2007

Source memory and the picture superiority effect

Noelle L. Brown
Louisiana State University and Agricultural and Mechanical College

SOURCE MEMORY AND THE PICTURE SUPERIORITY EFFECT

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Arts

in

The Department of Psychology

by
Noelle L. Brown
B.S., University of Central Florida, 2003
December, 2007

TABLE OF CONTENTS

ABSTRACT

Two experiments were conducted to explore whether a picture superiority effect exists in source memory. To investigate this issue, participants studied a mixed list of pictures and words. Experiment 1 tested people's memory for an organizational source where half the pictures and words were studied on the left or right side of a computer monitor. In Experiment 2 an associative source was tested. During encoding half of the pictures and words were associated with a female voice and the other half with a male voice. At test, participants' memory for the location or voice of the pictures and words was assessed. On the memory test each participant saw half of the items represented in the same format (picture/word) they studied and the other half were shown in the other format. The results showed the typical picture superiority in item recognition; however, this effect was only found in source memory for the organizational source. There was also evidence that source memory is better when the study and test formats matched, but this effect of format was not found in item recognition. These results indicate that some manipulations may affect item recognition and source memory differentially.

INTRODUCTION

The difference in memory for pictures versus words has been investigated in memory research over the last few decades, with the general conclusion being that pictures are better remembered as compared to words (e.g., Madigan, 1974; Paivio & Csapo, 1973; Shepard, 1967). This advantage is generally referred to as the picture superiority effect. This phenomenon has been observed in both recall and recognition memory suggesting that the effect is rather robust. This study represents an exploration of the picture superiority effect in source memory. Source memory generally refers to the accessibility of qualitative details of memories (Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 1981). For example, memories may include perceptual and sensory information, semantic information, as well as other qualities. The amount and type of such qualitative information may be used in memory judgments, such as recall, recognition, and most commonly in judgments concerning the original source of a memory (e.g., read vs. heard, imagined vs. perceived). The primary goal of this study was to better understand how the availability of these qualitative characteristics may differ for concepts learned in a pictorial format as opposed to a printed word format.

The extant data pertaining to the memory differences between pictures and words are documented in different areas of memory research, but these differences have not been extensively explored within the source monitoring framework (Johnson et al., 1993). One motivating theoretical question is whether these qualitative features are better bound to pictures than words, perhaps contributing to the picture superiority effect. There are some data that bear on this question. Park and colleagues conducted a series of studies that explored the picture superiority effect in the context of aging. For the present purposes, I am focusing only on the conditions where younger adults were tested. For example, Park and Mason (1982) manipulated the background color of words and pictures in addition to the spatial location of words and pictures. In conditions where people were not expecting a test of spatial or color memory, they nonetheless showed more accurate memory for spatial location of pictures as opposed to words. Memory for color did not differ for pictures versus words, although memory for color was close to chance levels for both types of stimuli in these conditions. In another study, Park and Puglisi (1985) found that color memory for pictures was much better than for words, with color memory for words in this study close to floor under incidental source memory instructions. One notable methodological difference between the two studies is that the pictures in the Park and Puglisi study were painted to represent the color being part of the object (e.g., a yellow colored chair). Therefore, these pictures were more likely to represent ecologically valid stimuli. The pictures and the words in the Park and Mason study were white line drawings laid over a color patch on a slide. The authors suggested that presenting color as an intrinsic part of a picture makes it less effortful to encode and therefore may result in better memory for those items. In yet another study, Park, Puglisi, and Sovacool (1983) demonstrated better spatial memory for pictures as opposed to words, regardless of whether people expected to be tested about spatial location. Thus, across these studies, spatial memory was consistently better for pictures, whereas the results for color memory were more equivocal. The fact that color memory was always on floor for words in both the Park and Puglisi and the Park and Mason studies is also somewhat problematic. The present study represents another attempt to observe source memory for pictures versus words where the levels of performance are above chance.

A more general motivating question comes from work by Glanzer, Hilford and Kim (2004). They suggested that variables affecting item recognition should also affect source memory in a positively correlated manner. In a series of experiments these authors were able to show that source memory increased when they used manipulations that increased item recognition. For example when a level of processing manipulation was used item recognition and source memory for voice were better for items that were encoded deeply compared to items that were encoded more shallowly. Given that pictures produce a general memory advantage in item recognition, the Glanzer et al. (2004) hypothesis would predict an advantage for source detail bound to pictures as well. In addition to the consistent spatial memory advantage found by Park and colleagues, pilot data from our lab has shown that there is evidence of a picture superiority effect in source memory for list membership. However, we used a mixed list consisting of both pictures and words at study in the pilot data. The benefit of studying items as pictures in source memory was only significant when the test cues were presented as pictures and not when they were presented as words. These results may suggest that source memory for pictures is better than words only when the retrieval conditions match those of encoding. Notably, all of the aforementioned studies by Park and colleagues represent cases in which the test cues always matched the study format (e.g., if a picture was encoded, a picture cue was always provided at test). Park (1980) carried out a study where she crossed the type of studied material with what was presented on the test in a completely between-subjects design. Participants then made recognition judgments about either the same or different stimuli. Source memory was not measured in this study, however, the results showed better item recognition for items that were presented spatially over items that were presented with different background colors. Therefore, the present study also addresses whether study and test format are important in obtaining a source memory advantage for pictures.

Another motivating question is whether there are discrepancies regarding the picture superiority literature within the source monitoring framework. Some researchers have separated features that may be bound to an event or item in memory into separate source memory classes (e.g. Moscovitch, 1992; Troyer, Winocur, Craik, & Moscovitch, 1999; Troyer & Craik, 2000). Moscovitch (1992) determined that the amount of effort required to bind a feature to an item or event depends on the qualities of that source. Based on these findings Troyer and her colleagues classified sources into two separate categories. Associative qualities are more closely tied to the item and can help improve source memory without requiring a considerable amount of effort to encode the source. Features that are inherent to the way an item is presented are included in this class of sources such as color, voice and font. Organizational qualities are not as closely associated with the item and require more effort to encode these sources such as temporal order and spatial location (Moscovitch, 1992; Troyer et al., 1999; Troyer & Craik, 2000).

However, the results from Park and colleagues apparently do not fit these expectations. They found that source memory for spatial details was always better for pictures than for words, and also better than source memory for color. Moreover, source memory for color in that work benefited greatly from the knowledge that color would be tested, whereas source memory for spatial location increased very little (Park & Mason, 1982). Unless instructions specified that color memory was going to be assessed, memory for color was close to chance, even when color was presented as an intrinsic part of the stimuli. They interpreted these results to mean that encoding color information requires more effortful processing than encoding spatial information. These explanations are contrary to the hypothesis that spatial information is organizational and therefore should require more effort to encode as compared with color, which would be classified as an associative attribute. The Moscovitch (1992) scheme may also depend on the study-test format. For example, as discussed previously, pilot data from our lab only showed a temporal source memory advantage for pictures when pictures were also used as the test cues.

Participants in this study encoded two lists of pictures and words. On the test they were required to decide if an item was presented on List 1, List 2, or if it was a new item. This type of source falls into the organizational category and may explain why we did not obtain a picture superiority effect when the test cues were words.

Given the foregoing discussion, the goals for this study are threefold. The first is to test whether a picture superiority effect for source information depends on the amount of processing required to remember the source, as suggested by the findings of Park and colleagues. We chose to use voice as the associative attribute and spatial location as the organizational attribute. We chose voice partly because color memory in previous work was very near floor levels for word stimuli (e.g., Park & Mason, 1982) and also because the results with colored pictures depended much on methodological details (Park & Puglisi, 1983). Voice can be encoded along with the stimuli, but will not vary in the "amount" of perceptual detail the way that color might, such as how much color is in the stimulus or whether the color is in the background or the foreground of the stimulus. Second, aside from our pilot work, there are no studies that have manipulated study-test format in demonstrating picture superiority for bound source information. Park and colleagues always kept the test format consistent with the study format (i.e., word-word or picture-picture) when testing item *and* source memory. Third, this study will represent another test of the Glanzer et al. (2004) hypothesis that source memory should be better when bound to information that is better recognized.

In reference to the organizational/associative dimension, the predictions are ambiguous. Although Park and colleagues most consistently showed that an organizational attribute (i.e., spatial location) displayed superiority for pictures, their associative attribute (i.e., color) did not. In contrast, data from our lab showed that temporal order memory, an organizational attribute, displayed picture superiority only when pictures were also used as test cues. Considering these seemingly equivocal results, it is difficult to make a prediction based solely on the organizational/associative scheme developed by Moscovitch (1992). However, a tentative prediction regarding Glanzer et al.'s (2004) item-source correlation hypothesis and the study-test format manipulation can be made. First, we would expect pictures to produce better source performance than words regardless of the type of source, because pictures generally produce better item recognition. Second, several studies have shown better recognition performance when the encoding and retrieval conditions match (e.g., Morris, Bransford, & Franks, 1977). Therefore, for both words and pictures, when the test format is consistent with the study format as opposed to inconsistent, source memory for both voice and spatial location should improve.

Experiment 1 was conducted to test our hypotheses for memory of an organizational source, spatial location, and in Experiment 2 we assessed memory for the associative source, voice. In both experiments we also evaluated item recognition for pictures and words.

EXPERIMENT 1

Method

Participants. Participants for Experiment 1 were 127 Louisiana State University psychology students. Participants volunteered for extra credit or to fulfill a course research requirement. There were 63 people in a picture label condition and 64 in a no label condition.

Materials and Design. One hundred four items were taken from a subset of the pictures and their corresponding labels used by Snodgrass and Vanderwart (1980). All of the pictures were simple black and white line drawings of concrete nouns that the authors established norms for familiarity, name agreement, visual complexity and image agreement. An example of a picture without a label is shown in Figure 1 and a picture with a label is shown in Figure 2.

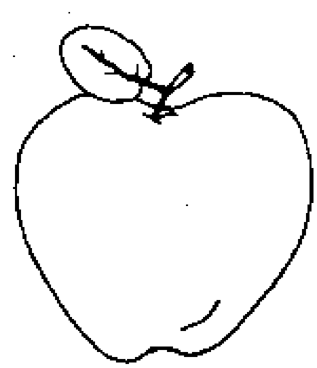
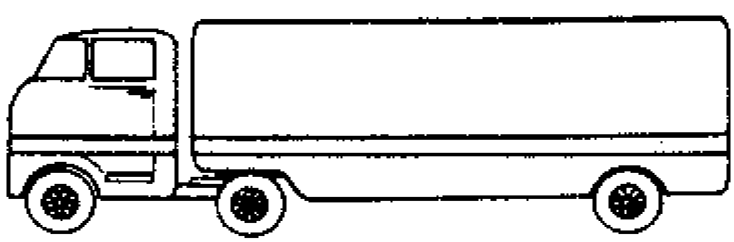


FIGURE 1: No Label Condition.



TRUCK

FIGURE 2: Label Condition.

Sixty-four items were used in the study phase, half words and half pictures. Eight additional items were fixed as buffer items at the beginning and end of the study list. The memory test consisted of all of the old items along with 32 new items. The non-buffer items were grouped into six lists of 16 items to ensure they were presented as pictures or words at encoding, pictures or words at retrieval, or new an equal number of times. Each list of 16 had roughly the same familiarity, name agreement, visual complexity and image agreement. Half of the items were shown on the left side of the computer screen and the other half on right side of the screen. All of the items were counterbalanced across participants such that items that appeared on the left side of the screen appeared on the right side an equal number of times.

To assess whether picture naming affects the encoding or retrieval process, close to half of the participants were provided with labels. Pictures were presented at the top of the screen with their corresponding word label underneath. However, this manipulation could influence the distinctiveness of the picture format. Pictures have two visual codes, the actual picture and the label, whereas words only have one. Thus, the remaining participants did not see pictures with a verbal label. These two conditions were tested across subjects and will be referred to as the label and no label conditions from this point forward. All participants were randomly assigned to the label and no label conditions. The words in both conditions were positioned on the bottom of the screen (on the left or the right side) in Courier New font. The labels for the pictures in the label condition were in the same font and location on the screen as items presented as words. Pictures and words will be referred to as item type and was manipulated within-subjects. There were two words and two pictures at the beginning and end of each list that served as buffer items to help minimize any recency and primacy effects. The third variable of interest was congruence between study and test format, and was also a within-subjects manipulation. Half of the studied pictures and half of the studied words appeared in the same format at test (picture-picture or word-word), with the remaining items appearing in the other format (picture-word or word-picture). An equal number of items from each source (left/right) were represented in each study-test format condition.

The memory test was created to asses both recognition memory and source memory. The test items were presented either as words or as pictures in the center of the screen with the source and recognition choices (Left, Right, or New) below them. Picture test items in the label condition were viewed as pictures with their corresponding verbal labels underneath. Picture test items in the no label condition were tested only as pictures.

Procedure. Only one participant was tested at a time. Each participant viewed one study list with half of the items presented as words and the other half presented as pictures either with the verbal labels written underneath or only the picture. Half of the items appeared on the left side of the screen and the other half on the right side of the screen. Each item appeared at a rate of 2750 ms with 250ms interstimulus trial. Before presentation of the items participants were instructed to remember each item to the best of their ability because their memory for them will be tested later. The instructions for the encoding phase did not specify that their memory for the location of each item would also be tested. Item presentation was randomized anew for each participant.

After the encoding phase participants took part in a three minute filler task where they were asked to name as many United States capital cities as possible. Paper and pencil were provided for participants to record their answers. This task was only to attempt to eliminate any recency effects and the results from this task will not be included in any reported analyses. After three minutes the instructions for the memory test appeared on the computer screen.

The memory test was self paced and the items were tested randomly. The test included all 64 studied items plus 32 new items. Each participant was required to decide which source (left/right) the item was presented in or if it was a new unstudied item. They were instructed to call items old if they were studied in any form, regardless of the test format. After participants finished the memory test they were debriefed and thanked for their participation.

Results

In this and any subsequent analyses, I used an alpha level of .05. There were no meaningful differences between the label and no label groups. Therefore all of the following analyses are collapsed across the label conditions. Recognition performance was slightly better for items presented on the right; however there was no bias for participants to attribute new items to the left or the right. Seven participants were dropped from the analyses because their recognition performance was below chance.

Recognition Memory. To obtain a good measure of recognition memory the hit rate, false alarm rate, and corrected measures of recognition for pictures and words were calculated. The hit rate was calculated as the total number of items reported as old divided by the total number of old items in that condition. The false alarm rate was the proportion of new items incorrectly called old divided by the total number of new items. Because the design was completely within-subjects, the word false alarm rate is the same for the picture-word and word-word conditions and the picture false alarm rate is the same for the picture-picture and word-picture conditions. I then calculated difference scores and *d*' values as measures of corrected recognition. As seen in Table 1, the conditions are listed first by the study format followed by the test format. For example, items studied as words and tested as pictures are represented as WP. The hit rates were subjected to a 2 (study format: picture vs. word) × 2 (test format: picture vs. word) within subjects analysis of variance (ANOVA). As expected, recognition memory for items studied as pictures was significantly better than recognition for items studied as words, F (1, 119) = 150.74, $p < .05$, $\eta_p^2 = .559$. There was also a significant main effect of test format, F (1, 119) = 40.36, $p < .05$, $\eta_p^2 = .253$, favoring studied items tested as pictures. However, the main effects were qualified by a significant interaction, F (1, 119) = 22.32, $p < .05$, $\eta_p^2 = .158$, due to better recognition of items studied as pictures and tested as pictures than items studied as pictures and tested as words. Post hoc comparisons confirm that the PP produced higher hit rates than the PW condition, t (119) = 9.05, $p < .05$, whereas performance in the WW and WP

TABLE 1: Organizational Recognition Memory Results. HR = hit rate; FAR = false alarm rate; Diff = Difference score. Standard deviations are presented in parentheses.

	HR	FAR	Diff	*d'*
PP	.92 (.09)	.30 (.21)	.62	1.97 (.76)
PW	.81 (.15)	.21 (.19)	.60	1.86 (.87)
WP	.70 (.20)	.30 (.21)	.40	1.16 (.83)
WW	.69 (.18)	.21 (.19)	.48	1.41 (.77)

conditions did not differ, t (119) = -1.21, $p > .05$. The results support prior findings in showing the typical picture superiority effect, though they do not seem to support a dual coding view of the picture superiority effect. According to the dual coding hypothesis pictures are automatically processed in a visual and semantic manner creating both a perceptual and conceptual code. Therefore pictures should be protected against any cost of changing the test format from the study format whereas words should suffer (Mintzer and Snodgrass, 1999). Our results show greater costs associated with a format change for pictures and no cost of a format change for studied words. These results will be fully addressed in the discussion section.

A t-test was carried out to compare the false alarm rates for picture and word items. The results showed that the false alarm rate for pictures was significantly higher than the false alarm rate for words $t\,(119) = 6.45$, $p < .05$. These results are interesting if one considers that the picture superiority effect is partially due to pictures being more distinct than written words. If pictures are truly more distinct than written words, then one would expect the false alarm rate for pictures to be lower than that of words. Others have found similar results in the false alarm rates. Both Mintzer and Snodgrass (1999) and Stenberg, Radeborg, and Hedman (1995) found higher false alarm rates for pictures than words when they used mixed study lists. Using pure study lists, only words or only pictures, reversed the trend of the false alarm rates. Studying pure lists resulted in higher false alarm rates for words than pictures, creating a mirror effect and showing evidence that pictures are more distinct than words. Mintzer and Snodgrass (1999) suggest that the reversed mirror effect in the mixed list design is due to participants adopting a strategy based on the familiarity of the test items. Under these circumstances new pictures may not evoke the same amount of familiarity as old pictures, but they may evoke enough familiarity to be attributed to an old item studied as a word. Therefore participants' higher false alarm rate for pictures is based on the belief that they were studied as words and not pictures. Using this strategy the word false alarm rate should not vary between items studied as pictures or words because the advantage that studied words may have by being tested in the same format they were studied can be offset by the distinctiveness of studied pictures. Mintzer and Snodgrass (1999) were able to show support for this notion when they had participants study pure lists.

d' scores were also subjected to a 2 (study format: picture vs. word) × 2 (test format: picture vs. word) within subjects ANOVA. Once again there was a significant effect of study format, $F\,(1, 119) = 184.70$, $p < .05$, $\eta_p^2 = .608$; however, the main effect of test format was not significant, F (1, 119) = 1.80, p = .183, η_p^2 = .015. The interaction was significant, F (1, 119) = 30.69, p < .05, η_p^2 = .205. Post hoc comparisons show that the PP and PW conditions were marginally different from each other, t (119) = 1.67, p = .09, and the WW and WP conditions were significantly different from each other, t (119) = 3.96, p < .05. These results are similar to the results obtained in the hit rates; however, the effect of test format is no longer significant. Although the interaction was also significant, the pattern is different from the hit rate results. Corrected recognition data show that the cost of changing the test format is greater for the items studied as words than items studied as pictures. This is predicted by the dual coding view, but as previously mentioned may be a result of the design. As can be seen in Table 1, the high false alarm rate for items tested as pictures and lower false alarm rate for items tested as words greatly reduced the difference between items studied as pictures, but increased the difference between the items studied as words. If pure study lists were used the *d'* analysis may be more reminiscent of the hit analysis in showing support for a distinctiveness view of the picture superiority effect. Inferential statistics are not reported on the difference scores because the interpretation of the results was the same as the *d*' analysis.

The recognition results show a general advantage of studied pictures over words and the effect of changing the test format differs depending on the measure used. Evaluating the cost on hits of changing the test format shows a greater cost to pictures. On the other hand, evaluating the cost of changing the test format using corrected recognition shows a greater cost to words.

Source Memory. Source memory was calculated as a conditionalized score: The proportion of words or pictures presented on the left or right side of the screen that were correctly attributed to the correct location divided by the total number of items identified as old and are reported as an Average Conditional Source Identification Measure (ACSIM). There was no bias to claim one source over the other, but source memory was slightly better for items presented on the right. This difference was significant; however, the ACSIM scores are averaged across items presented on the left and right. Approximate hit and false alarm rates were calculated for items presented on the left. The hit rates are the number of items presented on the left and correctly called left (L_L) plus 0.5 divided by the number of items presented on the left identified as old plus 1. The false alarm rate was then calculated as the number of items presented on the right and incorrectly identified as left items (R_L) plus 0.5 divided by the number of items on the right identified as old plus 1. These corrections were done to correct for instances where the hit rate for any participant may have been 1 and/or a false alarm rate equal to 0.0 (Macmillan and Creelman, 1991). Difference and *d'* scores were then calculated from these hit and false alarm rates. Table 2 summarizes the source memory results.

TABLE 2: Organizational Source Memory Results. Standard deviations are reported in parentheses.

	ACSIM	L_L	R_L	Diff	*d'*
PP	.79 (.16)	.74 (.18)	.23 (.15)	.51	1.60 (.96)
PW	.77 (.16)	.71 (.17)	.19 (.13)	.52	1.61 (.88)
WP	.64 (.17)	.59 (.19)	.36 (.20)	.23	.67 (.84)
WW	.67 (.17)	.60 (.20)	.31 (.19)	.29	.88 (.92)

A within-subjects ANOVA was conducted on the ACSIM scores. This analysis revealed a main effect of study format, F (1, 119) = 85.56, $p < .05$, $\eta_p^2 = .418$, showing an advantage in source memory for items studied as pictures. The main effect of test format was not significant, F (1, 119) = .503, $p = .480$, $\eta_p^2 = .004$, suggesting there was no difference in source memory when test items were pictures or words. The interaction between study and test format was significant, *F* (1, 119) = 6.70, $p < .05$, $\eta_p^2 = .053$. Post hoc comparisons reveal that the difference between the PP and PW conditions was not significant, *t* (119) = 1.34, $p > .05$; however, the difference between the WW and WP conditions was significant, t (119) = 2.01, $p < .05$. The results show a picture superiority effect in source memory for spatial location. The results also show some support for transfer appropriate processing: matching test & study formats show better source memory than when the two conditions do not match. This effect was not observed in the recognition data.

A within-subjects ANOVA was also conducted on the source memory *d'* scores. This analysis replicated the ACSIM results. There was a main effect of study format, *F*(1, 119) = 131.95, $p < .05$, $\eta_p^2 = .526$, no effect of test format, *F*(1, 119) = 2.90, p = .09, $\eta_p^2 = .024$, and the interaction was marginally significant, *F*(1, 119) = 3.01, p = .09, $\eta_p^2 = .025$. Post hoc comparisons show that source memory for items encoded as words is significantly different when the test format is changed, *t* (119) = 2.33, $p < .05$. This effect was not significant for items encoded as pictures *t* (119) = -.10, $p > .05$. The *d'* results, like the ACSIM results, show an advantage for pictures in source memory; however, there is no longer evidence for transfer appropriate processing for items studied as pictures, although the effect persists for items studied as words.

In sum, the results show an advantage for pictures in both recognition memory and source memory, supporting the Glanzer, et al. (2004) hypothesis that item recognition and source memory are positively correlated. The results also show memory well above chance for an organizational source, similar to that obtained by Park and colleagues.

EXPERIMENT 2

The purpose of Experiment 2 was to replicate the findings of Experiment 1 and to compare the results obtained from an organizational source with that of an associative source. For that reason the source in Experiment 2 was changed to male and female voices.

Method

Participants. Participants for Experiment 2 were 130 Louisiana State University psychology students who volunteered for extra credit or to fulfill a course research requirement. There were 62 people in the label condition and 68 in the no label condition.

Materials and Design. The materials and design were the same as in Experiment 1 with the exception of the source. Half of the items were presented in a digitized male voice and the other half were presented in a digitized female voice. All of the items were visually displayed in the middle of the screen with pictures located on the top portion and words on the bottom portion. All of the items were counterbalanced across participants such that items that were presented in a male voice were also spoken in a female voice an equal number of times.

I again manipulated the presence of the verbal label between subjects to which participants were randomly assigned. The study-test format manipulation was the same as Experiment 1. The memory test was also the same as experiment one except that the source question evaluated voice and not spatial location.

Procedure. The general procedures were the same as Experiment 1, with the exception of source presentation. Participants heard the male and female voices over a set of headphones connected to the computer. The presentation of the items was slightly different from Experiment 1. All of the items were horizontally centered on the screen with pictures on the top portion and words on the bottom. The rate at which items appeared remained the same as well as the filler task. The test procedures were the same as well, with the exception of the source that was tested.

Results

Again, there were no meaningful differences between the label and no label conditions so all of the following analyses were collapsed across this factor. Recognition performance was the same for female and male items and participants were not biased to call new items male or female. One participant was dropped from the analyses because his/her recognition performance was below chance and two others were removed because they did not follow directions.

Recognition Memory. The hit rates, false alarm rates, and corrected measures of recognition were calculated in the same way as Experiment 1. The means for these measures are reported in Table 3.

TABLE 3: Associative Recognition Memory Results. HR = hit rate; FAR = false alarm rate; Diff = Difference score. Standard deviations are presented in parentheses.

	HR	FAR	Diff	*d'*
PP	.90 (.11)	.34 (.22)	.56	1.77 (.75)
PW	.78 (.16)	.24 (.21)	.54	1.68 (.82)
WP	.72 (.17)	.34 (.22)	.38	1.08 (.73)
WW	.70 (.16)	.24 (.21)	.46	1.45 (.72)

A 2 (study format) × 2 (test format) within subjects ANOVA was run on the proportion of correctly recognized items. Recognition memory for items studied as pictures was significantly better than recognition for items studied as words seen in the main effect of study format, F (1, 126) = 126.18, $p < .05$, $\eta_p^2 = .500$. There was also a significant main effect of test format, F (1, 126) = 32.15, $p < .05$, $\eta_p^2 = .203$. Once again, the latter outcome seems to suggest that there is a benefit for items to be tested in a picture format. These findings replicate the results of

Experiment 1 and were also qualified by a significant interaction, F (1, 126) = 29.33, p < .05, η_p^2 = .189. Post hoc comparisons on the hit rates show that performance in the PP condition was significantly better than performance in the PW condition, t (126) = 9.27, p < .05 and performance between the WW and WP conditions did not differ significantly, t (126) = -.59, p > .05. These findings replicate the recognition results from Experiment 1 and others (e.g. Mintzer and Snodgrass, 1999).

A t-test was carried out on the false alarm rate for pictures and words. The results replicate those of Experiment 1, the false alarm rate was significantly greater for pictures than the false alarm rate for words, t (126) = 5.77, p < .05, in contrast to a distinctiveness view of the picture superiority effect. Again, it is possible that this result can be attributed to participants adopting a strategy based on the familiarity of the test items, as in Experiment 1.

A 2 (study format) × 2 (test format) within subjects ANOVA was conducted on the *d'* scores. The main effect of study format was significant, F (1, 126) = 131.51, p < .05, η_p^2 = .511. Unlike Experiment 1, main effect of test format was also significant, F (1, 126) = 4.64, p < .05, η_p^2 = .033. These results are similar to the hit rate results, except the grounds for the difference are different. As can be seen in Table 3, items tested as pictures have, on average, lower d' values (*M* = 1.43) and therefore participants were less accurate at recognizing these items than items tested as words (*M* = 1.57). The pattern of results is the same as Experiment 1 even though the main effect of test format was not significant for d' scores in Experiment 1. However, these main effects were qualified by a significant interaction, F (1, 126) = 48.99, p < .05, η_p^2 = .280. Post hoc comparisons show that the picture study conditions were not significantly different from each other, t (126) = 1.13, p > .05, and the word study conditions were significantly different from each other, t (126) = 5.13, p < .05. The interaction also replicates the results from

Experiment 1, showing that the cost of changing the test format from the study format is greater for items studied as words than the change format costs associated with studied pictures. As previously mentioned, our design may be the reason for the high false alarm rate for items tested as pictures and lower false alarm rate for items tested as words. These false alarm rates greatly reduced the difference between items studied as pictures, and increased the difference between the items studied as words. If pure study lists were used the *d*' analysis may be more comparable to the hit analysis showing support of a distinctiveness view of the picture superiority effect. The analysis of the difference scores was the same as that of the *d*' analysis, therefore the results are not reported here.

Overall, the recognition results of Experiment 2 replicate those of Experiment 1. The picture superiority effect was replicated. Again, the interpretation of the effect of changing the test format differs depending on the measure used. Evaluating the cost on hits of changing the test format shows a greater cost to pictures. On the other hand, evaluating the cost of changing the test format using corrected recognition shows a greater cost to words.

<u>Source Memory.</u> Source memory was calculated in a similar fashion as in Experiment 1. ACSIMs are the proportion of words or pictures presented in a male voice or female voice that were correctly attributed to the correct voice divided by the total number of items identified as old. There was no bias to claim one source over the other, and source memory performance was the same for male and female items. Approximate hit and false alarm rates were calculated for items presented in a male voice. The hit rates are the number of items presented by a male and correctly called male plus 0.5 divided by the number of items presented by a male and identified as old plus 1 (M_M). The false alarm rate was then calculated as the number of items presented by the female and incorrectly identified as male plus 0.5 divided by the number of items

presented in the female voice and identified as old plus 1. Difference and *d'* scores were then calculated from these hit and false alarm rates. Table 4 summarizes the voice source memory results.

A within-subjects ANOVA was conducted on the ACSIM scores. This analysis did not reveal any significant effects, $F\ (1, 126) < 1$, $p > .05$, $\eta_p^2 = .001$ for study format, $F\ (1, 126) < 1$, $p > .05$, $\eta_p^2 = .001$ for test format, and $F\ (1, 126) = 2.13$, $p = .15$, $\eta_p^2 = .017$ for the interaction. These results show that there was no advantage for items studied as pictures in voice memory, nor was there an effect of changing the study format at test. However the means in Table 4 show a trend of transfer appropriate processing.

TABLE 4: Associative Source Memory Results. Standard deviations are reported in parentheses.

	ACSIM	M_M	F_M	Diff	*d'*
PP	.57 (.14)	.56 (.17)	.44 (.17)	.12	.32 (.69)
PW	.55 (.16)	.57 (.19)	.48 (.18)	.09	.27 (.78)
WP	.55 (.15)	.52 (.17)	.44 (.20)	.08	.24 (.73)
WW	.57 (.17)	.57 (.19)	.47 (.18)	.10	.30 (.79)

A within-subjects ANOVA was also conducted on the source memory *d'* scores. This analysis replicated the ACSIM results. There were no significant effects**,** $F\ (1, 126) < 1$, $p > .05$, $\eta_p^2 = .002$ for study format, $F\ (1, 126) < 1$, $p > .05$, $\eta_p^2 = .000$ for test format, $F\ (1, 126) = 1.25$, $p > .05$, $\eta_p^2 = .010$ for the interaction**.** The *d*' results can be interpreted in the same way as the ACSIM results.

Contrary to the results from Experiment 1, the picture superiority effect was only found in recognition memory and not source memory. Also, the results from this experiment do not support the Glanzer, et al. (2004) hypothesis that improving item recognition will also improve source memory. These results, together with the results from Experiment 1, contradict the Troyer and colleagues' scheme of source memory. This scheme predicts that source memory should be better for associative sources than organizational sources because they require less effort to encode and are an intrinsic part of the item itself (Troyer and Craik, 2000; Troyer, et al., 1999). The current results show the opposite of these predictions. Source memory for spatial location, an organizational source, was well above chance, and almost as good as item recognition. Source memory for voice, an associative source, was close to floor. An explanation of these results is given in the discussion.

DISCUSSION

The purpose of the present experiments was (a) to explore whether a picture superiority effect for source information is dependent upon the amount of processing required to remember the source, (b) to observe any effects of a study-test format manipulation on both recognition and source memory, and (c) to investigate the hypothesis that source memory should be better bound to information that is better recognized (Glanzer, et al., 2004). An advantage for pictures was found in recognition memory, supporting the typical picture superiority effect; however, this advantage was not found for both sources. The effect of study-test format showed that recognition memory tended to be better for items tested as pictures regardless of study format, where as source memory was improved when the test format matched the study format. Together the results contrast the hypothesis that improving item memory also improves source memory. If the latter were the case we should have found an advantage for pictures for both sources as well as found better source memory for items tested as pictures. Each of these results will be discussed further in terms of recognition memory and source memory.

Recognition

The recognition results from both experiments show a clear picture superiority effect supporting previous results (e.g., Paivio, 1971, 1975; Park, 1980). This pictorial advantage in item memory was observed regardless of the type of source information participants encoded. There were also no differences in recognition performance for items that were presented as associative sources or organizational sources.

According to dual coding theory the advantage of pictures is the result of two sources: the first is that pictures are automatically encoded in a verbal and image form, the second is that the image code is naturally stronger and more memorable than the verbal code (Paivio & Csapo, 1973; Snodgrass & McClure, 1975). However, dual coding would also predict that the study-test format manipulation should have a greater effect on items studied as words than items studied as pictures. If pictures are automatically encoded in a verbal form, then changing the test format to a verbal form should not affect memory for these items. The hit rate results do not support the dual coding view of the picture superiority effect. The hit rates in both experiments dropped significantly when items studied as pictures were tested as words from when they were tested as pictures. Items studied as words were relatively unaffected by the form change manipulation. However, these results may be explained by participants showing a bias towards calling picture test items old. This bias should systematically increase the hit rate of old items tested as pictures. Therefore, the PP hit rates and WP hit rates should be falsely inflated (Snodgrass & McClure, 1975). The inflated WP hit rates result in what seems to be better memory for these items than WW items and also mask any possible effects of studying and testing items in the same format, (i.e. transfer appropriate processing). Nonetheless, it can still be argued that the image encoding of pictures predicted by dual coding should make them more memorable and distinguishable from new pictures. Also, storing a verbal code for pictures should still protect against the costs of changing the test format to words.

The *d'* results, on the other hand, do seem to support the form change costs predicted by dual coding. The costs were greater for items encoded as words than items encoded as pictures; however, as previously mentioned, this result can be attributed to the uncharacteristically high false alarm rate for items tested as pictures compared to that of items tested as words. This is caused by the same bias observed in the hit rates, towards calling new items pictures. While the inflated false alarm rates for picture test items mask the effect of the form change on picture study items it exaggerates the effect on word study items. For this reason, Mintzer and Snodgrass

(1999) assessed peoples' memory for pictures and words using a between subjects paradigm. Participants studied pictures and words between lists and their memory for these items were tested in both formats. This design allowed the authors to assess false alarm rates for each condition separately (PP, PW, WP, WW). The results from this experiment still showed an advantage in the hit rates for items tested as pictures; however, the corrected recognition data showed evidence of transfer appropriate processing and importantly a greater cost of changing the test format for picture study items than word study items. These results show that when the bias to call test items pictures is accounted for the results no longer support the dual coding view of the picture superiority effect. I also have unpublished data in which participants studied pictures and words between subjects that show a similar pattern of results. The data is shown in Tables 5 and 6.

TABLE 5: Organizational Recognition Memory Results for Unpublished Data. HR = hit rate; FAR = false alarm rate; Diff = Difference score. Standard deviations are presented in parentheses.

	HR	FAR	Diff	*d’*
PP	.90 (.09)	.17 (.17)	.73	2.40 (.82)
PW	.83 (.13)	.22 (.21)	.61	1.93 (.87)
WP	.81 (.11)	.51 (.25)	.30	0.90 (.61)
WW	.82 (.12)	.42 (.24)	.40	1.17 (.63)

TABLE 6: Associative Recognition Memory Results for Unpublished Data. HR = hit rate; FAR = false alarm rate; Diff = Difference score. Standard deviations are presented in parentheses.

	HR	FAR	Diff	*d'*
PP	.88 (.09)	.24 (.22)	.65	2.10 (.80)
PW	.73 (.15)	.19 (.18)	.54	1.87 (.81)
WP	.79 (.15)	.51 (.26)	.28	0.85 (.89)
WW	.78 (.13)	.40 (.23)	.37	1.24 (.94)

Another interpretation of the picture superiority effect is that it is caused by the distinctive features of pictures and/or more distinct processing of pictures during encoding. The significant difference between the dual coding and distinctiveness view is the effect of manipulating the study-test format. The distinctive sensory/semantic features of pictures should make them more dependent upon a same format test. Words, on the other hand, do not rely on distinctiveness of the memory codes and therefore should be less vulnerable to a format change at test (Mintzer & Snodgrass, 1999). Hunt (2006) also makes the argument that distinctive processing can only occur if the original processing is also applied at the time of test, in other words transfer appropriate processing must occur for distinctive items. The current hit rate results support this view; however, the corrected recognition results do not. As previously argued, the present corrected recognition results are most likely a product of the design and strategy that participants employed. It is assumed that participants discriminated new test items from old ones based on the familiarity of these items. If this were the case new pictures may induce enough familiarity to be incorrectly identified as old words, but not old pictures. The use of this strategy would cause higher false alarm rates for pictures based on the belief that they were originally studied as words. In fact, the Mintzer and Snodgrass (1999) and my unpublished data show support for this assumption. The hit rates *and* corrected recognition results from these experiments show that picture items rely more heavily on the similar processing at both encoding and retrieval than do word items.

Although the current results show an apparent advantage for items studied as pictures irrefutable evidence for the grounds for this effect cannot be drawn from this study alone. However, these results along with other studies support the distinctiveness view over dual coding.

Source Memory

The source memory results are less consistent across experiments than the recognition results. One of the concerns of this study was to address whether a picture superiority effect was dependent upon the amount of processing required to encode a source. The results show evidence of a picture superiority effect in memory for spatial location (organizational source), but not voice memory (associative source). These results imply that processing differences may be a moderating factor in how well source information is bound to pictures versus words. One explanation of these results comes from Troyer and colleagues' classification scheme of sources into organizational or associative groups. However, the present results contradict their predictions. According to their scheme organizational sources require more cognitive resources to encode and memory for these items should be reduced compared to memory for associative items. The existing results show superior memory for an organizational source compared to memory for an associative source. However, the current results support previous findings. Park and her colleagues consistently found an advantage of pictures over words in source memory for spatial location. This advantage was not found for color memory which would be classified as an associative source by Troyer's scheme. Troyer's source classification scheme is not supported by these results or the results from the current study. Though, there is one possible drawback with the manner in which the voice source was presented. Voice information was always encoded along with visual presentation of each item. Consequently, the visual information presented in the study may have received precedence and the auditory information may have been perceived as superfluous information and therefore disregarded. Associative sources are thought to require less effort to encode because they are intrinsic to the item itself. It can be argued that the current experiment is not a fair test of the organizational/associative classification scheme because voice presentation in this study was not intrinsic to the item and therefore should not be classified as an associative source.

Another objective was to observe any effects of a study-test format manipulation on source memory. The results from both experiments show evidence of transfer appropriate processing. Source memory was consistently better when items were tested in the same format as the format they were studied. This result was not obtained in recognition memory; however, there are two possible explanations for this difference. The first can be explained by the aforementioned bias to call items pictures. The second is that source memory requires more criterial evidence to call an item old than does item recognition. In other words, item recognition can be based on many different aspects of the encoding event where source memory is extremely specific. For this reason, source judgments may be more dependent upon the encoding and retrieval processes matching than item recognition.

The last goal of this study was to test the positive item-source correlation proposed by Glanzer et al. (2004). This hypothesis was not supported by the current results. First, transfer appropriate processing was observed in source memory but not item recognition. Again, this difference may be an artifact of the design. Second, the advantage of pictures over words was not found for both sources but it was found in item recognition regardless of source. To further test this hypothesis correlations were calculated for the hit rates and ACSIM scores. The results did not show any systematic relationships between recognition and source memory for any of the conditions. According to the source monitoring framework recognition and source monitoring vary in the amount and type of information used to make each judgment. Depending on the situational demands, the processes used to make recognition decisions and source judgments can vary considerably. The more these processes overlap the more likely it is that the two decisions will be correlated (Johnson, Hashtroudi, & Lindsay, 1993). The present results seem to suggest that the processing involved in spatial memory overlaps more with item recognition under the present conditions than does memory for a particular voice. Some considerations of these differences are considered next.

Evidence from several neuroimaging studies have shown that some of the areas that are recruited for item processing are also recruited for spatial location, whereas auditory processing occurs in different brain regions (Slotnick, 2004). Finke, Bublak and Zihl (2005) found that both left and right parietal lobes were involved in working memory (WM) for visual shapes, and right parietal lobes were engaged during visual spatial WM tasks. Together these results along with the current discrepancy in source memory performance between spatial location and voice memory support predictions from the source monitoring framework. These findings bolster the notion that source memory performance relies on an overlap of item processing and source processing.

Summary

In closing, the results from this study replicate those of previous findings in showing a picture superiority effect in recognition memory. The manipulation of study and test format helps to illuminate how the nature of retrieval cues affects source retrieval in general. These results imply that source memory is somewhat dependent upon the processes at encoding matching those at retrieval. This finding is important because it can be used to influence source memory performance in subsequent experiments. The main concern of this study was to examine whether or not there are differences in how the picture superiority effect will be present in other contextual attributes of the stimuli. Although a vast literature exists pointing to the generally better episodic memory for pictures over words, whether this advantage extends to orthogonal features connected to pictures and words has received much less attention (e.g., Park & Mason, 1982). The source memory results appear to suggest that an advantage for pictures will be present if the cognitive processes evoked during item encoding overlap with those used to encode the particular source. The present results do not directly confirm this assumption; however, the results echo those from neuroimaging studies (e.g. Slotnick, 2004). Further research in this area may be able to reveal some new and/or unexplored features of the picture superiority effect as well as to provide insight into the cause of the effect.

REFERENCES

Finke, K., Bublak, P., & Zihl, J. (2005). Visual spatial and visual pattern working memory: Neuropsychological evidence for a differential role of left and right dorsal visual brain. *Neuropsychologia, 44*, 649-661.

Glanzer, M., Hilford, A., & Kim, K. (2004). Six regularities of source recognition. *Journal of Experimental Psychology: Learning, Memory & Cognition, 30*, 1176-1195.

Hunt, R. R. (2006). The concept of distinctiveness in memory research. In R. R. Hunt and J. B. Worthen (Eds.), *Distinctiveness and Memory.* (pp.3-25) New York: Oxford University.

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

Johnson, M. K., & Raye, C. L. (1981). Reality Monitoring. *Psychological Review, 88*, 67-85.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.

Madigan, S. (1974). Representational storage in picture memory. *Bulletin of the Psychonomic Society, 4*, 567-568.

Mintzer, M. Z., & Snodgrass, J. G. (1999). The picture superiority effect: Support for the distinctiveness model. *The American Journal of Psychology, 112*, 113-146.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior, 16*, 519-533.

Moscovitch, M., (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience, 4*, 257-267.

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

Paivio, A. (1975). Coding distinctions and repetition effects in memory. In G. H. Bower (Ed.), *The psychology of learning and motivation*. (pp. 179-214). New York: Academic Press.

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

Park, D. C. (1980). Item and attribute storage of pictures and words in memory. *American Journal of Psychology, 93*, 603-615.

Park, D. C, & Mason, D. A. (1982). Is there evidence for automatic processing of spatial and color attributes present in pictures and words? *Memory & Cognition, 10*, 76-81.

Park, D. C., & Puglisi, J. T. (1985). Older adults' memory for the color of pictures and words. *Journal of Gerontology, 40*, 198-204.

Park, D. C., Puglisi, J. T., & Sovacool, M. (1983). Memory for pictures, words, and spatial location in older adults: Evidence for pictorial superiority. *Journal of Gerontology, 38*, 582-588.

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior, 6*, 156-163.

Slotnick, S. D. (2004). Visual memory and visual perception recruit common neural substrates. *Behavioral and Cognitive Neuroscience Reviews, 3*, 207-221.

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

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

Stenberg, G., Radeborg, K., & Hedman, L. R. (1995). The picture superiority effect in a cross-modality recognition task. *Memory & Cognition, 23*, 425-441.

Troyer, A. K., & Craik, I. M. F. (2000). The effect of divided attention on memory for items and their context. *Canadian Journal of Experimental Psychology, 54*, 161-170.

Troyer, A. K., Winocur, G., Craik, I. M. F., Moscovitch, M. (1999). Source memory and divided attention: Reciprocal costs to primary and secondary tasks. *Neuropsychology, 13*, 467-474.

VITA

Noelle Brown is from Peachtree City, Georgia, and graduated from McIntosh High School in 1998. She received a four-year athletic scholarship to attend the University of Central Florida where she received her Bachelor of Science degree in psychology in 2003. In 2005 she was accepted to Louisiana State University's doctoral program in the field of cognitive psychology, and awarded a Huel Perkins Fellowship. She currently works with Dr. Jason Hicks performing research on human memory. She has presented posters at conferences, including the annual meetings of the American Psychological Society and the Psychonomic Society.