Modality and Recall

Sarah M. LeStourgeon

Hanover College
The main focus of this study looks at the way humans process information and store it into memory when a stimulus is presented either visually, aurally, or a combination of the two. Based on the way information is processed, the researchers can then conclude which modality would therefore be most efficient for immediate recall. For example, if people process information aurally, then auditory stimuli should have a greater chance of recall compared to visual stimuli. There are two main theories that apply to this study. One theory includes the use of working memory and the other includes a separate memory store called short term memory (STM).

Working memory is related to information processing because it is specifically involved in controlling what information is processed and what information is thrown out while a person is performing some kind of cognitive task. The cognitive psychology textbook by Solso, MacLin, & MacLin (2005) further discusses working memory and exactly how it processes information. One theory which is examined in the text was proposed by Baddley (1986). He proposed two different strategies for processing which occurred in a verbal and visual store. The verbal store in working memory holds information based on what Baddeley called the phonological loop. This phonological loop suggests that people have to rehearse information aurally in order to hold that information until it is further processed or disposed. Similarly, Baddeley also proposed a visual store he called the visuospatial sketchpad. Like the phonological loop, information must be rehearsed in order to be held in memory. Baddeley has demonstrated in his research that these two memory stores work independently, but further research has shown even more concrete evidence that different parts of the brain are activated during a visual versus an auditory task. Using brain-imaging techniques Cabeza
& Nyberg (1997) demonstrated that certain parts of the brain were activated during a verbal memory task while Haxby, Ungerleider, Horwitz, Rapoport, and Grady (1995) showed that completely separate parts of the cortex were activated during a visual task.

Although Baddeley’s theory of working memory has previously been viewed independently, for the purpose of the present study his theory will be compared to that of Paivio (1986) who proposed the dual-coding theory of STM. Paivio suggested that there are two systems which can be activated simultaneously, verbal and visual. Words and language are coded into the verbal memory store while images are coded into the visual store. His theory specifically states that when both stores are activated at the same time, memories are enhanced and are able to be recalled more often and more accurately. In relation to Baddeley’s phonological loop and visuospatial sketchpad, Paivio’s theory could mean that when information is being processed within these two domains simultaneously the memory has more data and therefore, creates an availability of that memory for future recall whether immediate or delayed.

Further supporting Paivio’s dual-processing model is the study by Moreno & Mayer (2002). Based on the previous research investigating a dual-processing model, Moreno and Mayer investigated the enhancement of memory and recall when the two streams of visual and auditory information occurred simultaneously. To test this, they had participants listen to a scientific narration describing the process of lightning formation while viewing on-screen text at the same time. Results showed that participants who were presented with the information visually and aurally retained significantly more than those who did not. From these results it is evident that Moreno and Mayer are suggesting that memory recall is more abundant and accurate when two modalities are incorporated into a task compared to when
only one modality is incorporated. Therefore, this led the researchers to the first part of the hypothesis: When two streams of visual and auditory processing are activated simultaneously, participants will recall a greater number of words compared to when the words are presented only visually or only aurally.

In the same text which investigated working memory and information processing (Solso, Maclin, & Maclin, 2005), it also introduces another theory that investigates the way information is processed in STM. This second theory was developed by Conrad (1963, 1964) and states that STM is dominated by an auditory processing model. This is evident in Conrad’s Memory Span experiment in which participants were presented with chunks of letters that either sounded different or chunks of letters which sounded similar (s and f or c and v). The results indicated that participants recalled more letters when they sounded different than when they sounded the same. Demonstrated in Figure 1, the same study was conducted in a cognitive psychology class and the same results were found. This suggests that our memory processing errors are mainly auditory and therefore, people predominantly process visual information aurally. Based on Conrad’s results, the researchers developed a second part of the hypothesis which predicted that participants will be more able to recall words presented aurally. It would take longer to process visual information because it has to enter an extra step as it is coded from visual to auditory. Therefore, this leads to the second hypothesis that more words in the auditory condition will be immediately recalled compared to those in the visual condition.

Method

Subject
Modality and Recall

There were a total of 34 participants, 84 percent of which were female, which ranged in age from 18-23 years old. Those who participated were all students at a liberal arts college in the midwest and were all capable of reading, hearing, and comprehending spoken and written English. They volunteered after they were informed by word-of-mouth or by flyers posted in the psychology department.

Equipment

Microsoft Powerpoint (2003) was used to create a slideshow on a Sony Vio laptop computer (model PCG-9231) with Windows 98. Words were presented on a projector screen (resolution 1024x780) by using a Sharp Notevision 5 (serial number: XG-NV5XB) projector. Also, in the verbal conditions, auditory stimuli were presented with Boston Speakers (model BA 265).

Stimuli

Looking at Table 1, it displays three lists of words used for the three conditions. Each list contains nine words which were presented either visually, aurally, or both. Within the conditions with a visual stimulus, written words were presented on a blank white screen in size 22 Times New Roman, black font. Each word was presented for two seconds on its own slide in capital letters. Words presented verbally were recorded into Microsoft PowerPoint (2003) by a female voice. They were spoken clearly and were chosen so that they would not be easily confused with similar words. All words chosen were concrete, unrelated, and no two words were presented in the same condition or any other condition.

The modality (visual or verbal) and the combination of modalities (visual and verbal together) in which the words were presented is the independent variable. A mean percent
accuracy of recall was calculated for each condition and serves as the dependent variable in
this study.

Procedure

At the beginning of the study participants signed an informed consent form. Then, to
account for order effects, participants ran condition one (visual stimuli), condition two (verbal
stimuli), and condition three (visual and verbal simultaneously) in different order
arrangements based on which of the three hours the study was completed. Those who came in
the first hour of the study completed conditions 1, 2, and 3. Those who entered the study the
second hour completed conditions 2, 3, and 1. Finally, the students who came to the third
hour of the study completed conditions 3, 1, and then 2. After each condition participants
immediately wrote down as many words as they could recall in any order, with no time limit.
The experiment was run within one day and the researchers then compiled the words recalled
and mean accuracy percentages for each condition into a spreadsheet.

Results

In order to compare the mean percentages for words recalled within the three
conditions, the researchers ran a within subjects one-way Analysis of Variance (ANOVA).
This test indicated whether modality significantly affected the amount of words recalled by
the participants. The result of the ANOVA showed that there was no significant main effect
of modality on participants’ ability to recall a list of words, $F(2,58) = .668$, $p = .517$. 
Therefore, these results do not support the researchers’ hypothesis that presentation of visual
and verbal stimuli simultaneously enhances our ability to immediately recall words. The
results also do not show evidence for the researchers’ second hypothesis that aurally presented
words would be more significantly recalled than those words presented visually.
Discussion

One possible reason why the mean percent accuracies between the visual and verbal conditions were not significantly different is because different strategies of rehearsal were not accounted for. De Beni and Moe (2003), however, did control for rehearsing strategies (visually or aurally) when presented with an oral or written stimulus. They found that written presentation had better recall with the use of verbally based strategies while oral presentation had better recall with the use of imagery based strategies. In comparison, the present study used words which could be easily rehearsed both visually and verbally, therefore it did not account for the underlying differences in how visual and auditory information is processed. If this is the reason why significance was not found, then according to past research a solution to finding more profound differences would be to present participants with abstract words (not easily visualized) as well as easily visualized words. By having two types of stimuli would account for the way in which the list of words was rehearsed and all known confounding factors would be controlled in order to see a definite trend of visual versus auditory processing.

Another explanation for the similarity in results for the visual and verbal conditions could be that the researchers underestimated the strength of the visual code in STM. Originally, with the emphasis of Conrad’s Memory Span experiment, there was a notion of auditory dominance during the processing of information. Perhaps this theory is overestimating the difference in auditory and visual processing ability. However, if there truly is an auditory dominance, the process in which visual information is converted to auditory information could be so automatic so that there is no difference in time length between the integration of written and verbal information into memory.
Interestingly, by looking at Figure 2 it is evident that the pattern of mean percent accuracy follows the hypothesis that auditory information would have higher recall ($M = .64$) compared to visual information ($M = .63$). Also, the figure demonstrates the pattern that auditory and visual stimuli presented simultaneously did, in fact, increase immediate recall ability ($M = .68$). The absence of significant differences could have been caused by the lack of abundant participants which, if there were more students participating in the present experiment, would presumably define the results and show significant differences.
Table 1: Modality and 9-item word list for the three conditions

<table>
<thead>
<tr>
<th>Word order</th>
<th>Visual only (Condition 1)</th>
<th>Verbal only (Condition 2)</th>
<th>Verbal and Visual (Condition 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ROSE</td>
<td>CLOCK</td>
<td>BALL</td>
</tr>
<tr>
<td>2</td>
<td>TRUCK</td>
<td>BAG</td>
<td>RAKE</td>
</tr>
<tr>
<td>3</td>
<td>DANCE</td>
<td>TRAIN</td>
<td>HAIR</td>
</tr>
<tr>
<td>4</td>
<td>LIME</td>
<td>NEST</td>
<td>CAMEL</td>
</tr>
<tr>
<td>5</td>
<td>BRIDGE</td>
<td>KITE</td>
<td>SOCK</td>
</tr>
<tr>
<td>6</td>
<td>CHURCH</td>
<td>HAMMER</td>
<td>SHELF</td>
</tr>
<tr>
<td>7</td>
<td>APPLE</td>
<td>CAN</td>
<td>GLASS</td>
</tr>
<tr>
<td>8</td>
<td>CHAIR</td>
<td>STAR</td>
<td>MOUSE</td>
</tr>
<tr>
<td>9</td>
<td>DUCK</td>
<td>TREE</td>
<td>RING</td>
</tr>
</tbody>
</table>

Figure 1: Class data for Memory Span experiment comparing recall for similar and different sounding letters.
Figure 2: Mean percent accuracy for recall across modalities.

The graph shows the percent accurate recall for visual, verbal, and visual and verbal modalities. The visual modality has a percent accurate recall of approximately 0.62, the verbal modality has a percent accurate recall of approximately 0.60, and the visual and verbal modality has a percent accurate recall of approximately 0.68.
