Pictures, Words, and Sounds: From Which Format Are We Best Able to Reason?

PAULA GOOLKASIAN
Department of Psychology
University of North Carolina, Charlotte

ABSTRACT. The effect of presentation format on reasoning was studied with a sentence verification task. Background information was presented in single-format and combined conditions that included pictured, printed, or spoken versions of the stimulus items. In Experiment 1, a test sentence appeared together with the background at varied stimulus onset asynchronies, to study how format influences the acquisition of the stimulus information. In Experiments 2 and 3, however, the test sentence followed the presentation of the background, to test the effect of format on memory. Reaction time responses to the test sentences showed a consistent picture advantage. However, when participants responded to materials stored in memory, both pictured and spoken formats provided quicker responses in comparison to printed words, and the format difference was smaller than when materials were readily available on the screen. Multimedia presentations, when compared with single-format conditions, did not provide additional benefits.

Key words: format effects, multimedia, picture-word differences

PROBLEM SOLVING with stimulus materials that include several formats generally shows a picture advantage. Goolkasian (1996) compared picture and word formats and found that participants could make an inference to verify the accuracy of a test sentence more quickly when background material appeared in picture format. Similarly, Bauer and Johnson-Laird (1993) compared verbal and diagrammatic presentation of problems in deductive reasoning and found considerable improvement in the speed and number of valid conclusions with the diagrams.

Although format effects have been the subject of much research (Goolkasian & Park, 1980; Kroll & Corrigan, 1981; Pellegrino, Rosinski, Chiesi, & Siegel, 1981),

The author thanks Tim Brown, Mandy Tarantino, and Kelly Barrett for their assistance with data collection and data analysis. A portion of these findings were reported at the annual meeting of the Psychonomic Society in Philadelphia, November 1997.

Address correspondence to Paula Goolkasan, Department of Psychology, University of North Carolina, Charlotte, 9201 University City Blvd., Charlotte, NC 28223; send e-mail to pagoolka@email.uncc.edu.
1977; Potter & Paulconer, 1975; Smith & Magee, 1980), the results have varied with the kind of task, and there is a lingering debate regarding the nature of the representations that are developed from picture and word stimuli. Some researchers claim that format effects result from differences in the way that pictures and words are stored (Glenberg & Langston, 1992; Paivio, 1971, 1975, 1978), whereas others emphasize perceptual differences among the stimulus formats. For example, Larkin and Simon (1987) indicated that text and diagrams containing the same information are not necessarily equivalent in terms of the processing required to extract the information because some features may be directly represented in one that may be inferred in the other. They identified picture-word differences in the efficiency of the search for information and differences in the explicitness of the information.

Most recently, this work has broadened to consider effects of multimedia presentation techniques. Sweller and his associates (Sweller, Chandler, Tierney, & Cooper, 1990; Tindall-Ford, Chandler, & Sweller, 1997) outlined the conditions under which problem solving with instructional materials may benefit from a dual rather than single mode of presentation. Attending to multiple sources of information requires participants to mentally integrate disparate information prior to problem solving and can produce a split attention effect that interferes with problem solving. However, when the material presented in varied formats is physically integrated, problem solving is facilitated because the load on working memory is reduced. Interference from split attention effects can also be reduced if information is presented in more than one sense modality (Mousavi, Low, & Sweller, 1995). Following Baddeley's (1992) description of working memory that includes two separate and independent processors—a visual-spatial sketch pad and a phonological loop for verbal materials—presentations that involve more than one sense modality have increased working memory capacity in comparison to single-format presentations.

Mayer and Sims (1994) also reported an advantage for multimedia presentations with problem-solving tasks. They recently adapted Paivio's (1971, 1986) dual-coding theory to explain multimedia learning. When information is presented verbally and visually, representations of that information are encoded in separate verbal and visual systems within working memory, and referential connections between the two representations are also strengthened. Mayer and Sims predicted that multimedia methods promote the formation of all connections and, as a result, are more likely to promote transfer of information in problem-solving tasks compared with single presentation methods. Data consistent with their theory were obtained when participants with high and low spatial ability viewed an animation simultaneously or successively with a narration. Among the high spatial learners only, problem solving was better when the materials were presented simultaneously rather than successively (Mayer & Sims, 1994).

In the present study, I considered whether the multimedia presentation advantage that had been demonstrated with instructional materials could be found
with simple problem-solving items that required reasoning from background material. Such a finding would suggest that the effect could be generalized beyond instructional items to include a variety of reasoning tasks. Two kinds of problem-solving items—probability judgments with colors and category inclusion—were used as the stimulus material. (See Table 1 for examples of each.) Two kinds of items provided at least two contexts for the investigation of format effects. A multimedia advantage with both would suggest effects that were characteristic of all items. Probability judgments with colors seemed, at least at an intuitive level, to depend more on a visual representation, whereas the category inclusion items emphasized membership in categories. Each item consisted of three lines of background information (in pictured, printed, or spoken formats) and a test sentence. To respond accurately to the test sentence, participants had to make an inference from the three lines of information.

Previous work (Goolkasian, 1996) showed a picture advantage when the background information was presented in either picture or printed word format. The present study extended that effort by including (a) a spoken version among the single-format conditions and (b) combined conditions in which two or three single versions appeared together. Each version of the background contained the same information but differed in presentation format.

Multimedia presentation effects were assessed in at least two ways. First, by comparison of response times (RTs) in the single and combined format conditions, any advantage of multiple presentations either across format (pictured/printed word) or across modality (pictured/spoken or printed/spoken word) would result in a shortened RT for that condition when compared with its respective single-format condition. Although multimedia researchers emphasize the advantages of integrating materials across modalities, others have found an advantage to combined presentation of formats within one sense modality. An advantage to combined presentation of text and pictures has been predicted by

<table>
<thead>
<tr>
<th>Kind of item</th>
<th>Background information</th>
<th>Test sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability judgment with colors</td>
<td>7 red squares</td>
<td>True: Yellow squares are least likely</td>
</tr>
<tr>
<td></td>
<td>5 blue squares</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 yellow squares</td>
<td>False: Red squares are least likely</td>
</tr>
<tr>
<td>Category inclusion</td>
<td>3 horses</td>
<td>True: There are fewer flowers than animals</td>
</tr>
<tr>
<td></td>
<td>3 cows</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 roses</td>
<td>False: There are fewer animals than flowers</td>
</tr>
</tbody>
</table>
those researchers (Glenberg & Langston, 1992; Hegarty & Just, 1993) who believe that participants are constructing mental models and using these propositional representations in reasoning about the material. Hegarty and Just explained that the advantage of the combined presentation comes from the fact that the diagrams act as an external memory aid. When a picture is used, compared with other presentation formats, memory resources are not needed to visualize the display, and there is more processing capacity available for information acquisition. As a result, representations based on pictures are richer and more elaborate than representations based on text alone.

Second, multimedia presentations were assessed by taking into consideration the format difference between the background information and the test sentence. For each stimulus item, the background appeared in one of several formats, but the test sentence was always printed words. When the background appeared as a printed word, participants could integrate the background material with the test sentence, using a common format. When the background appeared in another format, however, this process was not possible because of the format difference between the pictured and spoken information and the test sentence. According to Sweller et al. (1990), when the background appeared in a format other than the printed word version, participants were required to split their attention among multiple sources of information, and some mental integration would be required before an accurate response could be provided. This was especially the case with the spoken version because it represented another sense modality. The split attention effect was expected to influence RTs depending on format condition. When the background appeared in pictured or spoken format, responding to the test sentence required integration of material across formats. Any effects due to the split attention to two modalities or two formats would be evident in comparisons with the printed word condition.

In Experiment 1, I investigated the underlying perceptual differences across format by presenting the background information and the test sentence together in the same display with varied stimulus onset asynchronies (SOAs). In the simultaneous condition, it was assumed that the test sentence would guide the acquisition of information from the background, but with the other SOA conditions some preview of the background would have occurred prior to the presentation of the test sentence. In Experiments 2 and 3, however, the influence of format on memory processes was studied by presenting the background for 6 s prior to the presentation of the test sentence. When responding to the test sentence, participants had to rely on a memorial representation of the background information.

The basic questions concerned the format effect. When a person responds to an inference statement, does the format of the background information matter? Can we reason just as efficiently from spoken information as from pictured or printed words? What advantage if any would be provided by the combined format conditions? If presenting information in two sensory modalities (visual and auditory) enhances working memory capacity (Sweller et al., 1990) or strengthens ref-
erential connections between internal representations (Mayer & Sims, 1994), then the combined conditions that include visual and auditory modes would provide some advantage over the other dual or single presentation conditions.

EXPERIMENT 1

The format of the background varied such that information was presented in pictured, printed word, or spoken word formats; in combined conditions the same material was repeated in two or three different formats. The background information remained on the screen, and the test sentence appeared with varied SOAs from 0 to 1500 ms. By manipulating the onset of the test sentence, I investigated the effect of format on the time course of processing the background material. Larkin and Simon (1987) suggested that picture/word differences may result from perceptual differences in the explicitness of information. For example, making an inference from a word format may require more reasoning than from a picture because pictures are more direct. Support for such an explanation would be found if there were a picture advantage irrespective of SOA condition or single versus combined presentation condition.

The participants made a true–false speeded response to the test sentence. Each test sentence was an inference presented in printed word form. Both true and false statements for each of the two problem-solving items were used. The analyses tested for main and interaction effects of format, SOA condition, test statement accuracy, and kind of problem-solving item.

Method

Participants

The participants were volunteers from the University of North Carolina, Charlotte, who had normal or corrected to normal (20/20) vision and no evidence of color blindness. Students participated in only one of the experiments and obtained extra credit points toward their psychology class grade. Experiment 1 involved 20 men and women, Experiment 2 involved 33 men and women, and Experiment 3 involved 30 men and women.

Stimulus Materials

Pictured, printed word, and spoken word versions of the background were developed for each kind of problem-solving item. Figure 1 presents an example of the pictured and printed word formats. The spoken version accessed a sound file with a female voice speaking the printed word version. The single-format conditions presented the background information once, whereas the combined conditions presented the same background material in two or three formats. In the
spoken-word-alone condition, the sound file was heard against a blank screen, whereas in the combined conditions the sound file played as the pictured and/or printed word versions of the background were displayed. Also, in the single-format conditions, the background information was centered on the screen, whereas in the combined condition, each of the formats appeared side by side and the left/right placement of picture and printed words was counterbalanced across stimulus items.

As much as possible, the sizes of the pictured and printed word backgrounds were equated. In all cases, when viewed from a distance of 30 cm, the stimuli...
were larger than 3 degrees of visual angle. The specific dimensions of each of the stimuli are identified in Figure 1. The sizes of the picture and word versions of the probability judgment with color stimuli were approximately the same.

The test sentences were developed from problem-solving items identified by Brainerd and Reyna (1993). Table 1 contains some true and false examples. The inference statements required some reasoning—that is, the statements were not explicitly presented in the background information. In the probability judgment item, the inferences required participants to make judgments of which color was most or least likely. In the category inclusion example, statements questioned whether the background input contained more or fewer examples of members of a particular category.

The stimuli were displayed on an Apple Color High Resolution RGB 13" monitor. The monitor had a P22 phosphor with a medium-short persistence. Stimulus presentation and data collection were controlled by SuperLab running on a Macintosh II computer.

Procedure

Each trial consisted of three stimulus events. A fixation point and mask appeared for 5 s, followed by background information and then the test sentence. Figure 1 shows the mask (5.5 cm × 8 cm) that covered the stimulus display. The procedure varied such that the test sentence appeared either together with the background information or after a delay (750 or 1500 ms). To control precisely the time to encode the background information across format, a 200-ms wait was used before presenting the background information because it took longer to draw the pictured than the printed word version to the screen. The inclusion of the wait at the beginning of each trial kept the screen invisible until the image was drawn and ready to be seen. Response times measured the time period between the presentation of the test sentence and the keypress response.

Participants were seated so that their eyes were 30 cm from the monitor. They were tested individually in sessions of approximately 45 min. They used a chin rest to stabilize their head movements, and they were instructed to study the material presented and to respond to the test sentence as quickly as possible without sacrificing accuracy. Instructions indicated that the background information would be spoken, pictured, presented in a printed word format, or some combination of these conditions. Moreover, the participants were made aware that in the combination conditions the same background information would be presented in different formats. When the test sentence appeared, participants were told to respond "true" if the test sentence contained material that could be inferred from the background material and to respond "false" otherwise. Responses were made by pressing T or F on the keyboard. There were six practice trials prior to the experiment. The screen locations of the background information and the test sentence permitted both to be viewed simultaneously. The background informa-
tion appeared in the upper center portion of the screen, and the test sentence appeared in the lower center.

There were 336 trials, representing four replications of 84 experimental conditions. Each participant received a random arrangement of trials that represented the seven format conditions factorially combined with two kinds of items, true and false statements, and three SOAs.

**Results**

In each of the experiments, means were computed from the correct RTs obtained from each participant across the 4 trials within each of the experimental conditions. RTs in excess of 6 s (less than 2% of the responses) were not included in the analyses. Also recorded were the incorrect responses. The $F$ tests included the Geisser–Greenhouse correction to protect against violation of the homogeneity assumption. A $7 \times 2 \times 2 \times 3$ repeated measures analysis of variance (ANOVA) was used on the data collected in Experiment 1 to test for the effects of format, statement accuracy, kind of item, and SOA interval.

RTs were found to decrease with increasing SOA, $F(2, 38) = 270.91, MSE = 275853.84, p = .0001$. Mean RTs in order of increasing SOA were: 2877 ms, 2377 ms, and 2160 ms, respectively. As expected, responses were quicker when the background information appeared in advance of the test sentence. Also, there were notable differences in RTs across the seven format conditions, $F(6, 114) = 165.95, MSE = 487489.53, p = .0001$.

Figure 2 presents the mean RTs for the significant format by SOA interaction, $F(12, 228) = 3.66, MSE = 173451.23, p = .001$. The absence of a three-way interaction with format and SOA, however, shows that the interaction was consistent across the two kinds of items. There was neither a Format $\times$ Kind of Item interaction, $F < 1$; nor a Format $\times$ SOA $\times$ Kind of Item interaction, $F(12, 228) = 1.90, MSE = 129885.35, p = .09$; nor a Format $\times$ Statement Accuracy $\times$ Kind of Item $\times$ SOA interaction, $F < 1$.

Tests for simple effects of format were significant ($ps = .0001$) at each of the SOA conditions. Follow-up post hoc tests of the format effect ($ps < .05$) showed that under the simultaneous presentation condition (0 SOA), the RTs were grouped according to whether the single or combination conditions included pictured, printed word, or spoken word formats. Responses to pictures were the quickest, followed by responses to printed words. Spoken materials took the longest because participants had to wait several seconds to hear enough information to respond. Background information was delivered instantaneously in the other displays. The Format $\times$ SOA interaction indicated that with increasing SOA there is a narrowing of the RT difference across format. For example, at the longer SOAs, the picture/printed word difference was still significant but considerably smaller than the difference observed at 0 SOA.

Post hoc comparisons within each of the SOA conditions also tested com-
FIGURE 2. Mean RT and errors as a function of format and SOA in Experiment 1. The error bars represent standard errors.

Comparisons among relevant single and combined formats. In general, when background information appeared in multiple formats, participant RTs were similar to whichever single format included in the combination led to the quicker response. RTs in the picture/spoken word condition were not different from RTs in the picture-only condition, and RTs in the printed word/spoken word condition were not different from RTs in the printed word condition. In some instances (such as the picture/spoken word/printed word and the picture/printed word condition), the
combination condition resulted in significantly longer RTs when compared with the relevant single-format condition.

Consistent with previous findings (Goolkasian, 1996), true statements were responded to more quickly than false statements, $F(1, 19) = 19.42, MSE = 246803.30, p = .0003$. Mean RTs for true and false statements were 2424 ms and 2531 ms, respectively. Statement accuracy also interacted with format, $F(6, 114) = 2.93, MSE = 132294.48, p = .02$. Format differences were consistent with both statements; however, the effect was larger when statements were true rather than false.

There were also differences in RTs to the two kinds of items, $F(1, 19) = 16.71, MSE = 1131484.28, p = .0006$. Mean RT for the probability judgments was 2371, versus 2583 ms for the category inclusion item. Kind of item interacted with SOA condition, $F(2, 38) = 17.53, MSE = 181317.34, p = .0001$, and with statement accuracy, $F(1, 19) = 7.65, MSE = 108791.59, p = .01$. Simple effects of kind of item at each of the SOA conditions showed that making probability judgments about colors was quicker than category inclusion responses, but only under the two shortest SOA conditions ($ps < .05$). When the background information appeared 1500 ms prior to the test sentence, RTs for two kinds of items did not differ.

Analysis of the errors showed significant main effects of format, $F(6, 114) = 5.87, MSE = .01, p = .0002$, and SOA, $F(2, 38) = 4.11, MSE = .01, p = .02$. As can be seen in the bottom panel of Figure 2, these effects result from a decline in the error rate with increasing SOA and an increase in the error rate for some of the format conditions that included spoken words. Average error rates for the spoken words, printed words/spoken words, and printed words/spoken words/picture conditions were 7% versus 3.5% for the other format conditions. None of the other effects were significant.

**Discussion**

When participants reasoned from background material that was readily available, presentation format and the time for preview of the background material influenced RTs. Consistent with results of previous studies (Bauer & Johnson-Laird, 1993; Goolkasian, 1996), the data show a considerable pictorial advantage. For both kinds of items, reasoning was facilitated under the single and combination conditions that included a picture format in comparison to the other formats. Because it takes time to present spoken materials, spoken formats took the longest to process and produced the most response errors. Participants needed to wait several seconds to acquire enough information to answer the test questions, whereas in the other formats the information was available instantly. Experiments 2 and 3, which used a 6-s presentation time for the background material, provided another test of the auditory format without this problem.

The data are consistent with those of Larkin and Simon (1987), who suggested that pictures provide a more direct access to information when compared
with printed words. More time was needed to reason when information appeared as printed words rather than being pictured. However, this explanation does not account for why the picture advantage decreased with SOA. The facts that the pictured and printed word difference (a) was largest when both background material and test sentence appeared simultaneously and (b) showed some evidence of decreasing when participants were allowed some time to process the background material in advance of the presentation of the test sentence, suggest that presentation format was influencing the way that the information was acquired and initially processed. Given the obvious picture advantage in encoding, in the next two experiments I explored the degree to which format may influence later processing stages, such as those that may be associated with responding to a test sentence from memory.

The finding that was of particular interest, however, was the comparison among the single and combined format conditions. The combined format conditions resulted in RTs that were either the same as or a little longer than a comparable single-format condition. When making an inference from background materials that are readily available, redundancy of the information does not provide an advantage. This was especially the case when the combined condition included all three formats.

EXPERIMENT 2

In Experiment 2, the background information appeared for 6 s and was replaced by the test sentence. I expected that because the presentation time for the background was long enough to allow complete encoding of the material prior to receiving the test sentence, differences in encoding pictured, printed word, or spoken information would not influence RT. Format differences in RTs should result only from the processing that followed the encoding of the background material. Such differences in the way that pictures and words are stored have been identified (Glenberg & Langston, 1992; Paivio, 1971, 1975, 1978). For example, mental models suggest that representations from pictures are richer and more elaborate than representations from other formats (Glenberg & Langston, 1992; Hegarty & Just, 1993). Findings consistent with this explanation would show that RTs to test sentences vary across format conditions. The analyses tested for the main and interaction effects of format, test statement accuracy, and kind of problem-solving item.

Method

Although the stimulus materials were the same as in Experiment 1 and the same three events occurred on each trial, the procedure varied. The fixation point and mask appeared for 500 ms, followed by the background information for 6 s. Pilot tests showed that 6 s were sufficient for the participants to encode
any one of the single or combined versions of the background input. Then a test sentence appeared and remained on the screen until the participant made a key-press response. The screen locations of the background input and the test sentence were adjusted so that both were in the center of the screen (as compared with Experiment 1, in which both appeared centered in the top and bottom of the display).

There were 112 trials that represented four replications of 28 experimental conditions. Each experimental session consisted of a random arrangement of trials that represented the seven format conditions factorially combined with both true and false statements and two kinds of problem-solving items.

Results

Data from 2 participants were excluded because error rates produced below chance accuracy in 3 or more experimental conditions. A $7 \times 2 \times 2$ repeated measures ANOVA was used on the RT and error data to test for effects of format, statement accuracy, and kind of item. The ANOVA on the RTs showed significant main effects of format, $F(6, 180) = 5.39, MSE = 147717.23, p = .0003$; statement accuracy, $F(1, 30) = 7.85, MSE = 317106.06, p = .009$; and kind of item, $F(1, 30) = 4.61, MSE = 527981.79, p = .04$. Kind of item was found to interact with format, $F(6, 180) = 5.39, MSE = 156995.06, p = .0003$, and with statement accuracy, $F(1, 30) = 7.80, MSE = 144561.27, p = .009$. None of the other interactions were significant.

Figure 3 presents the interaction of Format $\times$ Kind of Item. Tests for simple effects of format with each kind of item show significant differences in RTs among the format conditions with both kinds of items ($ps < .05$). Format differences were much larger, however, when participants were making category inclusion judgments than when they were asked to make probability judgments with colors. Post hoc comparisons of the format conditions (at the $p < .05$ level of significance) for both items showed that spoken words provided a response advantage in comparison to pictures. The picture advantage over printed words was significant with category judgments but not with probability judgments about color. In addition, a comparison of the combined with the single-format conditions did not result in any noticeable advantage to multimedia presentation formats. Unlike the results of Experiment 1, however, RTs in the combined format conditions were not similar to whichever single-format condition included in the combination led to the quicker response. For example, the combination conditions that included spoken words resulted in significantly longer RTs than the spoken word condition.

Consistent with past findings, false statements took longer than true, and the true–false difference was larger with the probability judgment item than with category judgments. Also, there was an overall difference in responding to each kind of item that was in distinct contrast to the results of Experiment 1. The category
inclusion item was responded to more quickly than the probability judgment item. Mean RTs for each item were, respectively, 2.088 s versus 2.194 s. A possible explanation for the inconsistency between the results of Experiments 1 and 2 may be found in the different procedures that were used. The main effect of kind of item may reflect a difference in item difficulty, and it is possible that this changed between the experiments. In Experiment 1, probability judgments were easier.
than category inclusion items because participants were reasoning from backgrounds that were readily available and probability judgments seemed to require a more visual judgment. In Experiment 2, however, when participants were responding to test sentences from memory, category judgment items were responded to more quickly—perhaps because of the conceptual nature of the task. The fact that kind of item interacted with SOA interval in Experiment 1 and showed a more pronounced difference with no or short SOA intervals supports the interpretation that the amount of time that was available for processing the background prior to the test sentence had a significant influence on RTs. When the background information appeared 1500 ms in advance of the test sentence, there was no difference in RTs to the two kinds of items. Because the procedure used in Experiment 2 reflected memory differences rather than perceptual effects, category inclusion may have been an easier kind of item.

The error analysis showed only an effect of Format × Kind of Item, $F(6, 180) = 3.66, MSE = .01, p = .003$ (see bottom panel of Figure 3). Error rates across all format conditions represented less than 6% of the responses, with the exception of spoken words in probability judgments with color and printed words and printed words/spoken words in category inclusion items. This finding partially replicates previous work (Goolkasian, 1996) in which it was shown that more errors were made when participants made inferences from printed words compared with picture formats. There were no significant main effects of format, $F(6, 180) = 1.70, MSE = .01, p = .14$; statement accuracy, $F(1, 30) = 2.34, MSE = .02, p = .14$; or item, $F < 1$, nor were any of the other interactions significant.

**Discussion**

Although format effects were more evident with category inclusion items than with probability judgment items, the data show an unexpected advantage to spoken background material. The advantage of picture over printed words obtained in previous research was evident with only the category inclusion item. There are several reasons why these findings may have occurred. The spoken information differed from the other formats in the way that the material was distributed over the 6-s presentation time. With both the pictured and printed word formats, the background material was presented all at once and was available for the entire 6 s, whereas the auditory format presented the background word by word across the 6-s presentation period. Participants may have been forced to process the spoken material as it was presented, whereas with the other formats, the material was available all at once and participants could have processed it in any order. Interestingly, the combination conditions that included a spoken format did not lead to faster RTs. So, when given a choice, participants do not rely on spoken words for processing the background material, even though this format resulted in faster responses in comparison to others. Because error rates were not uniformly elevated in the spoken word format condition (with the exception
of true probability judgment statements), the RT advantage cannot be attributed to a lower accuracy rate. Experiment 3 was conducted to test the hypothesis that distributing the background information across the 6-s presentation interval would provide some advantage to responding.

The finding that was of particular interest in this study, however, was in the comparisons among the single and combined format conditions. Even though some combined format conditions produced RTs that were as quick as those in the single-format conditions, in no case did any of the combined conditions result in performance advantages that were not available when background information appeared in single formats. When reasoning from background materials about probability judgments of color or making category inclusion inferences, redundancy of the background information across modality (picture/spoken word, printed word/spoken word) or across format (picture/printed word) did not provide a benefit. This was especially the case when the combined presentation included all three formats. The advantage of multimedia presentation identified with instructional examples drawn from math, science, and technology (Mayer & Sims, 1994; Sweller et al., 1990) did not extend to the simple problem-solving items used in these experiments.

However, the finding that the spoken version provided some RT advantage among the single-format conditions is consistent with the split attention effect of Sweller et al. (1990). Responding to the test sentence was faster when it involved integration of spoken rather than pictured or printed word formats. The RT benefit may have resulted from an increase in working memory capacity that accompanies dual presentation modalities, as suggested by Sweller et al., or it could have resulted from the gradual presentation of the background material. Experiment 3 was conducted to address this issue.

The absence of a clear picture/word advantage in the RTs to the probability judgment with color items was puzzling, given the previous work (Goolkasian, 1996), but most likely resulted from the fact that seven format conditions were tested.

**EXPERIMENT 3**

Experiment 3 tested whether the sound advantage found in Experiment 2 was due to the gradual presentation of the background material as compared with the simultaneous presentation in the other formats. Only single presentation conditions were used and, in addition to the formats described in the previous experiments, there were picture and printed word conditions in which the background information appeared line by line in successive 2-s displays. An advantage in gradual presentation of the background material should be shown with these conditions as well as in the spoken word condition. As in Experiment 2, the test sentence included both true and false inference statements with the probability judgment with colors and category inclusion items.
Method

The stimulus materials and procedure were the same as in Experiment 2, except that only single-format conditions were used. The pictured, printed word, and spoken versions of the background were used together with two additional formats in which the pictured and printed word material were divided into three separate displays, each with a line of information. The stimulus materials in the line-by-line conditions were identical in every respect to the original displays except that each of the three lines was presented in successive 2-s time frames rather than all at once. Each line replaced the preceding one, so that the participant saw only one line at a time during the 2-s display interval.

There were 80 trials that represented four replications of 20 experimental conditions. Each experimental session consisted of a random arrangement of five format conditions factorially combined with both true and false statements and two kinds of problem-solving items.

Results

A $5 \times 2 \times 2$ repeated measures ANOVA was used on the RT and error data to test for effects of format, statement accuracy, and kind of item. The RT analysis showed significant effects of all three variables. There were main effects of format, $F(4, 116) = 2.98, MSE = 207900.00, p = .03$; statement accuracy, $F(1, 29) = 22.06, MSE = 206900.00, p = .0001$; and kind of item, $F(1, 29) = 4.66, MSE = 454700.00, p = .04$. Figure 4 presents the means for each of the format conditions. They were averaged across kind of item because there were no significant interactions of Format $\times$ Item, $F(4, 116) = 1.52, MSE = 144800.00, p = .21$, or Format $\times$ Item $\times$ Statement Accuracy, $F < 1$, or Format $\times$ Statement Accuracy, $F(4, 116) = 1.09, MSE = 143000.00, p = .36$.

Post hoc comparisons (at the $p < .05$ significance level) within the format effect showed that when participants were reasoning from information stored in memory, the line-by-line picture representation was faster than either the spoken or printed word condition. However, the line-by-line picture condition did not differ from the original picture condition.

The other effects were consistent with those obtained in Experiment 2. Responses to false statements took longer than responses to true statements (mean RTs were 2347 ms versus 2172 ms, respectively); and reasoning with probability judgments items took longer than category inclusion problems. The mean RTs for the two kinds of items were, respectively, 2319 ms versus 2200 ms. There was also a significant interaction of kind of item and statement accuracy, $F(1, 29) = 10.32, MSE = 156000.00, p = .003$, which showed a larger difference between true and false statements with probability judgment items than with category inclusion items.

The analysis of the errors showed an average error rate of 8.5%, and this
rate did not vary by format, $F(4, 116) = 1.31, MSE = .014, p = .27$; statement accuracy, $F(1, 29) = 2.63, MSE = .046, p = .12$; or kind of item, $F < 1$. The only significant effect was an interaction of Statement Accuracy x Kind of Item, $F(1, 29) = 16.12, MSE = .019, p = .0004$. This effect resulted from an increase in the error rate to 12% in response to true category inclusion statements when compared with the other conditions. None of the other effects in the analysis reached significance.
Discussion

The data from Experiment 3 show that presenting the background information gradually (line by line) rather than all at once can have advantages in this task, but only with the picture format. The gradual presentation of the word format did not significantly influence responses. Part of the spoken word advantage in Experiment 2 may have resulted from the piecemeal manner in which the information was presented across the 6-s presentation interval. The format effect from Experiment 3 is also more consistent with previous findings (Goolkasian, 1996) when compared with the findings of Experiment 2. Among the single-format conditions, pictures were processed more quickly than were printed words.

GENERAL DISCUSSION

Taken together, the results of the three experiments show a consistent picture advantage. The advantage was considerable when the task measured the format effect on the acquisition of the background material. However, when participants were reasoning from material stored in memory, the format effect was much smaller, and spoken materials provided a performance advantage that was similar to pictures. Combining formats so that the background material would appear in two or three ways did not significantly influence RTs when compared with single-format conditions.

The findings support the supposition that pictures provide a more direct access to information, as suggested by Larkin and Simon (1987). The data from Experiment 1 showed a distinct processing advantage when participants reasoned from pictured versions of the background in comparison to the other formats. Even though all versions of the background consisted of the same information, problem solving from pictured material was quicker than from the verbally presented material. When the test sentence appeared simultaneously with the background, extracting information from the pictured version was quicker than in all other representations. However, such a simple theory does not go far enough to explain why the picture advantage decreased with SOA condition in Experiment 1 and why format effects were present in Experiments 2 and 3. If the format advantage were simply a matter of the difference in access to information, then the format effect would have been consistently large across SOAs in Experiment 1 and would have disappeared in Experiments 2 and 3 when the procedure emphasized the processing that followed encoding.

An explanation of the format effects in Experiments 2 and 3 may be related to short-term memory differences suggested by the split attention effect of Sweller et al. (1990). Problem solving was faster when participants were required to integrate materials across format or sense modalities. The RT benefit was demonstrated across sense modalities in Experiment 2 (with the sound advantage) but across format in Experiments 1 and 3 (with the picture advantage). Sweller
and his associates suggested that the load on working memory is reduced because dual mode presentations may have increased working memory capacity in comparison to single-format presentations.

A similar emphasis on increased processing capacity as the underlying explanation for format effects can be found with predictions from mental models (Glenberg & Langston, 1992; Hegarty & Just, 1993). For these researchers, pictures facilitate reasoning because they act as an external memory aid, freeing up processing resources in a manner not matched by the other presentation formats. As a result, representations formed with pictures are richer and more elaborate than representations of verbal materials. Such an explanation can be used to partially explain why pictures, whether presented alone or in combination with other formats, were responded to more quickly than printed words. When this explanation is combined with Sweller and colleagues' split attention effect, the format effects from Experiments 2 and 3 make sense. Experiment 3 showed that some of the benefit of the spoken material was in the piecemeal manner in which the information was presented, but data from Experiment 3 show clearly that both pictured and spoken formats provide some response advantage relative to printed word representations. Presenting the background material line by line may produce an advantage because it forces participants to encode the material as it is presented, providing equal time to each section of the display. Interestingly, however, the line-by-line presentation did not facilitate responding when words were used as the presentation format. Printed words proved to be the least effective.

However, the problem-solving benefit from presenting learning materials in dual formats was not evident with these simple problem-solving items. There is no evidence from either Experiment 1 or 2 that presenting information in two or three different ways facilitates problem solving. The fact that redundancy was not important suggests that, unlike dual coding theory, which places emphasis for format effects on the manner in which information is stored, the underlying mechanism for the format effect is in some differential manner of acquiring information from picture, word, and sound. Pictures may have an advantage in the perceptual layout of the information because more stimulus features may be represented than with verbal formats. However, this alone is not the only process underlying format effects. Mental model theory, with its suggestion of differential picture/word processing during encoding, when combined with Sweller and colleagues' split attention effect when materials are presented across sense modality, come closest to an explanation for these data. Multimedia learning theory (Mayer & Sims, 1994) may provide an explanation for problem-solving tasks that require a higher level of reasoning (or reasoning with more complex materials) than that tested with the sentence verification task used in the present experiments.

The answer to the question raised at the outset, regarding which format provides the most effective way to reason, would need to be a qualified one. When people reason from materials that are readily available, there is no doubt about the
picture advantage with both RTs and errors. Spoken materials have somewhat of a disadvantage within this context because of the need to present information across time. When the presentation of the test sentence requires reasoning from material stored in memory, however, both spoken and picture formats provide similar advantages. If one were to make a generalization to a classroom situation from these findings, the prediction would be that when a test statement is presented that requires some reasoning from material stored in memory, there is some slight edge to materials that were originally spoken and to pictured materials.

REFERENCES


*Manuscript received August 30, 1999*

*Revision accepted for publication March 1, 2000*