Presentation format and its effect on working memory

PAULA GOOLKASIAN and PAUL W. FOOS
University of North Carolina, Charlotte, North Carolina

In three experiments, we examined the separate cognitive demands of processing and storage in working memory and looked at how effective the coordination was when items for storage varied in format/modality. A sentence verification task involving arithmetic facts was combined with a span task involving two to six items presented in picture, printed word, or spoken word format. The first two experiments were the same, except for the added requirement of articulation of the math sentence in Experiment 2. Experiment 3 varied the length of the span item and compared recall with recognition performance. The results showed that both spoken words and pictures produced superior recall and recognition, as compared with printed words, and are consistent with Baddeley and Logie's (1999) and Mayer's (2001) models of working memory. Also, the differences in processing performance across spans varied with the difficulty of the task but showed the strongest support for the resource allocation model (Foos, 1995).

There has been much empirical interest in the processing of information from stimuli presented in varying formats. However, the early work concentrated on either picture-word differences (Goolkasian & Park, 1980; Kroll & Corrigan, 1981; Pellegrino, Rosinski, Chiesi, & Siegel, 1977; Smith & Magee, 1980) or auditory-visual differences. More recently, this work has been broadened in order to consider the effects of multimedia presentation techniques (Mayer, 2001) that included video and sound, together with illustrations and textual material. The present experiments extended the latter work by examining working memory for material presented in different modalities (auditory vs. visual) and different formats (pictures vs. printed words) in the same modality (visual).

Picture-word studies (Smith & Magee, 1980) have shown a naming advantage for word processing but a picture advantage when problem-solving tasks or text-processing tasks were used. When Potter (1976) studied short-term memory for pictures, she found that pictures were detected better than printed words. Short-term memory for visual patterns has been found to resemble verbal memory in its susceptibility to similarity effects and in its capacity limitation to around three or four objects (Luck & Vogel, 1997; Wheeler & Treisman, 2002). There are well-recognized differences in encoding stimuli across these two formats (Theios & Amrhein, 1989), and it is generally acknowledged that comprehension of instructional materials is enhanced when illustrations appear together with textual material (Glenberg & Langston, 1992; Larkin & Simon, 1987).

In terms of input modality, Penney (1989) suggests a model of verbal information in which auditory and visual information (pictures or words) are processed in separate streams. Several studies have shown that when the mode of presentation has been auditory rather than visual, superior recall for auditory items occurs, and researchers have referred to this advantage as the *modality effect* (e.g., Greene, 1985; Greene, Elliott, & Smith, 1988). This superior performance consists of at least two components: a sustained sensory memory for auditory input and another that is brief and can be eliminated by presentation of a similar auditory item (the suffix effect; Kellogg, 2001; Nicholls & Jones, 2002).

The long-term modality effect resides in the processing of auditory input that takes place in working memory (Gardiner, Gardiner, & Gregg, 1983). Auditory input of words results in superior performance when compared with visual input of the same words. This long-term effect is not eliminated by the presentation of a suffix item and is found for both serial and free recall (Greene & Crowder, 1986). In the present work, we investigated the advantages of auditory input and compared it with two visual formats, pictures and printed words. Moreover, an attempt was made to separate the effects that were due to sensory plus working memory (Experiment 1) from those that were due only to working memory (Experiment 2 and Experiment 3). It was expected that in all conditions, auditory input would result in significantly better performance than would presentation of the same printed words and, when the sensory memory component was eliminated, would result in the same high performance as that expected with picture presentation.

Recently, Mayer and his colleagues (Mayer, 2001; Mayer & Sims, 1994) developed a cognitive theory of multimedia learning that explained why learning is often facil-

We thank Colleen Farrell, Matthew Hurley, and David Thao for their assistance with data collection and data analysis. Correspondence should be addressed to P. Goolkasian, Department of Psychology, University of North Carolina, Charlotte, NC 28223 (e-mail: pagoolka@email.uncc.edu).

itated by explanations that incorporate at least two formats. Mayer argued that there are two parallel channels, visual/pictorial and auditory/verbal, for processing information through working memory. Although the paths for processing pictures and spoken words are direct and isolated within a channel, processing of printed words was hypothesized to involve both channels in a more complex manner. According to Mayer, format/modality effects should be evident in both the storage and the processing components of the working memory task.

The present research extended prior efforts by examining both format and modality effects in working memory (Baddeley & Hitch, 1974; Baddeley & Logie, 1999). Working memory was thought to be a multicomponent system in which a central executive controls and regulates working memory by coordinating a visuospatial sketchpad and a phonological loop. It was expected that format and modality effects would be evident primarily in the operation of these two specialized stores, with picture formats represented in the visuospatial store and auditory or linguistic information in the other.

In our experiments, we also examined the separate cognitive demands of processing and storage in working memory. The dual task was modeled on a complex operations span task used by La Pointe and Engle (1990). Our processing component was a sentence verification task involving arithmetic facts. Performance was recorded by measuring reaction time (RT) and accuracy to each arithmetic sentence. The storage component was a span task involving two to six items presented for immediate recall. Span items were presented in picture, printed word, or spoken word conditions. Although these conditions represented variations in both format and modality, they were considered in these experiments as levels of a presentation format variable. We were interested in the degree to which presentation format effects would influence performance in the dual task. Previous work (Goolkasian, 1996) with a sentence verification task had shown a picture advantage, but those format differences were attributed primarily to encoding. Given the separate verbal and pictorial channels of processing through working memory and the different findings when picture format and auditory modality were compared with printed words, presentation format effects are predicted.

Finally, in our experiments, we examined predictions about the allocation of working memory resources. The central executive is responsible for the allocation of resources to processing and/or storage tasks and to the coordination of the two other systems. According to one version of the resource-sharing model (Daneman & Carpenter, 1980; Kahneman, 1973), as memory demand increases, there should be less capacity available for processing or storing information through working memory. As demand increases, performance on all tasks should decline. Another version of the resource-sharing model proposes that resource allocation by the central executive ordinarily follows a set of guidelines built into the system. One of these guidelines calls for resource allocation to processing tasks over storage tasks (Foos, 1995). In this case, no decline in

processing accuracy would occur with increased demands for storage, but fewer stored items would be recalled.

We measured the storage versus processing tradeoff by looking at the span effect and what would happen to recall, RTs, and verification accuracy when participants kept in mind from two to six items presented in different presentation format conditions. A large span effect on all of these measures would support the first version of the resource-sharing model, whereas an absence of a span effect on verification accuracy would support the allocation guideline hypothesis.

In sum, the present experiments tested the hypothesis that concrete nouns presented as pictures or as spoken words would result in better memory for those nouns, as compared with printed words. Such a finding would also support Mayer's (2001) predictions regarding the complexity of processing printed words. In addition, presentation format influences on both processing and storage were examined. One version of the resource-sharing model predicts that as span increases, performance on all measures will decline, whereas the allocation guideline hypothesis predicts that processing accuracy will not decline as span is increased.

EXPERIMENT 1

Method

Participants. The participants were 100 male and female students from the University of North Carolina at Charlotte. They were volunteers who participated in the experiment to obtain extra credit points toward their psychology class grade. Twenty-four students participated in Experiment 1, 23 in Experiment 2, and 53 in Experiment 3.

Stimulus materials. A complex operations span task was used. The processing task required the participants to verify the accuracy of a series of math facts presented in a sentence form. Following the participant's keypress response, a concept appeared as a printed word, a spoken word, or a picture, and the participants were required to recall from two to six of these when cued.

The sentences for the processing task were questions adapted from the math task used by La Pointe and Engle (1990). Each sentence contained two components: the multiplication or division of two integers and the addition or subtraction of an integer from the result. The numbers used in both components were integers between 1 and 10. Correct and incorrect examples of the sentences used in the processing task were, respectively, "Is $(8 \times 2) + 4 = 20$?" and "Is $(2/1) + 1 = 2$?" Numbers, rather than words, were used for the sentence verification task, because Campbell's (1994) work had shown a decline in performance when words were used for a math task. Also, to reduce math anxiety, only one integer above 5 was used for the first component of the sentence.

The 60 items used in the span task are presented in the Appendix. Each of the items appeared as a printed word, as a spoken word, and as a picture. The concepts used as the stimuli for the span task were concrete nouns that were chosen from a pool of several hundred clip art photos on the basis of ratings by an independent group of 40 students drawn from the same population as the participants. An item was chosen for the study if at least 90% of the raters agreed on a name for the picture. Each picture was imported into Adobe Photoshop, and its size was adjusted to approximately 4×4 cm for use in the study. When the items appeared as printed words, they were uppercase characters printed with a Geneva font in a character size of 24 cpi. The spoken words were sound files created by a female voice as a Macintosh system sound file.

The stimuli were displayed on an Apple Color High Resolution RGB 15-in. monitor. Stimulus presentation and data collection were

controlled by SuperLab running on a Power Macintosh 7000 series computer.

Procedure. The participants were run individually in sessions of around 30 min. They were introduced to the dual task and given five practice trials to familiarize themselves with the task. They were asked to verify the accuracy of a series of arithmetic facts while simultaneously remembering a varying number of items presented as pictures, printed words, or spoken words. The math sentence appeared first and remained on the screen until the participant made a *yes/no* response (with either the F or the J on the keyboard), using his or her right or left index finger. An item from the span task followed the participant's response. The item was a picture, a printed word, or a spoken word presented against a blank screen. After 1 sec, a second math sentence appeared again, followed by the second item from the span task. This continued until a string of three question marks cued the recall of the span items. The participants called out as many of the span items as they could remember, and their spoken responses were recorded by a research assistant. When ready, the participant used a keypress to begin the next trial. The span task varied between two to six items.

In all, there were 45 recall trials for each of the participants. Five span conditions (two to six items) were combined with the three formats, and each of the 15 conditions was replicated three times. Within each replication, the order of the format $\times$ span conditions was randomized. Also, within each of the recall trials, the item order was randomized each time the experiment was run.

There were 180 math sentences developed for the sentence verification task. Of the 60 used in each of the three replications, there was an equal number of multiplication and division and of subtraction and addition facts, with half correct and half incorrect examples. The incorrect answer was always within two units of the correct answer. The pairing of the items used in the span task with the math sentences was random, and within each of the 15 recall conditions, there was an equal number of correct and incorrect math sentences.

Our complex operation span task differed from La Pointe and Engle (1990) in several important respects. The computer recorded the RTs and accuracy scores from the sentence verification task so that we could look at both span and format effects on the processing component of the working memory task. The participants were encouraged to recall as many of the items as they could, irrespective of when the item was presented, and recall was spoken rather than written. Also, recall was scored as the proportion of presented items correctly remembered. Lastly, in the first experiment, the participants were not required to read out loud either the math sentence or the items from the span task. However, in Experiments 2 and 3, the participants articulated the math sentence prior to their verification response.

Results

Means were computed from the correct RTs to the sentence verification task and from the proportion of items correctly recalled from each participant across the three replications within each of the experimental conditions. Also recorded were the proportion of incorrect RT responses. A 5×3 repeated measures analysis of variance (ANOVA) was used on the RT and recall data to test for span and format effects. The F tests that are reported include the Greenhouse–Geisser correction to protect against possible violation of the homogeneity assumption.

Format and span effects were obtained on both processing speed and storage components of the dual task. Figure 1 presents the interactions. From the top panel, it is apparent that the ability to remember items varied by format [$F(2,46) = 43.57, p < .01; \eta^2 = .65$]. Scheffé tests (at

the $p < .05$ level of significance) showed that the three formats differed significantly from each other. The means for the three format conditions were .92, .85, and .78, respectively. Span effects were also quite strong [$F(4,92) = 45.40, p < .01; \eta^2 = .66$]. Recall fell in a linear relationship with the number of items held in memory ($p < .01$ for the linear trend). Although there was a significant format $\times$ span interaction [$F(8,184) = 2.40, p < .02; \eta^2 = .09$], the effect size was small, and it most probably resulted from a ceiling effect with the short spans. The higher recall of words when four items were held in storage was puzzling and did not replicate in Experiment 2 when the same condition was tested.

RTs to the sentence verification task (middle panel of Figure 1) were also found to be influenced by format [$F(2,46) = 5.82, p < .01; \eta^2 = .20$] and span [$F(4,92) = 11.39, p < .01; \eta^2 = .33$] and their interaction [$F(8,184) = 4.56, p < .01, \eta^2 = .16$]. Verification of the math facts was quicker when spoken words, rather than printed words or pictures, were held in memory and when short, as compared with longer, spans were tested. However, the format effect varied with span, and at the shortest span, printed words were associated with the quickest RTs.

Errors rates from the sentence verification task showed effects that were somewhat consistent with the RT data. Error rates were essentially the same across all the conditions, except when two or three spoken words were held in mind. In these two conditions, error rates were unusually low, as compared with the other format and span conditions. As can be seen in the bottom panel of Figure 1, this result produced a significant format $\times$ span interaction [$F(8,184) = 2.64, p < .01$] and main effects of format [$F(2,46) = 3.82, p < .05$] and span [$F(4,92) = 4.16, p < .01$].

Discussion

Both storage and processing time components of the working memory task showed substantial effects from variation in the presentation format and the number of items held in memory. There was a strong advantage to storing items presented as spoken words, and pictures were remembered better than printed words.

The present results do not provide clear support for either model of resource sharing. One model (Daneman & Carpenter, 1980) predicted an effect of span on processing accuracy such that accuracy would decline as span increased. An effect of span on accuracy was obtained. The effect, however, was nothing like that predicted, since error rates deviated only in two conditions and error rates in those conditions were unusually low and seemingly were due to format, in addition to low span. The allocation guideline hypothesis (Foos, 1995) might claim partial support here, since error rates were generally the same regardless of span.

The fact that for some selected span conditions, there was an almost perfect recall of spoken words, accompanied by faster processing time in the verification task and fewer errors in comparison with the other formats, suggests that the advantage for spoken words may reflect differences in

Figure 1. Mean scores for the format $\times$ span interaction from Experiment 1. The error bars represent standard errors. The top panel presents the proportion of items recalled in the storage task, and the middle panel presents the reaction times from the sentence verification task. The bottom panel presents the proportion of errors made in the sentence verification task.

sensory memory, in addition to differences in working memory for auditory as compared with visually presented material. These findings may be similar to the modality effect commonly obtained in short-term memory research, where spoken words have produced a recall advantage for the last

several items in a list, owing to the longer availability of auditory information, in comparison with visual (Penney, 1989). To separate out the effects from sensory memory and to better understand the format difference in working memory, Experiments 2 and 3 were conducted.

EXPERIMENT 2

The second experiment was a direct replication of the first, except that the participants were asked to read out loud the math sentence that appeared on every trial. The participant's articulation of the math facts was expected to interfere with sensory memory for the spoken material and to wipe out any sound advantage resulting from echoic memory. Any modality effects that resulted from working memory should be unaffected by this articulation.

Method

The stimulus materials and the procedure were the same as those in Experiment 1, except that the participants were asked to read out loud the math sentence as soon as it appeared and then to respond with a key press to indicate the statement's accuracy. As in Experiment 1, an item from the span task followed the participant's response. The item was a picture, a printed word, or a spoken word presented against a blank screen. After 1 sec, a second math sentence appeared, again followed by a second item from the span task. This continued until a string of three question marks cued the recall of the span items. Items ranged in number from two to six. The participants were another group of students drawn from the same population as those in Experiment 1. In all other respects, this experiment was the same as the first.

Results

Figure 2 presents the means from the data collected in Experiment 2. Recall of items presented in the span task varied by format [$F(2,44) = 28.48, p < .01; \eta^2 = .56$], but the format effect was limited to a lower recall of items presented as printed words. The advantage that spoken words had over pictures in Experiment 1 was not replicated. As in Experiment 1, recall declined in a linear fashion with the number of items held in memory [$F(4,88) = 126.20, p < .01; \eta^2 = .85$]; however, there was no interaction of span and format [$F(8,176) = 1.31, p = .24; \eta^2 = .05$].

The analysis of the RTs showed only an effect of span [$F(4,88) = 5.36, p < .01; \eta^2 = .20$]. Verifying the accuracy of math facts took longer when more (five or six), rather than fewer (two, three, or four) items were held in memory. RTs were not affected by format [$F(2,44) = 1.93, p = .16; \eta^2 = .08$] or by the interaction of format and span [$F(8,176) = 1.02, p = .42; \eta^2 = .04$].

The errors made in the sentence verification task were also found to vary as a function of span [$F(4,88) = 2.61, p = .04$], and there was a span $\times$ format interaction [$F(8,176) = 2.71, p < .01$]. As can be seen in the lower panel of Figure 2, these effects were due to more errors being made when five or six pictures were held in memory.

Discussion

The results of this second experiment support the hypothesis that the prior advantage of spoken words over pictures was due to sensory memory. Articulation eliminated that advantage, although strong presentation format effects were still obtained.

The fact that both spoken words and pictures produced superior recall, as compared with printed words, is consistent with past findings of an advantage for picture pro-

cessing and an advantage for spoken over printed words. This experiment extended those results by making a direct comparison between picture and spoken versions of the items and supported Mayer's (2001) prediction of lower performance for printed words because printed words involve two channels and more complex processing, whereas pictures and spoken words involve but a single channel.

The picture over word superiority effect also supports the more traditional models of working memory that specify separate components for the processing of visual and auditory information (Baddeley & Logie, 1999). Because pictures do not rely as extensively on the phonological loop for storage as printed words, they were at somewhat of an advantage when it came to recall of the span items, because visual working memory may not have been as disturbed by the articulation task as verbal memory.

The spoken word advantage over pictures was eliminated with articulation, but interestingly, there was still an advantage relative to printed words. It is possible that the articulation task prevented the participants from coding the printed words into verbal working memory, but since the spoken words were already in a phonological format (speech), their recall was not as impaired as that of the printed words by the articulation task. Wilson (2001) has suggested the possibility of such sensory coding effects in verbal working memory.

The format effects in Experiment 2 may also have resulted from the added requirement to articulate the math facts. However, other than eliminating the sound advantage to processing spoken words, there was no evidence that the differences between experiments interacted with format or span. The need to articulate the math facts prior to producing a verification response added to the load on the central executive and made the dual task more difficult.

Finally, the present results, like those from the first experiment, do not clearly support either model of resource sharing. The allocation guideline hypothesis predicts no effect of span on processing accuracy, and yet such effects were obtained in both experiments. These obtained effects are, however, nothing like those predicted by the other model and seem, instead, to be due to unusually low error rates with two or three spoken words in Experiment 1 and unusually high error rates with five or six pictures in Experiment 2. Error rates for all the other conditions seem to be about the same.

Before any conclusions could be made, however, regarding the varying explanations for the format effects, there were at least two lingering issues that needed to be clarified. The use of single- and multiple-syllable items may have made the printed words more difficult to recall, in comparison with the other format conditions. Short words are generally recalled at a higher rate than long words (La Pointe & Engle, 1990; Wilson, 2001). The effect is also related to the length of time it takes to articulate the printed words, rather than to the number of syllables. Since our procedure did not include articulation of the items presented in the storage task, it was unknown what influence word length would have on recall. It was also of

Figure 2. Mean scores for the format $\times$ span interaction from Experiment 2. The error bars represent standard errors. The top panel presents the proportion of items recalled in the storage task, and the middle panel presents the reaction times from the sentence verification task. The bottom panel presents the proportion of errors made in the sentence verification task.

interest to test whether the format effects in working memory would generalize when memory tests other than recall were used. Experiment 3 was conducted to examine both of these issues.

EXPERIMENT 3

In Experiment 3, we looked at the effect of item length and type of memory task on processing and storage in

working memory. The stimulus items were divided into single and multiple syllables, to create an item length variable. As in the two previous experiments, all the items were presented in picture, printed word, and spoken word formats. The processing component of the dual task was the same as that in Experiment 2. The participants articulated the math facts and then verified their accuracy. The storage component was reduced to only two span conditions—short (three items) versus long (six items)—and the two memory tasks (recall and recognition) were compared between groups. We wanted to make sure that the lower performance when printed words were used did not result from a word length effect and would also occur when another memory test was used.

Method

Participants. The 53 participants were randomly assigned to one of two memory test conditions. The data from one of the participants were eliminated because of lower than chance performance on the sentence verification task.

Stimulus materials. The span task involved 54 stimulus items that were divided into single and multiple syllables on the basis of length. The Appendix lists the items in each of the two categories. Although a number of stimuli were the same as the ones in use previously, some new items were added from the original pool. Each item appeared as a picture, a printed word, and a spoken word. In all cases, ratings indicated that there was at least 90% agreement on a name for the picture. The 54 math sentences for the processing task were randomly drawn from the sentences used in the previous experiments.

The stimuli were displayed on an Apple 15-in. flat screen monitor. Stimulus presentation and data collection were controlled by SuperLab running on a PowerPC G4. Spoken words were presented with stereo speakers. In all other respects, the stimuli were the same as those in the previous experiment.

Procedure. The participants were run individually in 30-min sessions. They were randomly divided into two memory test groups—recall and recognition. Both groups had 36 trials in which they had to either recall or recognize the items presented on the span task. There were 2 span conditions (three and six) combined with 2 length conditions (single and multiple) and three formats (picture, spoken word, and printed word). Each of the 12 experimental conditions was replicated three times. Within each of the replications, use of each item was counterbalanced across format. Also, within each replication, the order of the 12 experimental conditions was randomized.

As in the previous experiment, the memory task was cued by a string of three question marks that appeared on the screen. For the recognition test, the participants were provided a sheet with printed words that included the actual items presented in the span task and an equal number of distractor items randomly drawn from the same syllable condition. Distractor and original items were randomly ordered on the sheet. The participants were asked to check the items presented on that trial. A research assistant monitored the task to make sure that the participants saw only the recognition items from a specific trial. When ready for the next trial, the participant made a keypress.

The procedure for the recall task was the same as that in Experiment 2. When cued, the participants called out as many of the span items as they could remember, and their spoken responses were recorded by a research assistant.

There were 162 sentence verification trials. The 54 math sentences were used three times, once in each of the replications of the 12 experimental conditions. There was an equal number of multiplication and division and of subtraction and addition facts, with half correct and half incorrect examples. The pairing of the items used in the span task with the math sentences was random, and within each

of the 12 experimental conditions, there was an equal number of correct and incorrect math sentences. In all other respects, the procedure was the same as that in Experiment 2.

Results

Data from each of the memory tests were treated separately with a $3 \times 2 \times 2$ ANOVA to test for the within-group effects of format, syllable length, and span. Figure 3 presents the recall and recognition data for each of the format $\times$ span conditions. Performance in the recognition memory condition was consistent with previous results. Printed words were at somewhat of a disadvantage when compared with recognition memory for pictures and spoken words. Analysis of the recognition memory scores showed an effect of format [$F(2,50) = 69.91, p < .01; \eta^2 = .74$] and span [$F(1,25) = 308.96, p < .01; \eta^2 = .92$], but no interaction between these variables [$F(2,50) = 2.51, p < .09$].

Analysis of the recall data also showed effects of format [$F(2,50) = 41.43, p < .01; \eta^2 = .62$] and span [$F(1,25) = 480.12, p < .01; \eta^2 = .95$]. However, for this analysis, the interaction was also significant [$F(2,50) = 6.78, p < .01; \eta^2 = .21$]. Although printed words were at a relative disadvantage in comparison with other formats, pictures were recalled at a lower rate than were spoken words when short spans were tested.

The effect that syllable length had on the storage of information was not consistent with the interpretation that the poor recall of printed words occurred because the items included multisyllabic words. The ANOVA on the recall data did not show a main effect of syllable length [$F(1,25) = 1.70, p < .20$; means for single and multisyllabic items were .69 and .68, respectively], nor was there an interaction of syllable length with format [$F(2,50) = 2.42, p = .10$]. Rather, syllable length was found to interact with span [$F(1,50) = 4.93, p < .04; \eta^2 = .16$], and there was a triple interaction with span and format [$F(2,50) = 8.37, p < .01; \eta^2 = .25$]. These interactions occurred because, at short spans, single-syllable items were recalled slightly better than multisyllabic items and the effect was more noticeable for spoken words and pictures than for printed words.

A main effect of syllable length was found with the recognition data [$F(1,25) = 27.71, p < .01; \eta^2 = .51$]; however, the effect was opposite to that predicted, with the multisyllabic items being recognized at a higher rate than the single-syllable items (means for the multisyllabic and single items were .82 vs .78, respectively). Also, there was a significant syllable $\times$ span interaction [$F(1,25) = 19.47, p < .01; \eta^2 = .44$], which indicated that there was a larger syllable effect for the long spans than for the short spans. None of the other interactions with syllable length reached significance.

With the exception of a small recall advantage to spoken words, relative to pictures at short spans, format and span effects were consistent with the results of Experiment 2.

Consistent with the findings from Experiment 2, RTs from the recall task were faster when the participants were

Figure 3. Mean scores for the format $\times$ span $\times$ memory task interaction from Experiment 3. The error bars represent standard errors. The top panel presents the proportion of items recalled in the storage task, and the middle panel presents the reaction times from the sentence verification task. The bottom panel presents the proportion of errors made in the sentence verification task.

asked to store three, rather than six, items [$F(1,25) = 6.50$, $p < .02$; $\eta^2 = .20$]. There was also a span $\times$ format interaction [$F(2,50) = 3.87$, $p < .03$; $\eta^2 = .13$]. From the middle panel of Figure 3, it appears that this interaction occurred because the RTs did not increase with span for the spoken words, as they had in the other two format conditions. Item length was also found to influence RTs [$F(1,25) = 4.53$, $p < .04$; $\eta^2 = .15$]. Verification responses were slower when single-syllable items were stored, rather than multiple-syllable items. Means were 5,941 and 5,804 msec, respectively. Item length was also found to interact with format and span [$F(2,50) = 4.43$, $p < .02$; $\eta^2 = .15$]. However, the reason for this effect was not obvious.

Verification responses with the recognition data showed the same triple effect of item length, format, and span as

the previous analysis [$F(2,50) = 7.66$, $p < .01$; $\eta^2 = .23$]. There was also a syllable $\times$ format interaction [$F(2,50) = 3.19$, $p < .05$; $\eta^2 = .11$]. None of the other effects in the analysis of the RTs reached significance.

Errors rates from the sentence verification task are presented in the bottom panels of Figure 3. The only significant effects were found in the data from the recognition task. There was an increase in the number of errors for short spans, as compared with long spans [$F(1,25) = 6.33$, $p = .02$], an effect of format [$F(2,50) = 3.67$, $p = .03$], and an interaction of span and format [$F(2,50) = 5.41$, $p = .01$]. More errors were made to the sentence verification task when printed words, rather than the other formats, were presented in short spans. None of the other effects in the analysis of the errors reached significance.

Discussion

The present results replicated the first two experiments by showing advantages of pictures and spoken words over printed words. This finding occurred in both recognition and recall memory tests and with single- and multiple-syllable items. These results eliminated word length as a significant factor in explaining the presentation format effect obtained in these experiments.

Overall, the effects of presentation format and span were consistent across the two kinds of memory tasks, with a few exceptions. At the shorter span in the recall test, the advantage of spoken words over pictures was again present. Also at the shorter span, multiple-syllable items were better recognized than single-syllable items and showed slightly smaller presentation format effects.

GENERAL DISCUSSION

The present experiments clearly showed presentation format effects in working memory. The advantage of pictures over printed words was found in a dual-task, working memory paradigm, as in problem-solving and text-processing tasks (Smith & Magee, 1980). The advantage of auditory over visual presentation, the modality effect, was also found in this same paradigm. The advantages of pictures and spoken words over printed words were the same, except when auditory sensory memory contributed to the modality effect (Experiment 1) and at very short spans, where spoken words were recalled better than pictures.

These results support both Baddeley and Logie's (1999) and Mayer's (2001) models of working memory. Mayer's model predicts that some combinations of formats should result in superior performance in working memory tasks. Since both spoken words and pictures resulted in superior performance, perhaps their combination will result in even higher performance. Our future work will examine these possibilities.

Finally, some support was found for the hypothesis that the allocation of resources by the central executive to storage and processing tasks follows a set of built-in guidelines (Foos, 1995). Generally, accuracy on the processing task was unaffected by changes in information storage, except at very short spans with words and very long spans with pictures. When the task requirements were made more difficult by the need to articulate, in addition to process and store information, verification accuracy was maintained at a stable rate, whereas recall performance declined. Other, related work has shown that under some circumstances, the central executive can override guidelines and allocate resources in other ways (Foos, 1995). Perhaps, at a very short span with spoken words (Experiment 1), the storage task was so easy that even more resources could be allocated to processing. Although both models seem able to account for the findings from Experiment 1, Experiments 2 and 3 support the allocation guideline hypothesis.

Since concrete nouns representing common objects were used exclusively as the stimuli for the storage task, it would be inappropriate to generalize the presentation format effects to other kinds of material. However, there is no doubt

that these findings show that the manner in which information is presented has an influence on working memory. Material that is presented as printed words are the least effective of the presentation formats that were studied, and pictures and spoken words appear to provide equivalent benefits.

REFERENCES

- BADDELEY, A. D., & HITCH, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.
- BADDELEY, A. D., & LOGIE, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 28-61). Cambridge: Cambridge University Press.
- CAMPBELL, J. I. D. (1994). Architectures for numerical cognition. *Cognition*, **53**, 1-44.
- DANEMAN, M., & CARPENTER, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, **19**, 450-466.
- FOOS, P. W. (1995). Working memory resource allocation by young, middle-aged, and old adults. *Experimental Aging Research*, **21**, 239-250.
- GARDINER, J. M., GARDINER, M. M., & GREGG, V. H. (1983). The auditory recency advantage in longer term free recall is not enhanced by recalling prerecency items first. *Memory & Cognition*, **11**, 616-620.
- GLENBERG, A. M., & LANGSTON, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory & Language*, **31**, 129-151.
- GOOLKASIAN, P. (1996). Picture-word differences in a sentence verification task. *Memory & Cognition*, **24**, 584-594.
- GOOLKASIAN, P., & PARK, D. C. (1980). Processing of visually presented clock times. *Journal of Experimental Psychology: Human Perception & Performance*, **6**, 707-717.
- GREENE, R. L. (1985). Constraints on the long-term modality effect. *Journal of Memory & Language*, **24**, 526-541.
- GREENE, R. L., & CROWDER, R. G. (1986). Recency effects in delayed recall of mouthed stimuli. *Memory & Cognition*, **14**, 355-360.
- GREENE, R. L., ELLIOTT, C. L., & SMITH, M. D. (1988). When do interleaved suffixes improve recall? *Journal of Memory & Language*, **27**, 560-571.
- KAHNEMAN, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- KELLOGG, R. T. (2001). Presentation modality and mode in verbal false memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 913-919.
- KROLL, J. F., & CORRIGAN, A. (1981). Strategies in sentence-picture verification: The effect of an unexpected picture. *Journal of Verbal Learning & Verbal Behavior*, **20**, 515-531.
- LA POINTE, L. B., & ENGLE, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 1118-1133.
- LARKIN, J. H., & SIMON, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, **11**, 65-99.
- LUCK, S. J., & VOGEL, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.
- MAYER, R. E. (2001). *Multimedia learning*. Cambridge: Cambridge University Press.
- MAYER, R. E., & SIMS, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual coding theory of multimedia learning. *Journal of Educational Psychology*, **86**, 389-401.
- NICHOLLS, A. P., & JONES, D. M. (2002). Capturing the suffix: Cognitive streaming in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 12-28.
- PELLEGRINO, J. W., ROSINSKI, R. R., CHIESI, H. L., & SIEGEL, A. (1977). Picture-word differences in decision latency: An analysis of single and dual memory models. *Memory & Cognition*, **5**, 383-396.
- PENNEY, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, **17**, 398-422.
- POTTER, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 509-522.
- SMITH, M. C., & MAGEE, L. E. (1980). Tracing the time course of

- picture-word processing. *Journal of Experimental Psychology: General*, **109**, 373-392.
- THEIOS, J., & AMRHEIN, P. C. (1989). Theoretical analysis of the cognitive processing of lexical and pictorial stimuli: Reading, naming, and visual and conceptual comparisons. *Psychological Review*, **96**, 5-24.
- WHEELER, M. E., & TREISMAN, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, **131**, 48-64.
- WILSON, M. (2001). The case for sensorimotor coding in working memory. *Psychonomic Bulletin & Review*, **8**, 44-57.

APPENDIX

Items Used in Experiments 1 and 2

ball	frog	pretzel
bat	globe	rabbit
bike	guitar	radio
blender	horse	rocket
books	house	rose
boot	key	scales
box	lighthouse	shark
bus	lion	scissors
car	monkey	snake
cat	mushrooms	snowman
chair	owl	stapler
cheese	panda	tank
cherries	pen	television
cow	pencil	tie
crab	penguin	toilet
dog	phone	tractor
eagle	pie	turtle
egg	pig	umbrella
fish	plane	watch
football	plant	whale

Items Used in Experiment 3

Single Syllable	Multiple Syllables
ball	blender
bat	cherries
bike	umbrella
books	football
boot	guitar
box	lighthouse
bus	monkey
car	mushrooms
cat	panda
chair	pencil
cheese	penguin
cow	pretzel
crab	rabbit
dog	radio
whale	rocket
egg	scissors
fish	snowman
fog	stapler
globe	toilet
horse	tractor
house	turtle
key	calculator
plant	elephant
owl	balloons
pen	handcuffs
pie	microscope
pig	sunglasses