distractor task was .87 with a standard deviation of .16, significantly above chance (*t*(11)= 7.84, *d* = .37, *p*<.01). Moreover, successful completion of the distractor task was not different in bizarre (*M* =.87 , *SD*=.19) and common (*M* =.88, *SD*=.17) sentence conditions (*t*(11)= .23, *d* = .01, *p*>.8). An analysis of the conditional memory performance for access yielded the same results as the unconditionalized access scores, a main effect of sentence type, (*F*(1,22) = 9.53, MS_e= .022, *p*<.01), a main effect of experimental condition, (*F*(1,22) = 4.74, MS_e= .020, *p*<.05), and no interaction (*F*<1). The conditional total noun recall yielded the identical pattern: a main effect for sentence type (*F*(1,22) = 8.11 , MS_e = .018, *p* <.01), a main effect for the experimental condition (*F*(1,22) = 9.83 , MS_e= .017, *p* <.01), and no interaction (*F*<1).

An additional pooled analysis of Experiment 3 and Experiment 4 to test a non-selective distraction account of the bizarreness effect. A reduction in the bizarreness effect in this merged dataset would show that bizarreness effect might be responsive to just distraction in the slave systems, regardless of whether it is visual or verbal. In order to test this, the datasets of Experiment 3 and 4 were merged, and the visual and verbal distraction conditions were coded as distraction, and the verbal control and visual control conditions were coded as control. Then, access and total noun recall measures were submitted to a 2 (sentence type: bizarre vs. common) x 2 (experimental condition: distraction vs. control) ANOVA. For the access measure, there was a main effect for sentence type ($F(1,54) = 30.07$, $MS_e = .021$, $p < .01$), a main effect for experimental condition ($F(1,54) = 11.07$, $MS_e = .023$, $p < .01$), and no interaction ($F < 1$). The total noun recall followed the identical pattern: a main effect for the sentence type ($F(1,54) = 21.50$, $MS_e = .017$, $p < .01$), a main effect for the experimental condition ($F(1,54) = 17.72$, $MS_e = .023$, $p < .01$) and no interaction ($F < 1$).

Vividness ratings

Vividness ratings were submitted to a 2 (sentence type) by 2 (experimental condition) ANOVA. The only effect that reached significance was the sentence type ($F(1,22) = 32.03$, $MS_e = .479$, $p < .01$). Participants rated the common items ($M = 4.42$, $SD = .54$) as more vivid than the bizarre items ($M = 3.30$, $SD = 1.06$). The main effect of experimental condition was not significant ($F(1,22) = 2.45$, $MS_e = .88$, $p > .05$) nor was the interaction ($F(1,22) = 1.94$, $MS_e = .48$, $p > .05$).

Discussion

In this experiment, we further investigated the verbal rehearsal account of the bizarreness effect. The experiment was done to determine if verbal WM distraction involving active maintenance reduced the BE. It does not, to add to the list of other aspects of verbal WM which do not reduce the effect. The previous distractor tasks, in which participants either passively listened to irrelevant speech or counted from one to four repeatedly, might not have occupied the phonological loop sufficiently to produce any decrement in the performance. However, even when the distractor task was ‘strong’ enough to reduce general memory performance, the size of the bizarreness effect stayed the same, providing additional evidence that the phonological loop is not involved in the emergence of the bizarreness effect. The conditional analyses revealed that even when we excluded the trials in which participants did not pay full attention to the verbal distraction task (in which case they might have paid more attention to the sentences themselves), the size of the bizarreness effect remained the same.

Furthermore, the pooled analyses of Experiment 3 and 4 tested a non-selective distraction account of the bizarreness effect. A possible reduction that might emerge in this analysis could support the idea that the central executive might be involved in the emergence of the bizarreness effect, since the common denominator in both experiments would be related to the taxing of the central executive more in the distractor tasks, as the experiments led to a general decrement in memory performance. Yet, the results did not point to any reduction in the bizarreness effect, providing no proof that a non-selective distraction account disrupted the bizarreness effect.

As in the previous experiments, the analyses pertaining to the vividness ratings revealed that common sentences had higher vividness ratings than bizarre sentences. However, the difference between common and bizarre sentences was not altered by the experimental condition. This provides further proof that the vividness ratings are only moderated by within-subject variables, and the existence of verbal distractors do not change the subjective experience of how participants perceive the vividness of the bizarre or common sentences. The fact that the vividness ratings are not lower as a function of memory performance in control and verbal groups might also show that vividness ratings are unrelated to memory performance.

CHAPTER 6

EXPERIMENT 5

The previous experiments tried to disentangle the processes that contribute to the bizarreness effect by using distractor tasks to selectively impair verbal and visual processes in working memory. In experiments 1-4, I wished to maximize the bizarreness effect in the control condition to have a greater chance of finding interactions with the distractor condition, if one could be found. The prior research indicated that the vividness rating instructions would do so. The results of these experiments indicated that neither type of concurrent working memory task impaired the superior memory for bizarre items. During the proposal, I had proposed follow up experiments that do not explicitly require image formation as the vividness ratings do.

The following experiments were conducted to determine whether visual and verbal distractor tasks would reduce the size of the bizarreness effect when the encoding instructions do not focus on image generation. Two different types of instructions were used, intentional memory instructions (Experiment 5) and plausibility ratings (Experiments 6 and 7). In Experiment 5, participants were instructed to intentionally remember the sentences for a later memory test, and to provide memorability ratings. Intentional memory instructions are thought to encourage semantic encoding: When participants are told that they need to remember the sentences, they are more inclined to use semantic rehearsal processes, which are closer to verbal processes than the visual processes (Burns, 1996; Hirshman et al, 1989). The literature on the use of intentional memory instructions in the bizarreness effect has yielded inconsistent results. Some research has found the bizarreness effect with intentional memory instructions (e.g. Worthen & Roark, 2002), and some research failed to obtain superiority for the bizarre sentences (Burns, 1996). For this reason, it was important to determine if the BE occurred with our materials and procedures, under intentional learning instructions. Consequently, only the control condition was carried out initially.

As an additional motivation, I am interested in the beliefs of participants in terms of the memorability of the sentences. The main question is whether participants distinguish between common and bizarre sentences in terms of their memorability and whether this initial judgment at encoding is correlated with the accuracy of the sentences at retrieval.

Method

Participants

22 participants, from the University of North Carolina, participated in the experiment in exchange for course credit.

Design and Procedure

Sentence type (bizarre vs. common) was manipulated within subjects. The study phase was the same as the verbal control condition of Experiment 1, with the exception of changes in the instructions and the type of rating. Participants in this experiment were told that they would be presented with sentences which they should try to remember for a later (unspecified) test. They were also asked to rate each sentence for its memorability on a scale from 1 to 5, with 1 indicating that the sentence is not memorable at all and 5 indicating that the sentence is very memorable. Each trial started with a beep. Two seconds later, participants were presented with the sentence on the screen for 8 seconds. This was followed by a screen on which the words "Memorability (1-5):" were displayed. The screen was displayed until the participants entered a number, up to 4 seconds. Each trial lasted up to 14.5 seconds. The distractor and testing phases were identical to the previous experiments.

Results

Memory Accuracy

The descriptives for access, total noun recall and memorability ratings are presented in Table.5.

There was no significant difference in access to bizarre and common sentences ($t(21)=.56$, *ns*). Likewise, the total noun recall was not significantly different for bizarre and common sentences ($t(21)=1.54$, *ns*).

Memorability ratings

Memorability ratings were significantly lower for bizarre than common sentences ($t(21) = 2.39$, $p <.05$). Moreover, participants' memorability ratings were averaged over accuracy and sentence type and were submitted to a paired samples t-test separately for bizarre and common sentences. The result revealed that the memorability ratings did not predict accuracy for bizarre sentences ($t(21) = .730$, *ns*) or common sentences ($t(21) = 1.55$, *ns*).

Discussion

The intentional memory instructions eliminated the bizarreness effect for both recall measures. Because the bizarreness effect did not emerge in the control condition, the concurrent working memory task conditions were not run. As the instructions change from incidental to intentional instructions, the memorial advantage for bizarre items over common items was wiped out completely. When participants are given incidental instructions, they do not take any initiation for what they will memorize; thus, bizarre items might attract more attention naturally, leading to more elaborative rehearsal than common items. Yet, when they are told that they will have to remember the sentences, they might exert more control over their learning processes, which leads to equivalent memory for bizarre and common items.

The memorability ratings have shown that although the participants predicted that their memory would be worse for bizarre items, memory performance for both types of sentences was similar. Considering that the memorability ratings show same direction tendencies as the vividness ratings, the memorability ratings might be related to some subjective encoding process, such as the ease of encoding (Begg, Duft, Lalonde, Melnick & Sanvito, 1989; Hertzog, Dunlosky, Robinson & Kidder, 2003). To explain, when the common items are presented to participants, they can comprehend the sentences very easily and process them more fluently, which gives them the feeling that they will be able to remember the items more easily. For bizarre items, participants have more difficulty in comprehending these items; thus, this difficulty in comprehension of the items leads to a subjective feeling that they will not be able to remember those items, leading to lower memorability ratings. There is no direct evidence that might suggest that vividness ratings and memorability ratings are driven by the same system. However, self-paced studies of bizarreness in which the participants move on the next item after they create an image show that the formation of the image takes longer for bizarre sentences (e.g.

McDaniel & Einstein, 1986; Robinson-Riegler & McDaniel, 1994). Research from metamemory literature show that the longer it takes to form an interactive image for word pairs, the less memorable participants rate the item through both judgments of learning (similar to memorability)and quality of encoding (similar to vividness ratings). However, the memory performance at final recall is uncorrelated with the time that it takes the participants to form the images (Hertzog et al., 2003). Thus, participants' memorability ratings and vividness ratings might be driven by their successful encoding of the image. Yet, this fluency of formation does not necessarily help their retrieval of information, as shown by non-significant correlations between memorability and accuracy.

CHAPTER 7

EXPERIMENT 6

Experiments 3 and 4 showed that the size of the bizarreness effect was not reduced by either verbal or visual concurrent working memory tasks, even though the distractors reduced overall memory performance. However, the use of overt imagery instructions in these experiments may have induced participants to generate images in all conditions, in spite of the distractor task. That is, the directive to use imagery may have worked at cross-purposes with the visual distraction task, such that participants successfully used imagery despite the distractor task. For example, according to the multi-imagery process theory, bizarreness effect occurs because participants make multiple transformations on the images. If they are not given overt instructions to construct an image, the concurrent visual working memory distractors might become more effective, leading to a reduced bizarreness effect.

In Experiment 6, memorability ratings were replaced with plausibility ratings, in which participants judge whether the sentences describe events that can happen in real life. Even though many bizarreness studies use vividness ratings, previous research has successfully uncovered bizarreness effects using other types of ratings. Some studies have used "bizarreness" ratings (e.g. Kroll & Tu, 1988; McDaniel et al, 1995, 1999) and others have used pleasantness ratings (e.g. Worthen, 1997). The results suggest that although the bizarreness effect might be somewhat diminished, that the effect is still found, especially in the sentence access measure. Experiment 6 uses plausibility ratings, because we feared that bizarreness ratings might induce imagery processing to a similar degree as vividness ratings. The distractor task was the visual span task from Experiment 3, which successfully reduced overall memory performance. The visual-imagery hypothesis predicts that the bizarreness effect should be reduced under visual distraction.

Method

Participants

Thirty-two participants participated in this experiment in exchange for course credit or monetary compensation.

Materials and Procedure

The design was a 2 (sentence type: bizarre and common) x 2 (experimental condition: visual distraction vs. control). The methods were identical to Experiment 3, with a few modifications. The rating screen asked for “Plausibility (1-5)” instead of “Vividness (1-5)”. Participants were asked to think about the plausibility of the sentences that they read rather than generating images. They were instructed to rate the plausibility with a number from 1 to 5. A rating of 1 indicated that the event mentioned in the sentence could never happen in real life, and a rating of 5 indicated that it could definitely happen in real life.

Results

Concurrent WM Tasks

The descriptive statistics for the visual and control groups are presented in Table 6.

Both the unconditional and the conditional measures of access and total noun recall were submitted to ANOVAs with the sentence type and experimental condition (visual vs. control) as factors. The unconditional access measure revealed similar results to previous experiments: There was a main effect for the sentence type ($F(1,30) = 15.60$, $MS_e = .017$, $p<.01$), with bizarre sentences accessed more often than common sentences. There was a main effect for the experimental condition ($F(1,30) = 9.59$, $MS_e = .028$, $p<.01$), with more items accessed in the control condition, as compared to the visual working memory task condition. The interaction between the two factors was not significant ($F<1$). The total noun recall revealed a slightly less informative pattern than the access measure. The unconditional total noun recall revealed that there was a marginally significant main effect of sentence type ($F(1,30) = 3.54$, $MS_e = .012$, $p=.07$). The total noun recall performance for the bizarre sentences ($M = .30$, $SD = .13$) was not significantly higher than the recall for the common sentences ($M = .25$, $SD = .17$). The main effect for the experimental condition was significant ($F(1,30) = 171.90$, $MS_e = .029$, $p<.01$), with the control group ($M = .31$, $SD = .14$) performing better than the experimental group ($M = .22$, $SD = .14$). The interaction did not reach significance ($F<1$).

The mean correct completion proportion for the verbal distractor task was .75, significantly above chance ($t(15)= 9.97$, $d = .24$, $p<.01$). Moreover, participants' successful completion of the distractor task was not different for bizarre ($M =.75$, $SD=.12$) and common ($M =.75$, $SD=.14$) sentences ($t(16)= .15$, $d = .0$, $p>.8$). The conditional access measure followed the exact same pattern as the unconditional scores, with a significant main effect for the sentence type ($F(1,30) = 10.89$, $MS_e = .023$, $p<.01$), a significant main effect for the experimental condition type ($F(1,30) = 5.31$, $MS_e = .028$, $p<.05$) and no interaction ($F<1$). The conditional total noun recall also revealed the same pattern as the unconditional total noun recall, a non-significant advantage for the bizarre sentences (*M* = .31, *SD* = .14) as compared to the common sentences (*M* = .27, *SD* = .16), ($F(1,30) = 2.54$, $MS_e = .013$, $p<.10$), a significant advantage in the control group (*M* = .34, *SD* = .14) as compared to the visual group (*M* = .24, *SD* = .14) ($F(1,30) = 6.47$, $MS_e = .027$), $p<.01$) , and no interaction ($F<1$).

In order to argue against the role of visual encoding processes in the emergence of the bizarreness effect, it is necessary to conduct a pooled analysis of all experiments that test the visual account (Exp1-3 and 6). In order to realize this, all visual and visual control conditions of previous experiments were merged and submitted to a 2 (sentence type: bizarre vs common) x 2 (experimental condition: visual distraction vs control) ANOVA. The pooled results for the access measure yielded a main effect for the sentence type ($F(1,146) = 95.88$, $MS_e = .017$, $p<.01$), with bizarre sentences (*M* =.44, *SD*=.17) accessed more often than common sentences (*M* =.28, *SD*=.15). There was also a main effect for the experimental condition ($F(1,146) = 10.02$, $MS_e = .033$, $p<.01$), with more items accessed in the control condition (*M* =.39, *SD*=.15), as compared to the visual working memory task condition (*M* =.33, *SD*=.17). The interaction between the two factors was not significant ($F(1,146) = 2.00$, $MS_e = .019$, $p>.10$). The total noun recall measure revealed the same pattern: a main effect for sentence type ($F(1,146) = 58.104$, $MS_e = .015$, $p<.01$), a main effect for experimental condition ($F(1,146) = 8.71$, $MS_e = .031$, $p<.01$) and no interaction ($F(1,146) = 2.83$, $MS_e = .015$, $p>.05$).

Plausibility ratings

Plausibility ratings were calculated separately for each sentence type and experimental condition and are shown on Table 7. Then, they were submitted to a 2 (sentence type) by 2 (experimental condition) ANOVA. The results indicate that common items were rated as more plausible than bizarre items ($F(1,32) = 185.96$, $MS_e = .801$, $p<.01$). The main effect of the experimental condition was not significant ($F<1$), nor was the interaction ($F(1,32) = 2.75$, $MS_e = .801$, $p>.05$).

Discussion

In this experiment, we investigated the role of visual distractors paired with encoding instructions that did not explicitly focus on imagery. Despite this change, the results were generally consistent with the findings of Experiment 3, in which the same visual distractor task was paired with explicit imagery instructions. First, there was a significant access advantage for bizarre sentences. In addition, the visual distraction condition reduced access to sentences in general. However, the size of the bizarreness effect was not reduced by visual distraction. Quite the opposite, numerically (if not significantly) the bizarreness effect was larger in the distraction than control condition. The results stayed the same for the conditional access scores, indicating that even when we only take into account those trials that participants were clearly paying attention to the concurrent working memory task, the access to bizarre sentences was higher than common sentences, and the size of the BE was unaffected by distraction. Thus, at the end of 4 experiments, including all visual tasks (DVN, spatial tapping and visual pattern matrices), the failure to find any decrement in the bizarreness effect when participants are presented with visual concurrent tasks, we could conclude that the imaginal encoding hypothesis does not account for the superior memory for the bizarre sentences.

The total noun measure was somewhat less informative: Participants recalled equivalent number of nouns from bizarre and common contexts. The experimental condition lowered total noun recall in general, but this effect was not moderated by the sentence type, demonstrating that the failure to obtain the bizarreness in the total noun recall measure was not relevant to the concurrent visual memory task. Prior studies have also shown that ratings other than vividness do not produce the bizarreness effect in the total noun recall measure, or the effect is sometimes smaller (e.g. Worthen & Roark, 2002; Kroll & Tu, 1988; Worthen, 1997), so it is not quite surprising that total noun recall did not yield a significant bizarreness effect.

In a cross experimental comparison of the instructions (Experiment 3 vs. 6), the results show that when participants are asked to use the vividness ratings, the bizarreness effect is obtained with both measures of access and total noun recall. However, when they are asked to judge the sentences for their plausibility, the bizarreness effect is reduced in the total noun recall measure. Thus, rather than the visual working memory task, there seems to be something special about the instructions given to participants. An informal cross-experimental comparison of the two experiments shows that the only value that changes for the total noun recall is that participants are able to recall more nouns from the common sentences when they make plausibility ratings as compared to vividness ratings. Even though these experiments were conducted at different times of year, plausibility ratings seem to help participants increase their performance for common items. This might have to do with the involvement of semantic verbal processes in plausibility ratings. With vividness ratings, participants' attention are attracted to bizarre items as compared to common sentences, but when participants are asked to judge sentences for their plausibility, they might be making more semantic verbal elaborations for common sentences, and the difference between common and bizarre sentences might decrease, so plausibility differentially increases the semantic elaboration of common items, leading to a decreased bizarreness effect on the total noun recall. Therefore, in such a case, one would expect that, if the participants are doing a verbal task while they are judging the plausibility of the sentences, the bizarreness effect should re-emerge in the total noun recall measure, as the participants would not able to pay attention to the common sentences as much. As a consequence, it is important to see whether using plausibility ratings paired with verbal distraction tasks leads to the re-emergence of the bizarreness effect in the total noun recall, as compared to a control condition. This possibility will be tested in Experiment 7.

A pooled analysis was also conducted for all the experiments that had a visual distraction condition, in order to argue against the lack of power in each experiment. The results clearly showed that even though the visual distraction tasks caused a general decrement in the memory performance, there was no evidence that visual distractors reduced the size of the bizarreness effect. This analysis clearly argues against any visual explanations of the bizarreness effect. A more elaborate discussion of this finding can be found in the general discussion section.

CHAPTER 8

EXPERIMENT 7

The previous experiment showed that the use of visual distractors along with the plausibility ratings leads to a general decrement in memory performance. Moreover, plausibility ratings led to a reduced bizarreness effect possibly indicating that judgments of plausibility lead to greater verbal elaboration especially for the common sentences. If so, the use of plausibility instructions along with a verbal distractor task may lead to the re-emergence of the bizarreness effect. Up to now, it was hypothesized that the verbal distractors should lead to a reduced bizarreness effect, if the effect is moderated by verbal rehearsal. In this experiment, the exact opposite view is entertained, because Experiment 6 showed that plausibility judgments might be leading to increased elaboration for the common items, reducing the size of the effect, because common item are rehearsed more. Thus, one would expect that if participants are presented with verbal distractors, this might again decrease the selective rehearsal of the common items, leading to a re-emergence of the bizarreness effect in total recall measure. This experiment tries to assess this account by pairing the plausibility instructions along with verbal distractor tasks.

Another reason for conducting this experiment is to complete the evaluation of the slave working memory processes in the emergence of the bizarreness effect. In all the prior experiments, there was no evidence that either visual or verbal distractors reduced the BE. As a complement to Experiment 6, Experiment 7 investigated the effect of verbal distraction along with plausibility instructions.

Method

Participants

24 participants participated in this experiment in exchange for course credit or monetary compensation.

Materials and Procedure

The design was a 2 (sentence type: bizarre and common) x 2 (experimental condition: verbal distraction vs. control). The methods were identical to Experiment 4, with one change. Participants were asked to think about the plausibility of the sentences rather than generating images, and rate the plausibility as in Experiment 6.

Results

Concurrent WM tasks

The descriptive statistics for the verbal and the control groups are displayed in Table 8.

The proportion of access was submitted to 2 (sentence type) x 2 (experimental condition) ANOVA, revealing a main effect of sentence type, with bizarre sentences accessed more often than common sentences ($F(1,22) = 4.41$, $MS_e = .027$, $p<.05$). All other effects were nonsignificant (Fs<1). The analyses for the total noun recall revealed no significant effects. To explain, the bizarreness effect was not obtained when the plausibility ratings were used ($F(1,22) = 2.72$, $MS_e = .022$, $p > .05$). Moreover, the verbal task failed to impair the memory performance ($F<1$).

The mean correct completion proportion for the verbal distractor task was .91 with a standard deviation of .08, significantly above chance ($t(11)= 17.13$, $d = .41$, $p<.01$). A paired samples t-test showed that participants' successful completion of the distractor task was marginally different in bizarre (*M* =.86, *SD*=.11) and common (*M* =.95, *SD*=.08) sentence conditions (*t*(11)= 2.15, *d* = .08, *p*=.054). The conditional access measure revealed the identical pattern as unconditional scores, with a main effect for the sentence type ($F(1,23) = 4.96$, MS_e= 1.36, p<.05), and all the other effects insignificant (*F*s<1). Thus, even though the bizarreness effect for the access measure was obtained in this condition, the presence of a verbal task paired with the plausibility ratings did not reduce performance in general. Moreover, the bizarreness effect remained intact in the verbal distraction condition. The conditional scores for total noun recall were the similar to the unconditional score analyses: all the effects failed to reach significance (F<1), even though the sentence type was marginally significant in the expected direction ($F(1,22) = 3.03$, MS_e= .020, p=.06).

In order to argue against the role of verbal encoding processes in the emergence of the bizarreness effect, it is necessary to conduct a pooled analysis of all experiments that test the verbal account (Exp1, 2, 4 and 7). In order to realize this, all verbal and control conditions of previous experiments were merged and submitted to a 2 (sentence type: bizarre vs. common) x 2 (experimental condition: verbal distraction vs. control) ANOVA. The pooled results for the access measure suggested a main effect for the sentence type ($F(1,130) = 68.58$, MS_e = .022, p<.01), with bizarre sentences (*M* =.45, *SD*=.16) accessed more often than common sentences (*M* =.30, *SD*=.16). There was no main effect for the experimental condition ($F(1,130) = 2.21$, MS_e = .019, p<.01) and no interaction (*F*<1). The total noun recall measure revealed the same pattern: a main effect for sentence type ($F(1,130) = 45.60$, MS_e = .017, p<.01), no main effect for experimental condition ($F(1,130) = 1.48$, MS_e = .028, p>.10) and no interaction (*F<1*).

Plausibility Ratings

Plausibility ratings were submitted to a 2 (sentence type) by 2 (experimental condition) ANOVA. The results indicate that common items were rated as more plausible than bizarre items ($F(1,22) = 2124.89$, $MS_e = .069$, $p<.01$). The main effect of the experimental condition was not significant ($F(1,22) = 1.28$, $MS_e = .085$, $p>.20$), nor was the interaction ($F<1$).

Discussion

As in Experiment 6, Experiment 7 produced the bizarreness effect with the sentence recall measure, but not with the total noun recall measure, although the latter measure exhibited a trend for the effect. If anything, the size of the bizarreness is numerically higher in the verbal distraction condition than in the control condition. The finding that the size of the bizarreness effect was not reduced through the use of verbal distractors along with plausibility ratings produce further evidence that the phonological loop is not involved in the bizarreness effect, in agreement with the findings of Experiments 1, 2 and 4.

A surprising finding was that there was no main effect for the verbal distractor in reducing the general memory performance. This finding might have actually been a consequence of the participant pool, since the experiments for the verbal distraction group were run before the control group. If the control group of the experiment 7 is compared to the control group of experiment 6 (in which the timing and the instructions are the same), the experiment 6 has higher scores in general for the control group. This might actually provide further proof for the change in the quality of the participants in the pool.

In order to argue that the experiments that tested the role of verbal encoding processes had sufficient power, a pooled analysis of all experiments that test the verbal account was conducted. The results revealed that the size of the bizarreness effect was not reduced by the verbal tasks. This argues against the role of verbal elaboration in the emergence of the bizarreness effect. The general discussion section will elaborate on this finding further, and indicate what might be important in the emergence of the bizarreness effect.

CHAPTER 9

GENERAL DISCUSSION

Evaluation of the Role of Verbal and Visual Encoding Processes in the Emergence of the Bizarreness Effect and Possible Future Directions

The goal of the current project was to determine the reasons for the emergence of the bizarreness effect. There have been two categories of accounts for this effect. One of the accounts focuses on visual imagery processes. There are different variations of the hypothesis, but the general idea is that when participants generate images for both common and bizarre items, the images for the bizarre items are visually more distinctive (McDaniel, 1986), or require more transformations in their creation (Wollen & Margres, 1987), which leads to better encoding for these items and, therefore, greater recall. The other category of accounts depends on verbal rehearsal processes. One version of the verbal rehearsal account claims that when participants encounter sentences that are bizarre, these sentences are harder to understand; thus, they make more verbal elaborations to understand the sentence, leading to increased memory traces for bizarre sentences (Imai & Richman, 1991). Another version of the verbal rehearsal account focuses on the unusualness of the bizarre items. According to this account, in lists consisting of bizarre and common sentences, bizarre information lures attention, leading to increased elaboration for those items at the expense of decreased elaboration for the common items (McDaniel & Bugg, 2008). Extant research typically used encoding instructions to encourage one type of processing or the other, in order to understand the effect. But as reviewed in the introduction, most of these results do not unambiguously favor one type of account or the other. In this project, the main goal was to discourage one of the processes during sentence encoding through the use of a concurrent working memory task known to selectively impair either verbal or visual WM, to investigate whether the size of the bizarreness effect decreases in any of the distraction conditions.

The role of the visual processes in the emergence of the bizarreness effect was investigated in Experiments 1-3 and 6. Experiment 1 used dynamic visual noise, Experiment 2 used spatial tapping, and Experiment 3 used the visual patterns task. None of the distraction conditions reduced the size of the bizarreness effect, providing no evidence that the bizarreness effect relies on visual WM processes. It is important to emphasize that all these tasks are well known to selectively disrupt visual working memory processes. DVN disrupts visual processing, but leaves verbal rehearsal intact (e.g. Quinn & McConnell, 1996, 2006). Similarly, spatial tapping leads to decrements in visuospatial processing, but does not interfere with free recall of words (Salway & Logie, 1995). Visual patterns task leads to decrements in processing of other visual stimuli (Della Sala et al, 1999). One might argue that the initial tasks may not be strong enough to disrupt long-term memory at all in distraction conditions. However, even when the task was strong enough to disrupt overall memory performance (Experiments 3 and 6), the size of the bizarreness effects remained undiminished.

The role of the verbal processes in the emergence of the bizarreness effect was investigated in Experiments 1, 2, 4 and 7. Experiment 1 used irrelevant speech, Experiment 2 used articulatory suppression and Experiment 4 used the letter span task.

All these tasks are known to depress verbal working memory. For example, backward speech leads to decrements in free recall of word lists, but not visual information (Oswald, Tremblay & Jones, 2000). Articulatory suppression leads to decrements when paired with other verbal tasks, but not with visual tasks (Cocchini et al, 2002). Letter spans are depressed when used along with other verbal tasks (Logie, Zucco &Baddeley, 1990). Again, even though these tasks are well known to load on the phonological loop, none of these concurrent verbal WM tasks reduced the size of the bizarreness effect, indicating that the phonological loop does not contribute to the emergence of the bizarreness effect. Moreover, even those tasks which caused some decrement in overall memory performance (Experiment 4) did not reduce the bizarreness effect.

Experiments 1-4 used explicit imagery instructions (vividness ratings) to make sure that we obtained the greatest difference between bizarre and common items in the control condition to maximize the chance of observing any interactions with the experimental conditions. Experiments 5-7 used encoding instructions that do not explicitly focus on visual imagery in order to see if the sensitivity of the bizarreness effect to distraction changed. The reasoning behind this was to determine if the use of imagery instructions renders the bizarreness effect resistant to any type of distraction. Experiment 5 used intentional instructions which eliminated the bizarreness effect in the control condition. Thus, intentional instructions were not used any further. Experiment 6 and Experiment 7 used plausibility ratings, paired with visual and verbal distractor tasks. Earlier research made use of ratings such as bizarreness and pleasantness in addition to vividness and memorability. However, when participants are asked to judge how bizarre or pleasant a sentence is, they might have a tendency to form images. Thus, I chose to use plausibility ratings, where participants are asked whether the sentence would be plausible in real life. Thus, they need to make reference to their world knowledge and long term memory structures to indicate the level of plausibility. This manipulation does not necessarily rule out imagery processing; however, using these instructions might lessen the reliance on imagery processing during encoding and focus more on semantic processes. First, the results indicated that the bizarreness effect is obtained in the control group with instructions that do not explicitly require imagery. However, the superior performance for bizarre sentences is not as pronounced with plausibility ratings as it is with vividness ratings: the sentence access measure produced a significant effect, but total noun recall produced only a marginally significant effect in both Experiment 6 and 7. More centrally, the distraction conditions did not reduce the size of the bizarreness effect in either of the experiments. Thus, even when participants were not instructed to create mental images, which might have overridden any visual or verbal distraction processes, the results did not indicate any role for the phonological loop or the visuospatial sketchpad in producing the effect.

Perhaps the most important contribution of this study is that the results argue against the role of imagery in the bizarreness effect. Starting from ancient Greek memory traditions to modern self-help books, from general folk knowledge to modern psychological research, the bizarreness effect has often been considered as a consequence of imagery. The ancient Greek tradition instructed people to form strange images in order to memorize speeches (Yates, 1966). Today's self-help books claim that in order to memorize names, lists or numbers, people should try to generate images that are unusual or bizarre (Buzan, 1991; Lorayne, 2010; Lucas & Lorayne, 1973; Trudeau, 1997) . In daily life, students try to learn information by coming up with weird images of the events. In modern psychological research, the bizarreness effect has been accepted to rely on imagery to the extent that it is sometimes called the bizarre imagery effect. Thus, many different sources have considered the bizarreness effect to be a consequence of imagery. Yet, the results of this project have consistently shown that the bizarreness effect does not depend on visuospatial processing. The pooled analyses of the experiments that have tested the visual account has shown that the effect is even more apparent in the visual distraction condition (numerically if not significantly) than the control condition. Looking at the visual imagery accounts, McDaniel and Einstein (1986) have argued for the visual distinctiveness of the bizarre sentences as compared to common sentences. According to this account, vividness ratings lead participants to engage in imaginal processing, making the bizarre items processed extensively through the visual system. In evidence for this account, it has been found that the bizarreness effect is obtained consistently with the vividness instructions (e.g. Einstein & McDaniel, 1987), but not consistently with other types of orienting tasks (e.g. Burns, 1996). Another account, similarly, indicates that when participants are asked to create images for bizarre sentences, the created visual images have to go through more transformations, because the visual images are not schematic in nature. To explain, in order to create an image for the sentence *The dog rode the bicycle down the street*, the image of the dog needs to be changed more (e.g. dog rotated into a sitting position, have front legs extended to the handle bar, hind legs extended to reach the pedals), leading to more effort on the part of the participants, thus more memory traces for the visual transformations (Wollen & Margres, 1987). In evidence for this, some studies have found that the more visual transformations that the image has to go through, the better the recall is. Yet, there is also much evidence against the visual-imagery account. For example, the effect is obtained with different orienting tasks that focus on verbal rehearsal (i.e. Worthen, 1997; Worthen & Roark, 2002). Secondly, there is no relationship between imaging ability and the bizarreness effect (Anderson & Buyer, 1994). One would expect that bizarreness effect would be more pronounced for those participants with higher imaging abilities, if it depended on visual processing. When we look at the results in the current project in which visual working memory distractor tasks are employed, the results do not provide any evidence for the involvement of visual working memory, even though the distractor tasks are well known to reduce or interfere with visual processing.

The role of verbal processing is a little more complicated, as sometimes the instructions that focus more on the verbal rehearsal (plausibility & intentional instructions) reduce the effect, but again the pooled analysis of the results do not suggest any decrement in the emergence of the bizarreness effect in the verbal distraction condition.

This project tested the effect of verbal and visual working memory processes with different concurrent working memory tasks and different instructions. A systematic examination of distraction conditions and instructions showed that selectively impairing the phonological loop or the visuospatial sketchpad did not moderate the size of the bizarreness effect. If these two systems are not responsible for producing the effect, what mechanism might actually be responsible? One possibility is the central executive. As explained earlier, the working memory system consists of three different components. Phonological loop and visuospatial sketchpad are the two slaves systems that deal with storage and manipulation of verbal and visual information, and work independently of each other. On the other hand, the central executive is an amodal supervisory system, which controls the flow of information from and to its slave systems. It is hypothesized to bind information coming from a number of different sources, to coordinate the slave systems, and to selectively activate and inhibit information. Thus, it is quite possible that the bizarreness effect goes beyond the effect of the slave systems and is more appropriately attributed to the central executive. Baddeley (2003) defined the role of central executive in the working memory in two different components: The first part is the control of behavior by habit patterns and schemas, implicitly guided by cues provided by the environment. This component is more automatic and it does not require too many resources. The second comprised an attentionally limited controller, the supervisory activating system (SAS) that can intervene when routine control is not sufficient. The bizarreness effect might actually be moderated by one of these subcomponents or the interaction between these two subcomponents of the central executive system.

There are several possible ways that the central executive might be involved in the bizarreness effect. One possibility has to do with the automatic orientation of attention to distinctive information. To explain, one of the ways that the bizarreness effect can be thought of is in terms of its distinctiveness relative to general world knowledge and long term memory structures (Schmidt, 1999). When the target item is unusual with respect to the general world knowledge of the participants, it attracts attention. For example, when participants read sentences in a mixed list of common and bizarre sentences, they might have a tendency to use the more habitual system for common sentences, but bizarre sentences might have a tendency to attract attention automatically, leading to more extensive traces for bizarre sentences. Thus, simple material like word lists might be disrupted through the use of concurrent verbal tasks. Similarly, when participants are shown simple pictures, visual concurrent working memory tasks might reduce subsequent memory. However, lists that involve sentential material requiring more effort for the comprehension, like the bizarreness effect, the capture of our attention by the “unusual” is quite involuntary, and that orienting of attention and comprehension is sufficient to obtain the effect.

There is a phenomenon similar to the bizarreness effect, called meaning-after-effort effect, which might be indicative of how simple comprehension of sentences might enhance memory (Auble & Franks, 1978; Auble, Franks, & Soraci, 1979; Zaromb & Roediger, 2009). In this phenomenon, participants are presented with ambiguous sentences, such as “the clothes were ruined because the sign vanished”, which are presented with or without their disambiguating cues (i.e. wet paint). The results indicate that when participants are asked to remember the sentences later, the free recall for sentences with disambiguating cues are significantly better as compared to sentences with no disambiguating cues. Auble et al has attributed this memory advantage to the sudden transition from non-comprehension to comprehension, when an initially ambiguous stimulus is suddenly comprehended, as long as the stimulus is initially ambiguous and then comprehension is ultimately achieved. Zaromb and Roediger (2009) have categorized the meaning after effort effect as an effect associated with the item-order account. According to this explanation, the transition from non-comprehension to comprehension makes the comprehended items more distinctive. Thus, just the mere comprehension of the bizarre sentences and the automatic guiding of attention to those items might be responsible for the emergence of the bizarreness effect. One finding that might elucidate this similarity between bizarreness and the meaning-after-effort effect is that the experiments conducted with complex bizarre sentences and anomalous sentences. These experiments have revealed that in short periods of time, the bizarreness effect is not obtained; however, if sufficient time is provided to participants, the bizarreness effect emerges again. The extra time might be responsible for the transition from non-comprehension to comprehension of complex or anomalous bizarre sentences (Imai & Richman, 1991; Richman, 1994). Thus, the increased time might moderate the effect by enabling comprehension, and this transition itself might lead to increased memory traces for bizarre sentences, beyond any visual or verbal rehearsal.

Another possibility in the emergence of the bizarreness effect might be relevant to the other subcomponent of the central executive, which is involved in more controlled processes. The idea is that the process by which the bizarre items enhance memory is not automatic, and when participants encounter bizarre sentences, they start paying attention to the sentences, which leads to increased controlled processes that are associated with the verbal rehearsal. However, the results of the experiments presented here argue against the account for controlled verbal rehearsal for bizarre sentences, as we find that the size of the bizarreness effect is not diminished in the verbal distraction condition.

For the reasons indicated above, the central executive appears as a viable candidate for producing the bizarreness effect. Future studies should investigate the role of central executive processes through the use of distractor tasks that disrupt central executive processes. One way this could be realized is to use different tasks that interfere with different functions of the central executive. To be clearer, it is important to see by which process the central executive might guide the bizarreness effect. The central executive is hypothesized to be responsible for several functions such as binding information from different sources, selective attention, updating and control. Thus, if a manipulation could be found in order to occupy the automatic or the controlled processes of the central executive, this might lead to a reduced bizarreness effect. For example, random number generation, in which participants are asked to generate random numbers between two boundary numbers in a limited time, has been shown to put demands on the central executive (Barnard, Scott & May, 2001). In the bizarreness effect experiments, participants could be asked shown two numbers (e.g. 202-598) before they start reading the sentences and asked to generate four random numbers between these two numbers throughout the trial (one number per every other second that is signaled by a metronome over the headphones).

It is also important to note that the two subsystems of the central executive sound suspiciously similar to the automatic and controlled processes of attention (e.g. Engle, 2002). The attentional approach to the bizarreness effect is not a new approach. According to this view, bizarre items use up more of the cognitive resources during processing: They attract more attention and are elaborated more, leading to an advantage for bizarre items at the expense of decreased performance for common items. Many experimenters have included an attentional component in their explanation of the bizarreness effect (Cox & Wollen, 1981, McDaniel & Bugg, 2008, Wollen & Cox. 1981, Wollen & Margres, 1987). One of the ways researchers have tried to test how much attention is allotted to bizarre sentences was to manipulate the time provided to participants. The general reasoning was that fast processing of bizarre items should eliminate the increased attention allocated to bizarre items, leading to similar performance for common and bizarre items. For example, in a study by Waddill and McDaniel (1998), half of the participants were exposed to a mixed list of bizarre and common sentences for five seconds, whereas the control group received them in a slower pace of 18 seconds. The results revealed a main effect for sentence type and processing time, with an advantage for recall performance for the sentences processed for 18 seconds and an advantage for bizarre sentences, but there was no significant interaction between the processing time and the sentence type, contrary to the predictions. The concern in operationalizing attention as time allocation is that participants may continue to process bizarre items as they are presented with common items (selectively displaced rehearsal), which cannot be measured through the time allocation manipulation.

Another way that the attentional approach was tested was to see how intensively the resources are being used up while processing common and bizarre sentences. By this logic, Worthen, Garcia-Rivas, Green and Vidos (2000) operationalized resource allocation as the reaction time to a secondary task while trying to form images during the encoding of common and bizarre sentences. In a mixed-subjects design with the sentence type as the within subjects variable and type of secondary task as the between subjects variable, participants had to form images of common and bizarre sentences. In some trials, a message asking participants to "press any key" would appear either at 2 seconds or 2.5 seconds after the presentation of the sentences. The reaction time to this task was used to measure attentional allocation. Although the bizarreness effect was produced on the later memory test, response to the secondary task did not reveal any significant effect of sentence type (Experiment 1). The performance for the secondary task was not moderated by predominantly common lists (Experiment 2), or more immediate key press response (Experiment 3, Experiment 4). Worthen et al (2000) explained this finding by rejecting the role of the attentional resources in the emergence of the bizarreness effect. However, it is important to address how the secondary task in this experiment might have failed to interfere with the bizarreness effects. In terms of the resources that are used, secondary task might not have required a serious allocation of cognitive resources for completing the primary task, especially considering that the key-press response is an automatic response, leading to null effects. Thus, up to now, no rigorous studies of attention have been conducted to investigate its role in the emergence of the bizarreness effect. Future studies should focus on more rigorous attentional tasks that consume more resources that require participants to respond more continuously.

Evaluation of Vividness Ratings, Memorability Ratings and Accuracy

A secondary aspect of the current experiments was the relationship between bizarreness and various ratings for each sentence. A common finding that was replicated in the current research is that bizarre items are rated as less vivid than common items. There are two views regarding this relationship. Wollen and Margres (1987) argued that the lower vividness ratings were a consequence of the extra visual transformations required for creating bizarre images. The more the image differed from schematic images, the less vividly the participant could imagine it. Marshall, Nau and Chandler (1979, 1980) found that the more visual transformations that the participants had to make, the better the recall was, but they did not look at the relationship between vividness and accuracy.

Thus, one would expect that if the vividness ratings are negatively correlated with the number of transformations, and if the number of transformations is positively correlated with recall for both common and bizarre items, then less vivid ratings would predict greater memory. In order to see whether this was actually the case, for the first 4 experiments, the correlations between vividness, accuracy, experimental condition and bizarreness were analyzed (For a detailed explanation of the analyses, please refer to general description for the vividness-accuracy analyses in Appendix. B.) These analyses suggested that the relationship between the vividness ratings and the memory accuracy was quite inconsistent across experiments. For example, in Experiment 1 lower vividness scores were correlated with higher memory accuracy for bizarre sentences, but the reversed pattern was obtained for the common sentences. In Experiment 2, vividness ratings for bizarre sentences did not correlate with memory accuracy, but higher vividness scores for common sentences were correlated with more accurate memory. In Experiment 3, lower vividness ratings were correlated with more accurate memory for bizarre sentences, but there was no correlation between vividness and accuracy for the common sentences. Considering that the results are so inconsistent, it is difficult to argue that more transformations actually lead to better memory. Moreover, considering that we have no evidence for the role of visual working memory in the bizarreness effect, the vividness ratings might be related to some encoding condition that has no bearing over the recall.

Another more viable possibility is that vividness ratings are related to the subjective ease of encoding. One of the factors determining whether an item is rated as more vivid is speed of encoding. For example, Hertzog et al (2003) used paired-associate learning to investigate the hypothesis that the speed of generating an interactive image determined the vividness rating for the image (which they called, Quality of encoding). In their experiment, participants were given word pairs and asked to form interactive images and press a key as soon as they formed the image. This was followed either by quality of encoding ratings (QUE), in which participants indicated how good the image was (Experiment 1), or by judgments of learning (JOL), in which participants indicated how likely they were to remember the pair on a later memory test (Experiment 2). The results showed that the latency of the key press indicating successful image formation was negatively related to both QUEs and JOLs. On the other hand, actual memory performance was unrelated to the time to form an image. Self-paced studies of the bizarreness effect usually show that it takes longer to understand the sentences and create an image for bizarre than common items (Robinson-Riegler & McDaniel, 1994; Waddill & McDaniel, 1998). Thus, the vividness ratings might be driven by the time that it takes for participants to create the image. Certainly, there is no sure way of telling whether this was the case in the present experiments, as the participants were given a fixed time to read the sentence and create an image, but the subjective ease of encoding seems to be a good candidate for explaining why bizarre items are rated less vividly than common items.

One of the future directions could be to conduct self-paced studies of bizarreness to investigate two issues. The first issue is whether different ratings such as vividness, bizarreness, comprehensibility and memorability correlate with the time to encode these sentences. For example, the participants could be asked to press a key as soon as they successfully complete the task (e.g. comprehend the sentence, create an image or evaluate its plausibility) and see whether the time difference to encode bizarre and common sentences drive the correlations between ratings and timing. To explain, considering that in self-paced studies, bizarre sentences take longer to encode than common sentences, it is useful to see whether self-perceived difficulty, as measured by the time that it takes to encode these sentences, correlate with the ratings of vividness, comprehensibility or plausibility of the sentences.

The second issue to investigate is whether the extended time to comprehend bizarre sentences and create images correlates with actual memory performance. Interestingly enough, even though imagery has been used to investigate participants' beliefs in their future performance, the bizarreness effect has not been investigated in the context of how metamemorial judgments correlate with memory accuracy. For an experimental paradigm in which we know that the comprehension times are longer for the distinctive items, it is interesting that the researchers have not investigated the relationship between encoding time and recall.

Another interesting question that emerged in Experiment 5, where memorability ratings were used along with intentional memory instructions, was the dissociation between the memorability and the actual memory performance. The memory performance for bizarre and common items was not significantly different; yet, participants' ratings for bizarre sentences were lower than common sentences. The finding that bizarre sentences are rated as less memorable may provide further proof that the ratings are not necessarily driven by some diagnostic cue of recall, but solely the subjective ease by which the sentence is understood, as found in Hertzog et al (2003). Looking at the bizarreness effect from the perspective of metamemory may give us further clues as to whether and how monitoring of information for different types of sentences that might or might not fit into our schemas of general world knowledge, might enhance our memory.

Table 1.

Means and Standard Deviations for Memory Performance in Experiment 1

Experimental Condition		Access		Total Noun Recall	
		Bizarre	Common	Bizarre	Common
verbal distraction control	Mean	0.54	0.32	0.39	0.26
	Std. Dev.	0.12	0.15	0.09	0.14
verbal distraction	Mean	0.49	0.25	0.42	0.22
	Std. Dev	0.18	0.17	0.19	0.15
visual distraction control	Mean	0.45	0.39	0.34	0.31
	Std. Dev	0.15	0.16	0.14	0.16
visual distraction	Mean	0.59	0.3	0.49	0.25
	Std. Dev	0.19	0.15	0.15	0.14

Table 2.

Means and Standard Deviations for Memory Performance and Vividness Ratings in Experiment 2

Experimental Condition		Access		Total Noun Recall		Vividness	
		Bizarre	Common	Bizarre	Common	Bizarre	Common
Control	Mean	0.48	0.32	0.38	0.26	2.82	4.06
	Std. Dev.	0.13	0.15	0.12	0.14	0.89	0.69
Visuospatial WM task	Mean	0.43	0.26	0.36	0.21	3.36	4.17
	Std. Dev.	0.20	0.13	0.19	0.11	0.78	0.83
Verbal WM task	Mean	0.45	0.33	0.35	0.28	3.41	4.15
	Std. Dev.	0.19	0.18	0.19	0.16	0.96	0.81

Table 3.

Means and Standard Deviations for Memory Performance in Experiment 3

Type of score	Experimental Condition		Access		Total Noun Recall	
			Bizarre	Common	Bizarre	Common
Unconditional scores	Visual Distraction	Mean	0.33	0.20	0.22	0.14
		Std. Dev.	0.18	0.11	0.14	0.1
	Control	Mean	0.45	0.28	0.37	0.24
		Std. Dev.	0.18	0.15	0.19	0.16
Conditional scores	Visual Distraction	Mean	0.29	0.18	0.2	0.15
		Std. Dev.	0.18	0.13	0.14	0.11
	Control	Mean	0.45	0.28	0.37	0.24
		Std. Dev.	0.18	0.15	0.19	0.16

Table 4.

Means and Standard Deviations for Memory Performance in Experiment 4

Type of score	Experimental Condition		Access		Total Noun Recall	
			Bizarre	Common	Bizarre	Common
Unconditional scores	Visual Distraction	Mean	0.35	0.22	0.26	0.15
		Std. Dev.	0.13	0.14	0.11	0.12
	Control	Mean	0.46	0.29	0.39	0.25
		Std. Dev.	0.12	0.17	0.15	0.15
Conditional scores	Visual Distraction	Mean	0.34	0.24	0.24	0.17
		Std. Dev.	0.12	0.16	0.10	0.12
	Control	Mean	0.46	0.29	0.39	0.25
		Std. Dev.	0.12	0.17	0.15	0.15

Table 5.

Means and Standard Deviations for Memory Performance and Memorability Ratings in Experiment 5

Experimental Condition		Access		Total Noun Recall		Memorability	
		Bizarre	Common	Bizarre	Common	Bizarre	Common
Control	Mean	0.38	0.35	0.35	0.3	2.87	3.22
	Std. Dev.	0.16	0.18	0.15	0.14	0.53	0.50

Table 6.

Means and Standard Deviations for Memory Performance in Experiment 6

Type of score	Experimental Condition		Access		Total Noun Recall	
			Bizarre	Common	Bizarre	Common
Unconditional scores	Visual Distraction	Mean	0.34	0.19	0.26	0.18
		Std. Dev.	(0.12)	(0.18)	(0.10)	(0.17)
	Control	Mean	0.45	0.34	0.35	0.33
		Std. Dev.	(0.14)	(0.15)	(0.14)	(0.15)
Conditional scores	Visual Distraction	Mean	0.37	0.22	0.27	0.20
		Std. Dev.	(0.18)	(0.17)	(0.12)	(0.16)
	Control	Mean	0.45	0.34	0.35	0.33
		Std. Dev.	(0.14)	(0.15)	(0.14)	(0.15)

Table 7.

Plausibility Ratings for Experiment 6

Experimental Condition		Plausibility	
		Bizarre	Common
Visual	Mean	1.64	4.32
	Std. Dev.	(.94)	(1.01)
Control	Mean	1.28	4.71
	Std. Dev.	(.26)	(.25)

Table 8.

Means and Standard Deviations for Memory Performance in Experiment 7

Type of score	Experimental Condition		Access		Total Noun Recall	
			Bizarre	Common	Bizarre	Common
Unconditional scores	Visual Distraction	Mean	0.40	0.27	0.33	0.24
		Std. Dev.	(0.18)	(0.11)	(0.14)	(0.10)
	Control	Mean	0.36	0.29	0.31	0.26
		Std. Dev.	(0.20)	(0.18)	(0.16)	(0.16)
Conditional scores	Visual Distraction	Mean	0.42	0.27	0.35	0.24
		Std. Dev.	(0.19)	(0.16)	(0.14)	(0.14)
	Control	Mean	0.36	0.29	0.31	0.26
		Std. Dev.	(0.20)	(0.18)	(0.16)	(0.16)

Table 9.

Plausibility ratings for Experimental 7

Experimental Condition		Plausibility	
		Bizarre	Common
Verbal	Mean	1.27	4.83
	Std. Deviation	(.29)	(.20)
Control	Mean	1.23	4.68
	Std. Deviation	(.25)	(.34)

Appendix A:

Sentence Materials for Experiments 1-7

Common Sentences	Bizarre Sentences
The banker dropped the newspaper in the puddle.	The banker floated across the puddle on a newspaper.
The biscuits were visible through the oven window.	The biscuits screamed when the oven jumped out the window.
The butterfly was examined by the biologist using the microscope.	The butterfly examined the biologist using a microscope.
The baby put the rattle in his mouth.	The rattle put the baby in its mouth.
The doughnuts on the floor were covered with ants.	The ants spit the doughnuts on the floor.
The maid spilled ammonia on the table.	The maid licked ammonia off the table.
The goldfish was swimming in the bowl next to the sofa.	The goldfish was eating out of the bowl on the sofa.
The spider crawled on the plant which was on the porch.	The spider watered the plant on the porch.
The dog chased the bicycle down the street.	The dog rode the bicycle down the street.
The camper found a mosquito on his binoculars.	The mosquito sighted the camper through the binoculars.
The fisherman pulled the lobster out of the barrel.	The lobster pulled the fisherman out of the barrel.
The burglar found a necklace under the mattress.	The mattress strangled the burglar with a necklace.
The chimp swung from the rope in the cage.	The chimp hung himself with a rope in the cage.
The minister read the bible after dinner.	The minister ate the bible during dinner.
The teacher sat at the desk grading a paper.	The teacher balanced the desk on the paper.
The farmer left his hat by the tomatoes.	The farmer made a hat out of tomatoes.

Appendix B:

Vividness-Accuracy Analyses for Experiments 1-4

General description of the analyses

The relationship between vividness and memory accuracy was analyzed with respect to the sentence type and experimental condition. In order to realize this, each participant's memory accuracy for each trial was dummy coded. If the participants remembered even only one noun from a sentence, their score was coded as 1, and if they could not remember any nouns from a sentence, their memory accuracy was coded as 0. Then, each participant's vividness rating was averaged separately within accuracy and sentence type, giving four different scores for each participant: vividness ratings for bizarre inaccurate sentences, bizarre accurate sentences, common accurate sentences and common inaccurate sentences. Following this, two separate ANOVA analyses were conducted for bizarre and common sentences, accuracy as the within subjects factor, and the experimental condition as the between subjects. The results for each experiment are briefly explained below

Experiment 1 vividness ratings

The results revealed a significant main effect for accuracy ($F(1,44) = 6.22$, $MS_e = .31$, $p<.05$) such that the bizarre accurate sentence had lower vividness ratings ($M = 2.90$ $SD = .95$) than bizarre inaccurate sentences ($M = 3.18$ $SD = .97$). All the other effects were non-significant (Fs<1). For common sentences, the results yielded a reversed pattern for main effect of accuracy ($F(1,44) = 4.87$, $MS_e = .27$, $p<.05$): Accurate sentences ($M = 4.12$ $SD = .95$) were rated more vividly than inaccurate sentences ($M = 3.90$ $SD = .89$). Neither the main effect of experimental group (F(3, 43) = 2.54, MS_e = 1.29, p>.05) nor the interaction (F<1) reached significance.

Experiment 2 vividness ratings

The analyses for bizarre sentences revealed a main effect for the experimental group, (F(1,87) = 3.97, MS_e = 1.59, p <.05). Post hoc comparisons using Fisher's LSD revealed that the participants in the control group (M =2.81, SD=.97) rated all sentences less vividly than the participants in the AS (M =3.4, SD=1.10) and spatial tapping (M =3.33, SD=.87) groups. All other effects were insignificant (Fs < 2). The analyses for common sentences revealed that there was a main effect for accuracy such that participants gave higher vividness ratings to accurate common sentences M =4.24, SD=.90) than inaccurate common sentences (M =4.06, SD=.81) (F(1,87) = 6.81, MS_e = .239, p <.05). None of the other effects reached significance (F < 1 for experimental condition; F(1,87) = 2.02, MS_e = .239, p >.10 for interaction).

Experiment 3 vividness ratings

The analysis for the bizarre sentences revealed a marginally significant main effect for accuracy (F(1,30) = 3.58, MS_e = .32, p=.068), with accurate sentences rated lower than inaccurate sentences. The other effects did not reach significance (F<2). The analysis for the common sentences revealed a different pattern. None of the main effects were significant, but the interaction between the two factors reached significance (F(1,30) = 5.59, MS_e = .18, p<.05). The participants in the distraction group rated the accurate (M = 3.97, SD= .92) and the inaccurate sentences (M = 3.98, SD= .85) similarly, but the participants in the control group rated the accurate sentences (M = 4.28, SD= .67) as more vivid than the inaccurate sentences (M = 3.92, SD= .85)

Experiment 4 vividness ratings

The analysis for the bizarre sentences revealed no main effect for accuracy (F<1), a marginally significant main effect for experimental condition (F(1,22) = 3.80, MS_e = 2.057, p=.064), with the sentences in the control condition (M = 2.90, SD = 1.07) rated less vividly than the sentences in the verbal condition (M = 3.70, SD = 1.06) and no interaction (F(1,22) = 1.45, MS_e = .21, p>.10). The analyses for the common sentences revealed no main effect for accuracy (F(1,22) = 1.16, MS_e = .079, p>.10), no main effect for experimental condition (F<1) and a significant interaction between the two factors (F(1,22) = 10.21, MS_e = .079, p<.01). Posthoc comparisons with paired samples t-test revealed that there was no difference in the ratings for inaccurate and accurate sentences in the verbal distraction condition (t(9) = 1.67, p>.10), but the ratings for the inaccurate sentences (M = 4.20, SD = .75) were significantly lower as compared to accurate sentences in the control condition (M = 4.58, SD = .64) (t(10) = 2.82, p<.05).

REFERENCES

Anderson, D. C., & Buyer, L. S. (1994). Is imagery a functional component of the "bizarre imagery" phenomenon? *The American Journal of Psychology, 107*, 207-222.

Andrade, J., Kemps, E., Werniers, Y., May, J., & Szmalec, A. (2002). Insensitivity of visual short term memory to irrelevant visual information. *Quarterly Journal of Experimental Psychology, 55A*, 753–774.

Auble, P. M., & Franks, J. J. (1978). Effect of effort toward comprehension on recall. *Memory & Cognition*, *6*, 20-25.

Auble, P. M., Franks, J. J., & Soraci, S. A. (1979). Effort toward comprehension: Elaboration or "aha!"? *Memory* & *Cognition, 7,* 426-434.

Baddeley, A. D. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, *129,* 126-145.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature reviews: Neuroscience*, *4*, 829-839.

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (pp. 47–90). New York: Academic Press.

Baddeley, A. D., Logie, R.H., Bressi, S., Della Sala, S., & Spinnler, H. (1986). Senile dementia and working memory. *Quarterly Journal of Experimental Psychology, 38A*, 603-618.

Barnard, P.J., Scott, S.K., & May, J. (2001). When the central executive lets us down: Schemas, attention, and load in a generative working memory task. *Memory*, *9*, 209- 221.

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory & Language*, *28*, 610-632.

Bell, R., Buchner, A., & Mund, I. (2008). Age-related difference in irrelevant-speech effects. *Psychology and Aging, 23*, 377–391.

Boyle, R., & Coltheart, V. (1996). Effects of irrelevant sounds on phonological coding in reading comprehension and short-term memory. *The Quarterly Journal of Experimental Psychology, 49A,* 398-416.

Bugelski, B. R., Kidd, E., & Segmen, J. (1968). Image as a mediator in one-trial paired-associate learning. *Journal of Experimental Psychology, 76,* 69-73.

Burns, D. J. (1996). The bizarre imagery effect and intention to learn. *Psychonomic Bulletin and Review, 3,* 254-257.

Buzan, T. (1991). *Use Your Perfect Memory: Dramatic New Techniques for Improving Your Memory*. NY: Penguin Publishing.

Calvo, M. G., & Castillo, M. D. (1995). Phonological coding in reading comprehension: The importance of individual differences. *European Journal of Cognitive Psychology*, *7,* 365-382.

Christoffels, I. K. (2006) Listening while talking: The retention of prose under articulatory suppression in relation to simultaneous interpreting. *European Journal of Cognitive Psychology, 18,* 206-220.

Cocchini, G., Logie, R. H., Della Sala, S., & MacPherson, S. E. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition*, *30,* 1086-1095.

Collyer, S. C., Jonides, J., & Bevan, W. (1972). Images as memory aids: Is bizarreness helpful? *The American Journal of Psychology, 85,* 31-38.

Cornoldi, C., Cavedon, A., De Beni, R., & Pra Baldi, A. (1988). The influence of the nature of material and of mental operations on the occurrence of the bizarreness effect. *The Quarterly Journal of Experimental Psychology, 40A,* 73-85.

Cox, S. D., & Wollen, K. A. (1981). Bizarreness and recall. *Bulletin of the Psychonomic Society, 13,* 244-245.

Delin, P.S. (1968). Success in recall as a function of success in implementation of mnemonic instructions. *Psychonomic Science, 12,* 153-154.

Delin, P. S. (1969). An experiment examining list-learning with and without the mnemonic use of bizarre associations. *Journal of General Psychology, 81,* 249-260.

Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: A tool for unwelding visuo-spatial memory. *Neuropsychologia, 37,* 1189–1199.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19-23.

Einstein, G. O., & McDaniel, M. A. (1987). Distinctiveness and the mnemonic benefits of bizarre imagery. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individual differences, and applications* (pp. 78-102). New York: Springer-Verlag.

Graf, R. (1980). Two consequences of generating: Increased inter- and intra-word organization of sentences. *Journal of Verbal Learning and Verbal Behavior, 19,* 316-327.

Hanley, J. R., & Broadbent, C. (1987). The effects of unattended speech on serial recall following auditory presentation. *British Journal of Psychology, 78,* 287-297.

Hanley, J. R., & Bakopoulou, E. (2003). Irrelevant speech, articulatory suppression, and phonological similarity: A test of the phonological loop model and the feature model. *Psychonomic Bulletin & Review, 10,* 435–444.

Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue utilized for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 22-34.

Hirshman, E., Whelley, M. M., & Palij, M. (1989). An investigation of paradoxical memory effects. *Journal of Memory & Language*, *28,* 594-609.

Imai, S., & Richman, C.L. (1991). Is the bizarreness effect a special case of sentence reorganization. *Bulletin of the Psychonomic Society*, *29*, 429-432.

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The Phonological Store of Working Memory: Is It Phonological and Is It a Store? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30,* 656–674.

Kroll, N., Schepeler, E. M., & Angin, K. T. (1986). Bizarre imagery: The misremembered mnemonic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 42-54.

Kroll, N. E. A., & Tu, S.-F. (1988). The bizarre mnemonic. *Psychological Research*, *50,* 28-37.

LeCompte, D. C. (1994). Extending the irrelevant speech effect beyond serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20,* 1396-1408.

Logie, R. H. (1986). Visuo-spatial processing in working memory. *The Quarterly Journal of Experimental Psychology, 38A,* 229-247.

Logie, R. H., Zucco, G. M., & Baddeley, A. D. (1990). Interference with visual short-term memory. *Acta Psychologica*, *75,* 55-74.

Lorayne, H. (2010). *Ageless Memory: The Memory Expert's Prescription for a Razor-Sharp Mind.* NY: Black Dog & Leventhal Publishers.

Macken, W. J., & Jones, D. M. (1995). Functional characteristics of the inner voice and the inner ear: Single or double agency? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 436-448.

Lorayne, H., & Lucas, J. (1974). *The Memory Book.* NY: Stein and Day.

Marshall, P. H., Nau, K. L., & Chandler, C. K. (1979). A structural analysis of common and visual mediators. *Bulletin of the Psychonomic Society, 14,* 103-105.

Marshall, P. H., Nau, K. L., & Chandler, C. K. (1979). A functional analysis of common and visual mediators. *Bulletin of the Psychonomic Society, 15,* 375-377.

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., Dornburg, C. C., Guynn, M. J. (2005). Disentangling encoding versus retrieval explanations of the bizarreness effect: Implications for distinctiveness. *Memory & Cognition, 33,* 270-279.

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.

McDaniel, M. A., Einstein, G. O., DeLosh, E. L., May, C. P., & Brady, P. (1995). The bizarreness effect: it's not surprising, it's complex. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(2), 422-435.

Merry, R., & Graham, N. C. (1978). Imagery bizarreness in children's recall of sentences. *British Journal of Psychology, 69,* 315-321.

Nappe, G. W., & Wollen, K. A. (1973). Effects of instructions to form common and bizarre mental images on retention. *Journal of Experimental Psychology, 100,* 6-8.

Nicholas, S., & Worthen, J. B. (2009). Adult age differences in memory for distinctive information: evidence from the bizarreness effect. *Quarterly Journal of Experimental Psychology, 62,* 1983-1990.

Oswald, C. J. P., Tremblay, S., & Jones, D. M. (2000). Disruption of comprehension by the meaning of irrelevant sound. *Memory, 8,* 345–350.

Parker, A., & Dagnall, N. (2009). Concreteness effects revisited: The influence of dynamic visual noise on memory for concrete and abstract words. *Memory, 17,* 397-410.

Paivio, A. (1971). *Imagery & verbal processes.* New York: Holt, Rhinehart, & Winston.

Paivio, A. (1991). Dual coding theory: Retrospect & current status. *Canadian Journal of Psychology, 45,* 255-287.

Quinn, J. G., & McConnell, J. (1996). Irrelevant pictures in visual working memory. *Quarterly Journal of Experimental Psychology, 49A,* 200-215.

Quinn, J. G., & McConnell, J. (2006). The interval for interference in conscious visual imagery. *Memory, 14*, 241-252.

Richardson, J. T. E., & Baddeley, A. D. (1975). The effect of articulatory suppression in free recall. *Journal of Verbal Learning and Verbal Behavior, 14*, 623-629.

Richman, C. L. (1994). The bizarreness effect with complex sentences: Temporal effects. *Canadian Journal of Experimental Psychology, 48,* 444-450.

Robinson-Riegler, B., & McDaniel, M. A. (1994). Further constraints on the bizarreness effect: Elaboration at encoding. *Memory & Cognition, 22*, 702-712.

Salame, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior, 21,* 150-164.

Salway, A. F. S., & Logie, R. H. (1995). Visuospatial working memory, movement control and executive demands. *British Journal of Psychology, 86*, 253-269.

Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory & Cognition*, *19*(6), 523-542.

Smyth, M. M., Pearson, N.A., & Pendleton, L.R. (1988): Movement and working memory: Patterns and positions in space. *Quarterly Journal of Experimental Psychology, 40A*, 497-514.

Toppino, T. C., & Pisegna, A. (2005). Articulatory suppression and the irrelevant-speech effect in short-term memory: Does the locus of suppression matter? *Psychonomic Bulletin & Review, 12*, 374-379.

Trudeau, K. (1997). *Kevin Trudeau's Mega Memory: How to Release Your Superpower Memory in 30 Minutes or Less a Day.* NY: William Morrow and Company.

Van Gerven, P. W. M., Meijer, W. A., Vermeeren, A., Vuurman, E. F., & Jolles, J. (2007). The irrelevant speech effect and the level of interference in aging. *Experimental Aging Research, 33,* 323–339.

Waddill, P. J., & McDaniel, M. A. (1998). Distinctiveness effects in recall: Differential processing and privileged retrieval? *Memory & Cognition, 26*, 108-120.

Wallace, W. H., Turner, S. H., & Perkins, C. C. (1957). Preliminary studies of human information storage. Signal Corps Project No. 1326, Institute of Cooperative Research, University of Pennsylvania.

Weir, D., & Richman, C. L. (1996). Subject-generated bizarreness: Imagery or semantic processing. *The American Journal of Psychology, 109,* 173-185.

Wollen, K. A., & Cox, S. D. (1981). Sentence cueing and the effectiveness of bizarre imagery. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 386-392.

Wollen, K. A., & Margres, M. G. (1987). Bizarreness and the imagery multiprocess model. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individual differences, and applications* (pp. 103-128). New York: Springer-Verlag.

Wollen, K. A., Weber, A., & Lowry, D. H. (1972). Bizarreness versus interaction of mental images as determinants of learning. *Cognitive Psychology, 3,* 518-523.

Worthen, J. B. (1997). Resiliency of bizarreness effects under varying conditions of verbal and imaginal elaboration and processing mode. *Journal of Mental Imagery, 21*, 167-194.

Worthen, J. B. (2006). Resolution of discrepant memory strengths: An explanation of the effects of bizarreness on memory. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and Memory.* (pp. 65-88). New York: Oxford University Press.

Worthen, J. B., & Deschamps, J. D. (2008). Humour mediates the facilitative effect of bizarreness in delayed recall. *British Journal of Psychology, 99*, 461-471.

Worthen, J. B., Garcia-Rivas, G., Green, C. R., & Vidos, R. A. (2000). Tests of a cognitive-resource allocation account of the bizarreness effect. *The Journal of General Psychology, 127*, 117-144.

Worthen, J. B., & Roark, B. (2002). Free recall accuracy for common and bizarre verbal information. *The American Journal of Psychology, 115,* 377-394.

Yarmey, A. (1984) Bizarreness effects in mental imagery. In A. Sheikh (Ed.), *International Review of Mental Imagery,* Vol. 1 (pp. 57-76). New York: Human Sciences Press.

Yates, F. A. (1966). *The Art of Memory.* Chicago: The University of Chicago Press.

Zaromb, F. M. & Roediger, H. L. (2009). The effects of ‘effort after meaning” on recall: Differences in within- and between-subjects designs. *Memory & Cognition, 37,* 447-463.