Modality Effects in False Recall:

Dissociations Between Short- and Long-Term Memory

Grace M. Glowniak

University of Michigan

Abstract

In recent research, the Deese/Roediger-McDermott (DRM) paradigm has been used to elicit false memories for critical, non-presented words in both short-term (STM) and long-term (LTM) recognition following memory-list presentation in either the visual or auditory modality (Olszewska, Reuter-Lorenz, Munier, & Bender, in press). In STM, false memories were more frequent when lists were seen than when they were heard, whereas the opposite pattern occurred in LTM. The present study uses a hybrid short and long-term DRM paradigm to test modality effects when words are *recalled* during STM testing. The first aim is to determine if associative memory distortions exhibit the same modality effects during recall as they do during recognition. The second aim is to determine why differences in false memory frequency between STM and LTM in the two modalities arise. In Experiment 1, the same pattern of memory distortions as previously reported using recognition was demonstrated. In STM participants recalled more target lures following visual than auditory presentation, whereas in LTM this pattern was reversed. Experiment 2 showed that participants generated a similar number of semantically related words in the visual and auditory condition when tested using inclusion instructions in STM, which instruct participants to recall the studied items and other words that come to mind. In LTM, recognition in both modalities improved relative to the standard instruction and modality differences were no longer evident. These results suggest that the higher frequency of auditory false memories in LTM under the standard instruction is due to deeper encoding following the visual presentation.

Keywords: False memories, short term memory, long term memory, modality effects, inclusion instructions

Modality Effects in False Recall: Dissociations Between Short- and Long-Term Memory

It is a common belief that memory accurately depicts an experience, but research has shown that it is not always a reliable representation of the past (Roediger, 1996). When someone remembers an event that never occurred, or remembers something strikingly different from what actually happened, they are said to have had a false memory. Research into this phenomenon has grown significantly in the last two decades. According to Mazzoni (2002), there are two lines of research that demonstrate the existence of memory distortions: *suggestion dependent distortions* (e.g., Loftus, Miller, & Burns, 1978; Loftus & Palmer, 1974; Loftus & Pickrell, 1995; Takarangi, Parker, & Garry, 2006; van Damme & Smets, 2014) and *naturally occurring distortions* (e.g., Deese, 1959; Dewhurst, 2001; Dewhurst & Anderson, 1999; Roediger & McDermott, 1995). The first type refers to when an external suggestion is given about a past event, leading to misremembering it, whereas the second is a result of the innate imperfection of human memory. Suggestion dependent distortions arise under particular circumstances that people are exposed to e.g., misleading questions. This later leads to changes in memory content. However, naturally occurring distortions reveal that external manipulations are not needed to distort the content of memory. Therefore, the human mind cannot be viewed as a recording device, because these two approaches depict the reconstructive nature of memory (Bartlett, 1932). In this thesis the focus will be on false memories that result from *naturally occurring distortions*, using an experimental task designed to induce them.

Naturally occurring memory distortions

If false memories happen without any external trigger, then what is responsible for their occurrence? Underwood (1965) proposed the implicit associative response hypothesis as an explanation. This hypothesis states that when a word is presented, it also unconsciously activates other words that are strongly associated with the presented one. Even if a specific word is not presented, the participant could believe that it was, if it is related closely enough to the word they did see. Underwood tested this theory mostly through recognition, while Deese (1959) showed that this method worked through recall testing, as well.

The most common method to investigate false memories in the laboratory setting is the Deese/Roediger-McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995), which entails presenting subjects with a study list of up to fifteen words, all of which are semantically related to a predetermined target word, called the related lure, which itself is not presented (e.g the words *arachnid*, *tarantula*, *web*, *creepy,* etc. would be presented, and the related lure would be *spider*). After the presentation of the list, subjects are tested on the words through recall, and again through recognition following the presentation of multiple lists. It has been shown that subjects falsely recall the related lures about as often as words presented in the middle of the studied list, and falsely recognize the related lures over half of the time (Roediger & McDermott, 1995). Subjects are also confident in their false identification of the related lure, shown through remember/know judgments. Participants were instructed to rate the degree of their memory for a word, in which a “remember” judgment means they have a conscious memory of its presentation, while a “know” judgment means they can’t bring to mind the specific occurrence but they know it was presented. Participants give a “remember” judgment for the lures in over half of the recognition trials (Roediger & McDermott, 1995). These results indicate that the DRM paradigm can produce false memories in both recall and recognition tests in long term memory (LTM).

Associative memory distortions in short-term memory

Atkins and Reuter-Lorenz (2008) developed a version of the DRM task to investigate the possibility that memory can be distorted after short delays, which is called the short-term memory Deese/Roediger-McDermott paradigm, or STM-DRM (see also Coane, McBride, Raulerson and Jordan (2007)). They used lists of only four associated words so working memory capacity was not exceeded. Atkins and Reuter-Lorenz (2008) showed that a few seconds between the study period and a recognition test is sufficient for people to falsely recognize or falsely recall semantically related items that were not studied. Subjects made more false alarms in response to critical lures than to unrelated words, and took longer to reject the critical words than unrelated words when making correct negative responses. These results suggest the possibility that the mechanisms that underlie memory distortions may be shared by working memory and long-term memory.

To further test the hypothesis that short term and long term, episodic memory are formed through a unitary memory system rather than through multiple systems, Flegal, Atkins, and Reuter-Lorenz (2010) compared both STM and LTM within subjects using the hybrid-DRM paradigm, in which participants were either tested on words in short or long term memory, but not both. The expectation was that STM would result in fewer false recognitions, based on proposed theories that describe long-term memory as primarily semantic/associative and short-term memory as more sensory/perceptually based (Baddeley, 1972; Baddeley, Thompson, & Buchanan, 1975). However, the results showed that the adjusted frequency of false memories was equal between STM and LTM, providing more evidence that the mechanisms for the formation of false memories may be common across delay, and not particular to short or long term memory. These results also support the idea that semantic processes are just as prevalent in STM as they are in LTM.

Modality effects in LTM studies

In their original work, Roediger and McDermott (1995) used auditory presentation of studied lists, followed by recall and then recognition of items presented visually. As described above, they found that non-presented related lures were recalled and later recognized with high confidence. In a later experiment that instead used visual presentation along with visual recognition testing, the rates of false recall decreased significantly (Robinson & Roediger, 1997; Smith and Hunt, 1998). This suggested that the modality of word presentation at study has an effect on false memory frequency.

In follow-up work, Smith and Hunt (1998) corroborated the results by also showing a difference in frequency between the two testing modalities. They proposed the reason behind this was that implicit associate responses (Underwood, 1965) were present at the time of study, but that it was easier to distinguish between these and the words that were actually presented in the visual modality than in the auditory. To test this hypothesis, Smith and Hunt (1998) asked participants to rate the pleasantness of words after they were presented. The reason for using pleasantness ratings is that they require the participants to focus on the meaning of the word, rather than just its superficial features. This creates a deeper encoding in a manner that is unique to each word, a phenomenon known as item-specific processing. Deeper encoding increases memory for studied words, and therefore reduced memory errors that arose at the time of study, such as implicit associate responses. This resulted in a decrease in the overall frequency of false memories in both modalities, but showed the same relationship between them, wherein the visual modality led to fewer false alarms when compared with the auditory modality (however, see Maylor and Mo (1998) for an exception). Since the frequencies of false memories decreased for both auditory and visual presentation after a pleasantness rating, this suggests that the formation of memory distortions occurs, at least in part, at the time of study since false memory frequency for both modalities were affected equally.

Hypothesis and Predictions

As previously mentioned, the DRM task can be tested in two different ways: recall and recognition. It has been shown that subjects routinely perform better on recognition tasks than on recall (Hanawalt, 1937; Postman, Jenkins, & Postman, 1948; Andrew & Bird, 1938). Previous experiments that have tested for false memories in both auditory and visual presentation in short term memory have only tested through recognition. Olszewska, Reuter-Lorenz, Munier & Bendler (in press) replicated Smith and Hunt’s (1998) findings in long term memory, in which there was a higher frequency of false memories for the auditory modality than the visual, but showed this relationship was opposite in short term memory. The first purpose of this present study is to replicate this effect for recall, and make sure that the dichotomy between modality and memory type observed by Olszewska et al. (in press) is still present when participants have to actively come up with the words rather than recognize them.

The second purpose of this thesis is to examine why a difference in the modality effect arises when tested in STM as compared to LTM. One possibility is that words are encoded at a different depth if they are seen rather than if they are heard. If visually presentation items are encoded more deeply, then related lures would be more likely to come to mind during the auditory modality (Smith & Hunt, 1998). Another related possibility is that source monitoring (Johnson, Hashtroudi, & Lindsay, 1993) is more effective following visual presentation. Source monitoring is the collection of processes involved in determining the origin of memories, knowledge, and beliefs, like distinguishing between imagined events and those that actually occurred. If source monitoring is more effective in the visual modality, then it would be easier to reject the related lures, resulting in the lower frequency of false memories seen following visual presentation in LTM.

To test whether deeper encoding or greater source monitoring for visually presented items contributes to the modality effects, the inclusion instruction was used in the second experiment (Hege and Dodson, 2004). Inclusion instructions require the participants to report all test items that were studied and any related items that may have come to their mind during the study period. The reason this method is used is that it endorses both studied and non-studied words, and promotes deeper, item specific processing. Therefore, if shallower encoding following auditory presentation is responsible for the modality effects, it is expected that the difference in false memory frequencies between the modalities will be eliminated following the inclusion instruction, because the disparity in depth of encoding will be compromised. Alternatively, if there is still a modality effect and a similar number of related words are generated at encoding for both modality conditions, then more effective source monitoring for visually encoded items in LTM may be implicated as a potential basis for the visual advantage (e.g., the LTM modality effect).

Experiment 1

False recognition and recall of semantically associated lures were documented in visual working memory (Atkins & Reuter-Lorenz, 2008). False recognition was also revealed in auditory working memory, and was associated with fewer false memories than in visual working memory (Olszewska et al., in press). Experiment 1 tested whether the discrepancy in false memories between auditory and visual presentation, otherwise known as a modality effect, exists in free recall in short-term memory. The hybrid-DRM paradigm (Flegal, Atkins, & Reuter-Lorenz, 2010) was used to test if free recall produces the same modality effects in STM and LTM that was observed previously with recognition testing (Olszewska et al., in press).

Methods

Participants. Thirty-two University of Michigan students (22 females and 10 males) between the ages of 18 and 21 participated in this study for course credit. One male participant was excluded due to failure to follow the instructions. All were native English speakers.

Materials. The study was administered on a Dell computer running the software E-Prime 2.0. For the visual condition, words in black font were presented in the middle of a white screen, and for the auditory condition, over-the-ear headphones presented pre-recorded words at a pre-set, comfortable volume. An experimenter in the testing room with the participants recorded the responses for the short term memory (STM) condition, and a standard keyboard recorded the responses for the long term memory (LTM) condition.

Seventy-two lists of four semantically related words, as previously used by Flegal et al. (2010), were used in this experiment. Each word was a strong associate of a predetermined related lure, as done in the DRM paradigm (Deese, 1959; Roediger & McDermott, 1995). These seventy-two lists were divided into four, eighteen list blocks. Two of the blocks were presented visually, and two were presented auditorily, in either a V-A-A-V or A-V-V-A order.

Design and Procedure. As in Flegal et al. (2010) and Flegal and Reuter-Lorenz (2014), the four-word memory sets were tested either in STM or in a surprise recognition test following completion of all STM trials (i.e., LTM), in order to examine short-term and long-term memory distortions concurrently and within-subjects. Unlike these prior studies, in this experiment STM was tested by free recall. On half of the STM trials no recall instruction was given upon the offset of the green screen. On these trials participants were instructed merely to press an arbitrary key on the keyboard to proceed to the next trial. Unbeknownst to them, these trials would then be probed in a surprise LTM recognition test.

The visual condition began with two practice trials that were intended to familiarize the participants with the procedure. The sequence of events in each visual trial was as follows: a black fixation cross was presented in the middle of the screen for 500 ms, followed by four words for 1000 ms each with 500 ms inter-stimulus intervals. After the last word, there was a 500 ms interval before the onset of a green screen, which demarcated the 3000 ms retention interval and signaled the participants to recite continuously, “the, the, the”. After 3000 ms, the participants were shown one of two instructions in the middle of the white screen: either “Recall Now”, in which the participants audibly repeated the four words presented in the list while the responses were recorded by the experimenter, or “Press Now”, in which the participants simply hit any key on the keyboard to advance to the next list of words. The recall period was untimed. When the participants were done with recall, they were instructed to press any button on the computer’s keyboard to advance to the next list of words. There were nine Recall and nine Press trials within each block, and the order presented was randomly intermixed.

The auditory block also began with two practice trials. Each trial began with a beep that lasted 500 ms, followed by four spoken words at a rate of approximately one per second, with 500 ms inter-stimulus intervals. Five hundred ms after the fourth word, the computer screen turned green for 3000 ms while the participants repeated the word “the” out loud, as in the visual trials. Then the instructions “Recall Now” or “Press Now” appeared on the white computer screen. The participants audibly recalled the studied words, which the experimenter recorded, or pressed a button on the keyboard without saying anything out loud. As with the visual blocks, the recall period was untimed. Half of the trials in each block were selected randomly to be recalled.

Following the four STM blocks, the participants completed two LTM blocks using recognition. Participants were instructed to make a keyboard response to each presented word, by pressing the “M” key if they had seen or heard the word at any point during the STM blocks, or by pressing the “Z” key if they had not seen or heard the word. Once either key was pressed, the experiment proceeded to the next trial. There were 72 LTM recognition trials, 36 of which tested memory sets that were not tested in the STM phase (i.e., Press trials). Additionally, to match the proportions of yes/no responses across the STM and LTM tests, there were 12 trials of studied associates from memory sets that were recalled at STM (never including theme words from studied probe trials), and 24 trials of unstudied, unrelated lures that were taken from Flegal and Reuter-Lorenz (2014; see also Olszewska et al., in press) and matched for frequency and word length with the corpus of theme words used in our experiments. As in their prior work, these trials were not analyzed.

The 72 LTM trials were divided into two modality blocks, auditory and visual, such that in each block probes were presented in the same modality in which they or their associates were studied. The first LTM block was always in the modality opposite to that used in the last STM block. Probe words in each modality were presented at a rate of approximately one word per 3000 ms, permitting time for the yes/no recognition response.

Results

STM analyses

To assess possible differences in memory accuracy due to study modality, we compared correct recall for items seen and those that were heard. Recall was better for words presented auditorily (M = 3.89, SEM = 0.03) than visually (M = 3.66, SEM = 0.06) $t(30) = 4.68$; $p < .001$; $d = 0.83$. This replicates the results found in previous experiments that studied modality differences in STM through recognition testing (Olszewska et al., in press).

The next set of analyses focused on recall errors (see Table 1). To assess possible modality differences in the rate of memory distortions, we compared all semantic errors (related lure errors plus other semantic errors) to phonological errors. The main effect of modality $F(1, 30) = 9.7$; $p = .004$; $\eta_p^2 = 0.1$ and of error type $F(1, 30) = 7.09$, $p = .01$; $\eta_p^2 = 0.35$ were significant, indicating that participants made more errors of all types for visual than auditory study lists, and more semantic errors than phonological errors. The significant interaction between modality and error type $F(1, 30) = 15.9$; $p < .001$; $\eta_p^2 = 0.2$ indicated there were more semantic errors for the visual than auditory condition, $t(30) = 4.79$; $p < .001$; $d = 0.87$, however the phonological errors did not differ between the two modalities $t(30) < 1$.

Focusing exclusively on false recall of lure words indicated that participants produced more of these intrusions following visual than auditory presentation $t(30) =$

4.67; *p* < .001; *d* = 0.86. Again, as with correctly recalled words, these results replicate the modality effect found by Olszewska et al. (in press).

	Lure errors	Other semantic errors	Phonological errors	Total semantic errors	Total errors
Visual	1.0 (0.16)	0.74 (0.17)	0.58 (0.20)	1.78 (0.30)	2.35 (0.40)
Auditory	0.29 (0.11)	0.41 (0.16)	0.70 (0.14)	0.82(0.2)	1.42 (0.30)

Table 1. Mean number of lure errors, other semantic errors and phonological errors in Experiment 1. Standard error of the mean in parentheses.

LTM analyses

A t-test comparing correct recognition of studied probes ("*yes*" responses to studied probes) in LTM for the auditory and visual study conditions showed no differences $t(30) < 1$; *n.s*. (see Table 2 for proportion of "*yes*" responses for each of the three probe types in each modality in LTM). However, using the high-threshold measure (*Pr*) (Seamon, Luo, Kopecky, Price, Rothschild, Fung, & Schwartz, 2002) in which errors to unrelated items are subtracted from the hit rate, recognition accuracy was marginally higher in the visual than in the auditory condition $t(30) = 1.71$; $p = 0.09$; $d = 0.31$.

To determine whether the rate of false semantic memories differed between modalities, the proportion of "*yes*" responses to related and unrelated lures was compared using a two-way ANOVA with probe type (related, unrelated) and modality (auditory, visual) as repeated within-subject factors. There was a main effect of modality, *F*(1, 30) = 10.67, *p* = .002, η_p^2 = .26 indicating that more errors were made in the auditory condition and a main effect of probe type, *F*(1, 30) = 66.09, *p* < .001, η_p^2 = .69 indicating more errors were made to related lures than to unrelated lure probes. The interaction between modality and probe type was not significant *F*(1, 30) = 0.2; *p* = 0.62; *n.s*. The relationship of false alarms to related and unrelated lures differed between auditory and visual presentation, *t*(30) = 2.08; *p* = 0.04, *d* = 0.38 and *t*(30) = 2.23; *p* = 0.03, *d* = 0.41, respectively. When the sensitivity (*Pr*) measure was applied, false recognition did not differ between modalities *t*(30) = 0.5; *p* = 0.6; *n.s.*

	Studied	H-T correction	Related	H-T correction	Unrelated
Visual	0.69(0.04)	0.65(0.05)	0.25(0.03)	0.21(0.03)	0.04(0.01)
Auditory	0.66(0.04)	0.54(0.06)	0.38(0.05)	0.26(0.06)	0.12(0.04)

Table 2. Proportion of “*yes*” responses for each of the three probe types in each modality in LTM in Experiment 1. H-T correction = data corrected for baseline using the high-threshold procedure (*Pr*). Standard error of the mean in parentheses.

Discussion

The results reported in Experiment 1 demonstrate the existence and reliability of a modality effect in false STM when memory is tested with recall. In comparison with visual presentation, auditory presentation of words not only resulted in more correctly recalled studied items, but also reduced the level of false memories. This suggests that memory is more accurate in the short term when words are heard, than when they are read. This replicates the results seen in previous experiments when words were tested through recognition rather than recall (Olszewska et al., in press). However, there was a general tendency for this pattern to reverse in LTM when the base rate of “yes” responses (false recognition of unrelated words) was taken into account through the high-threshold procedure. Auditory presentation resulted in a higher rate of false recognitions for both related and unrelated lures. When this was corrected for, studied words were recognized more often and fewer false recognitions of lures occurred in the visual modality, suggesting better memory in the visual condition. This result partially replicates the effects reported in the Olszewska et al. paper (in press), wherein the auditory modality resulted in more false memories in LTM. However, they found no modality difference in the correct recognition of studied words. The decrease of correct recognition and increase in recognition of unrelated words during the auditory condition in the present experiment are possibly due to the use of recall during the STM phase of the task. Participants experienced significantly more auditory stimulation than visual, because they heard themselves say the words presented as they recalled them. This could lead to more liberal responding or greater uncertainty about auditorily presented items in LTM which could explain the greater number of false alarms and potentially the lower memory accuracy for the auditory condition relative to the visual condition and in comparison to the results found previously using recognition in STM. Alternatively, the disparity in the correct recognitions between the modalities could arise from a difference in the depth of encoding that occurs between modalities (Craik & Lockhart, 1972; Flegal & Reuter-Lorenz, 2014). Fewer correct recognitions were made in the auditory modality, so this suggests that words are encoded more shallowly when they are heard than when they are seen.

Experiment 2

Experiment 1 demonstrated a modality effect in short-term recall using a variant of the hybrid-DRM paradigm. This short-term modality effect included higher accuracy and fewer semantic intrusions for the lists that were heard than for lists that were read by participants. These results replicated the modality effect in recognition reported by Olszewska et al. (in press). There was a partial reversal of this modality effect in LTM, where memory performance in the auditory condition was worse than in the visual condition. This difference was evident in corrected recognition scores, and was largely due to more false alarms in the auditory condition.

In Experiment 2 we tested whether this modality effect in LTM may be due to a modality differences in semantic processing during encoding in STM. If there are modality differences in associative processing during encoding, these should be evident by instructing participants to report any associations that come to mind during list presentation. This is referred to as the inclusion instructions, and it is how participants were tested in the STM phase of Experiment 2. Alternatively, the inclusion instruction may promote associative processing more so than the standard instruction, and could potentially equate associative processing across the two modalities. If the modality difference in LTM observed in Experiment 1 is due to differences in processing depth, then the inclusion instruction could eliminate this LTM modality difference, by equating processing depth at encoding.

Methods

Participants. Thirty-two University of Michigan students between the ages of 18 and 21 participated in this study, 23 of whom were female. Course credit was granted as compensation, and all participants were native English speakers.

Materials. All materials were identical to those in the Experiment 1 with the exception of the instruction. Instead of standard instructions, participants were given inclusion instructions. When participants were given the instruction to “Recall Now” after a list presentation in the STM phase, they not only said the four words that were presented during the study period, but also any words that came to their mind while they were studying them.

Design and Procedure. The procedure was identical to that in the Experiment 1 except that the inclusion instruction was used that required participants to recall memorized words and report all words that came to their mind.

Results

STM analyses

To test memory accuracy, correct free recall of memory items following the retention interval was analyzed for both modalities. The participants better recalled words presented auditorily (*M* = 3.77, *SEM* = 0.04) than visually (*M* = 3.50, *SEM* = 0.06) $t(30) = 5.47$; $p < .001$; $d = 0.82$. This is the same pattern that was seen in Experiment 1.

Next, the number of semantically related words that were produced during recall[1] was analyzed (see Table 3 for a mean number of correctly recalled words, words

[1] Since the inclusion instruction required participants to recall all words that came to their mind while encoding, these responses were not called errors.

produced following an inclusion instruction and phonological errors). Modality differences in the rate of semantically related words produced under the inclusion instructions were tested by a paired *t*-test, which showed similar numbers of related words produced following visual and auditory presentations $t(31) < 1$. This implies that participants can engage similar relational processing during both visual and auditory encoding when instructed to be inclusive in their recall. Any LTM modality effects that would emerge in the present experiment therefore would likely arise from some other source.

	Lures	Other Semantic Associates	Phonological responses	Total Semantic Associates
Visual	5.81(0.4)	22.77(2.7)	0.48(0.17)	28.58(2.7)
Auditory	4.81(0.4)	24.35(2.9)	0.48(0.12)	29.16(2.9)

Table 3. Mean number of lures and other semantically related words produced under an inclusion instruction as well as phonological responses in Experiment 2. Standard error of the mean in parentheses.

LTM analyses

To determine whether accurate recognition of studied probes ("*yes*" responses to studied probes) differed due to modality in LTM, a *t*-test was conducted and showed no differences between them $t(30) < 1$; *n.s*. (see Table 4 for proportion of "*yes*" responses for each of the three probe types in each modality in LTM). The same results were obtained even when the standard high-threshold correction procedure (*Pr*) was applied, so correct recognition rates did not differ between modalities, either $t(30) = 0.2$; $p = 0.8$.

To determine whether the rate of false semantic memories differed between modalities, the proportion of "*yes*" responses to related and unrelated lures were compared using a two-way ANOVA with probe type (related, unrelated) and modality (auditory, visual) as repeated within-subject factors. There was a main effect of probe type, $F(1, 30) = 82.76$, $p < .001$, $\eta_p^2 = .73$ that showed more errors were made to related lures than to unrelated probes. A main effect of modality $F(1,30) = 2.52$; $p = 0.12$; *n.s.* was not significant, indicating the modality through which words were presented had no effect on false recognition. The interaction between probe type and modality was also not significant $F(1,30) = 0.3$; $p = 0.59$; *n.s*, indicating that neither false alarms to related or unrelated lures differed between modalities $t(30) = 1.65$; $p = 0.1$.

	Studied	H-T correction	Related	H-T correction -	Unrelated
Visual	0.72(0.04)	0.64(0.04)	0.41(0.04)	0.33(0.04)	0.08(0.02)
Auditory	0.76(0.03)	0.66(0.05)	0.47(0.04)	0.37(0.05)	0.1(0.04)

Table 4. Proportion of "*yes*" responses for each of the three probe types in each modality in LTM in Experiment 2. H-T correction = data corrected for baseline using the high-threshold procedure (*Pr*). Standard error of the mean in parentheses.

Comparison of Experiment 1 and 2

The corrected LTM results show no difference in false recognition between modalities when inclusion instructions are implemented, but standard instructions show that LTM is more accurate in the visual modality. Together, these results suggest that the type of instruction received during STM had an effect on long term false memories associated with auditory and visual presentation. In order to better understand how STM recall instructions affect STM performance, a 2 (instruction: standard vs inclusion) X 2 (modality: visual vs audio) mixed ANOVA was conducted to compare corrected STM recall rates between standard and inclusion instruction conditions. The results showed a main effect of instruction $F(1, 60) = 5.20$; $p = 0.03$; $\eta_p^2 = 0.1$, which means that participants who received the standard instruction recalled more studied words than those who received the inclusion instruction. There was also a main effect of modality $F(1, 60) = 51.12$; $p < 0.001$; $\eta_p^2 = 0.22$, meaning that participants recalled words they heard better than words they read, regardless of instruction. Quantitatively, words are most likely to be accurately recalled in the short term when presented in the auditory modality under the standard instruction, however the interaction of instruction and modality did not reach significance.

A similar ANOVA was computed to compare corrected (*Pr*) hit rates between standard and inclusion instruction conditions in LTM. This analysis revealed a main effect of modality $F(1, 60) = 6.09$; $p = 0.01$; $\eta_p^2 = 0.07$, indicating higher recognition accuracy following visual than auditory presentation, opposite of what was seen in STM.

False alarms to related words in the two instruction conditions were also compared. The 2 (modality: visual vs audio) x 2 (instruction: standard vs inclusion) mixed participants ANOVA was performed on (*Pr*) for related lures showed a main effect of instruction $F(1, 60) = 6.04$; $p = 0.02$; $\eta_p^2 = 0.09$ indicating that more errors were made under the inclusion instructions than under the standard instructions. The main effect of modality and the interaction were both not significant $F(1, 60) < 1$, meaning that there was no difference in long term false memories between the modalities. Even though the interaction between instruction and modality was not significant when only considering related lures, the data suggested that in the visual condition there more total errors when participants performed the STM phase under inclusion instructions as compared to a standard instruction $t(60) = 2.16$; $p = 0.03$; $d = 0.56$. Total errors include incorrectly recognizing related and unrelated words. For auditory conditions there was no effect of instruction on total errors $t(60) = 1.46$; $p = 0.14$, *n.s*. This suggests that the instruction given at the time of STM recall has a significant effect on visual false memories.

General Discussion

The two experiments reported in this thesis produced several patterns of importance. First, in STM, correct recall under standard instructions was better for items presented auditorily than visually, which is consistent with previous experiments that tested the STM-DRM through recognition (Olszewska et al., in press). While semantic intrusions were present in STM following presentation in either modality, more recall errors were generated when study items were presented visually rather than auditorily. Second, with standard instructions in STM, participants made more LTM errors following auditory than visual presentation; this modality effect is the opposite to the pattern seen in STM. However, when inclusion instructions were used in STM, the number of errors in LTM was similar for both modalities—in other words, the modality effect disappeared. Third, LTM accuracy was higher and led to more false recognitions for both modalities when STM encoding was performed with the inclusion instructions relative to the standard instructions. Each of these results will be considered in turn.

Experiment 1 replicated previous evidence for false STM using recall and extended these effects by demonstrating that visual study-test conditions are more prone to false recall than auditory conditions. That is, more false memories were generated during visual presentation than during auditory in short-term recall. Furthermore, this effect was generally reversed during LTM. These results indicate that testing the hybrid-DRM paradigm through recall in STM produces the same pattern of false memories that is seen when tested through recognition.

In Experiment 2, STM for studied words continued to be greater for the auditory than the visual condition, even with inclusion instructions. However, the frequency of associative responses generated did not differ between modalities. Furthermore, in LTM, the modality pattern was no longer evident: there was no difference between visual and auditory presentation in LTM for either accuracy of studied words or false recognition. In fact, the only modality difference evident in Experiment 2 was the increase in magnitude of accurately recognized words in LTM in the auditory modality relative to Experiment 1. This suggests that under the standard instructions, words presented auditorily may be encoded less deeply than those that are presented visually. The inclusion instructions may have promoted deeper encoding overall as indicated by greater correct recognition of studied items in the auditory modality and more false recognition of related items in both modalities (Flegal & Reuter-Lorenz, 2014). The greater improvement in the auditory condition may have occurred because deeper processing was less likely to be engaged automatically with the standard recall instructions in STM. However, this does not explain the results seen in Olszewska et al. (in press), in which more false memories occurred in LTM following auditory presentation, but it is quite possible that the use of recall during STM in the present study altered encoding, which would be a source of disparity between the two studies. Auditory memory was poorer overall compared to Olszewska et al. (in press), so recalling words out loud likely disrupted auditory memory, perhaps by introducing modality specific interference.

A few of the results in this thesis were unexpected. The first is that under standard instructions, the participants who heard words had poorer performance for correct recognition in LTM than those who read them. This was unexpected because in the Olszewska et al. (in press) paper, there was no difference between the two modalities in correct LTM recognition. The reason that a modality difference was seen in this study is probably because recall during STM required participants to say words out loud, which disrupted the auditory encoding, thus leading to fewer words correctly recognized in LTM following auditory presentation. To determine if this was the reason for the modality differences in correct LTM responses, a possible future study would be to have participants silently write down words during recall testing instead of vocalizing them. If this resulted in a similar number of correct recognitions between the modalities, it would support the previously mentioned reasoning. However, if the visual advantage persists, it would suggest that the modality difference is not due to the increased auditory noise. Another unexpected result is that under the inclusion instructions, a similar number of associated words were generated in the auditory and visual modalities. It cannot be assumed from this result that the two modalities use similar relational encoding processes, however, because the strategy subjects adopted to remember words through the standard instructions is not necessarily the same one they used for the inclusion instructions. A followup study in which subjects' brains are imaged while participating in a similar set of experiments would determine if different encoding strategies are used depending upon which instruction a participant receives. It has already been shown that tasks which elicit different semantic processing demands show non-overlapping activation patterns, suggesting different encoding strategies are used (Binder, Desai, Graves, & Conant, 2009). Therefore, it would be of value to show that different activation patterns occur when the standard instructions are given, versus the inclusion instructions. This would show they promote different encoding strategies, supporting the idea that the similar number of added associates following visual and auditory presentation does not indicate similar levels of relational encoding.

References

Arndt, J., & Reder, L. M. (2003). The effect of distinctive visual information on false recognition. *Journal of Memory and Language*, *48*, 1-15.

Atkins, A. S., & Reuter-Lorenz, P. A. (2008). False working memories? Semantic distortion in a mere 4 seconds. *Memory & Cognition, 36,* 74-81.

Andrew, D. M., & Bird, C. (1938). A comparison of two new-type questions: recall and recognition. *Journal of Educational Psychology*, *29*(3), 175-193.

Baddeley, A. D. (1972). Retrieval rules and semantic coding in short-term memory. *Psychological Bulletin*, *78*(5), 379-385.

Baddeley, A. D., Thompson, N., & Buchanan, M. (1975). Word length and structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575-589.

Bartlett, E. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge, UK: Cambridge University Press.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional and neuroimaging studies. *Cerebral Cortex*, *19*, 2767-2796.

Coane, J., McBride, D., Raulerson, B., & Jordan, J. (2007). False memory in a short term memory task. *Journal of Experimental Psychology, 54*, 62-70.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671-684.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58* (1), 17-22.

Dewhurst, S. A. (2001). Category repetition and false recognition: effects of instance frequency and category size. *Journal of Memory and Language*, *44*(1), 153-167.

Dewhurst, S. A., & Anderson, S. J. (1999). Effects of exact and category repetition in true and false recognition memory. *Memory & Cognition*, *27*(4), 665-673.

Flegal, K. E., Atkins, A., & Reuter-Lorenz, P. (2010). False memories seconds later: The rapid and compelling onset of illusory recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1331-1338.

Flegal, K. E., & Reuter-Lorenz, P. (2014). Get the gist? The effects of processing depth on false recognition in short-term and long-term memory. *Memory & Cognition, 42*, 701 – 711.

Hanawalt, N. G. (1937). Memory trace for figures in recall and recognition. *Archives of Psychology (Columbia University)*, *216*, 89.

Hege, A. C. G., & Dodson, C. S. (2004). Why distinctive information reduces false memories: Evidence for both reduced relational processing and the distinctiveness heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 787-795.

Johnson, M. K., Hashtroudi, S. & Lindsay, D. S. (1993). Source Monitoring. *Psychological Bulletin*, *114*, 3-28.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *4*(1), 19-31.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: an example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*(5), 585-589.

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*, *25*(12), 720-725.

Maylor, E. A., & Mo, A. (1998). Effects of study-test modality on false recognition. *British Journal of Psychology*, *90*, 477-493.

Mazzoni, G. (2002). Naturally occurring and suggestion-dependent memory distortions: convergence of disparate research traditions. *European Psychologist*, *7*, 17-30.

Olszewska, J. M., Reuter-Lorenz, P. A., Munier, E., & Bendler, S. A. (in press). Misremembering what you see or hear: dissociable effects of modality on short- and long-term false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Postman, L., Jenkins, W. O., & Postman, D. L. (1948). An experimental comparison of active recall and recognition. *The American Journal of Psychology*. 511-519.

Robinson, K. J. and Roediger, H. L. (1997). Associative processes in false recall and false recognition. *Psychological Science, 8*, 231-237.

Roediger, H. L. (1996). Memory illusions. *Journal of Memory and Language, 35*(2), 76-100.

Roediger, H. L. and McDermott, K. B. (1995). Creating False Memories Remembering Words Not Presented in Lists. *Journal of Experimental Psychology*, *21*, 803-814.

Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschild, L., Fung, N. S., & Schwartz, M. A. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory & Cognition*, *30*(7), 1054-1064.

Smith, R. E. & Hunt, R. R. (1998). Presentation modality affects false memory. *Psychonomic Bulletin & Review*, *5*, 710–715.

Takarangi, M. K., Parker, S., & Gary, M. (2006). Modernizing the misinformation effect: the development of a new stimulus set. *Applied Cognitive Psychology*, *20*(5), 583-590.

Van Damme, I., & Smets, K. (2014). The power of emotion versus the power of suggestion: memory for emotional events in the misinformation paradigm. *Emotion*, *14*(2), 310.

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, *70*, 122-129.