This study investigated four types of media formats: (1) talking head; (2) voice-under-text; (3) voice-under-video; and (4) multiformat (combination of talking head, voice-under-video, and text). Sixty-four suburban middle-school students were shown four different narrative stories, each presented in a different media format. Students were tested to determine what they remembered through verbal recall and a spatial task. Memory differed depending on format. The formats that most enhanced memory were the primarily visual formats (multiformat and voice-under-video styles). Students were less apt to remember stories presented in the primarily audio (voice-under-text and talking head) formats. In contrast, initial analysis of spatial recall was inconclusive. This may have been due more to the inadequacy of the scoring scheme than real differences in the drawings. Some suggestions are made for improving scoring through coding spatial relationships and verbal responses. Four tables present study findings. (Contains 40 references.) (Author/SLD)
Title:

The Effects of Multimedia Presentation Formats on the Spatial Recall of a Narrative

Authors:

Barbara Fields Ottaviani
Hunter College of the City University of New York
695 Park Avenue
New York, NY 10021

John B. Black
Teachers College
Columbia University
New York, NY
This study investigated four types of media formats: talking head, voice-under-text, voice-under-video, and multiformat (a combination of talking head, voice-under-video, and text). Sixty-four suburban middle school students were shown four different narrative stories: each story was presented to the students in a different media format. The four stories were exactly the same except for the format of presentation. The students were tested to determine what they remembered about each story through verbal recall and a spatial task. The verbal recall results in this study indicated that memory of a narrative differed significantly, depending on whether it was shown in a talking head, voice-under-text, voice-under-video or multiformat presentation. More particularly, the formats of presentation which appeared to most enhance memory of a narrative were the primarily visual formats (multiformat and voice-under-video). Students were less apt to remember stories presented in the primarily audio formats (voice-under-text and talking head). In contrast, initial analysis of students' spatial recall was inconclusive. However, we argue that this was due more to the inadequacy of the spatial scoring schemes to reflect the real differences in the drawings. This paper reports the verbal recall findings of the study and suggests how the spatial task findings can be reviewed and evaluated to further an understanding of the relationship of format of presentation to understanding.

INTRODUCTION

A multimedia presentation combines the inherent media attributes of different electronic technologies to produce a message which uses both auditory/verbal and visual/pictorial channels of information processing. Just as the choice of a medium is important to how a message is received and understood, so too is the choice of the format of presentation within that medium because of the possible effects it may have on the perception and ultimate understanding of the message. Research on the interplay between the auditory/verbal and visual/pictorial channels of information processing suggests that understanding of electronic media presentations is determined not only by the clarity of the message being presented, but also by the associations made between the two cognitive channels during processing.

There have been many investigations comparing children's understanding of visual and verbal information presented in television programs. Some have concentrated on defining the distinctions between iconic and symbolic modes of representation (Bruner, 1966; Glass, Holvoak & Santa, 1979). Other research has concentrated on the behavioral issues associated with visual-auditory presentations, such as vocal characteristics (Shosteck, 1973; Burgoon, 1978; Beighley, 1952; Hadwiger, 1970; Smith & McEwen, 1973) or eye contact with the camera lens (Tankard, 1970).

The inherent characteristics of a newscast itself also have been found to influence a viewer's perception. For example, viewers are more apt to believe (Baggaley, 1980), agree with (Allen, 1973), or recall unfamiliar information (Davey & Kapinus, 1985) presented first in a newscast. The pace or action of a televised presentation also appears to influence understanding: children are less apt to understand and recall televised information that is presented through dialogue than content presented with moderate to high rates of character action (Calvert et al., 1982; Hayes & Birnbaum, 1980; Watkins, Calvert, Huston-Stein & Wright, 1980). On the other hand, increasing the video pacing of a film does not appear to affect recall (Schlater, 1970). It has also been shown that there is a relationship between the visual and verbal memory of televised messages and the emotional valence of that message (Fiske, 1982; Bradley, Greenwald, Petry & Lang.
The left hemisphere of the brain is associated with positive messages and the right side of the brain is thought to process negative messages best (Lang & Friestad, 1993; Davidson, 1985). Furthermore, since it is generally believed that the left hemisphere of the brain is the verbal processor and the right hemisphere is the visual–spatial processor (Kinsbourne, 1982), there have been suggestions that there is a relationship between the valence of the message and its format of presentation with regard to the memory (Reeves, Lang, Thorson and Rothschild, 1988).

A tremendous amount of television research has focused on what Salomon (1974) refers to as the "inherent symbolic media attributes". In general these media attributes can be defined independently of the narrative content (Covnil, 1974). For example, cuts have been shown to increase interest in a film (Kraft, 1986) and serve as visual punctuation (Carroll & Bever, 1976) as well as a method of setting the pace of the film (Giannetti, 1982).

Viewer comprehension of an electronic media presentation is influenced by many factors and interactions including how the information is organized and linked to knowledge structures in long term memory. In general, these frames (Minsky, 1975); scripts (Shank and Abelson, 1977); or schemata (Anderson, Spiro and Anderson, 1978), as they are variously referred to in the literature, are used to facilitate a subject's ability to make inferences about a concept. Some researchers believe that to understand the influence electronic media has on learning it is important to understand the skills needed to process the form of a message as well as its content (Clark, 1975; McLuhan, 1964, Olson & Bruner, 1974). These skills include visual and verbal analysis, perspective taking and spatial manipulations. The media attributes which affect knowledge acquisition are not limited to the physical representation of the message. When an individual is presented with information about an event, a different set of rules will be applied if that event is read in a newspaper rather than presented on radio (Williams, Paul & Ogilvie, 1957; Wilson, 1974), or if the message is seen on television in a talking head format, rather than in a video format (Katz, Adoni, et al., 1977; Edwardson, Kent & McConnell, 1985). Furthermore, research suggests that there are hemispheric differences in information recalled from global versus local processing (Hansen, 1981) or positive versus negative emotions (Davidson, 1985).

The effect of media attributes on the message is perhaps most pronounced in multimedia productions because several electronic media forms are combined to create one presentation. The attributes of one medium interact with those of another and it is these interactions that ultimately form the attributes of multimedia. Therefore, it is necessary to analyze the parts of a multimedia presentation, such as format of presentation, to define the whole.

The formats of presentation examined in this experiment included those most often used in multimedia productions: talking head, voice–under–text, voice–under–video and a multiformat combination of all three. These formats are familiar to students as they are frequently used individually or in tandem in all electronic media presentations. The question asked in this experiment was a simple one: would children remember more about a narrative if it was presented in one format rather than another. Specifically, when subjects were given simultaneous, or near simultaneous signals from multiple sources.
how were those stimuli encoded by the subject to be either stored and utilized in long term memory, or discarded as irrelevant?

Method

Subjects

Sixty-four students from Westchester County, New York participated in the experiment. The 32 boys and 32 girls participating in this experiment ranged in age from 12 years 7 months to 15 years 1 month and were racially and ethnically mixed to include 51 White, 5 African-American, 4 Latino and 4 Asian children. English was the first language for all the students. The students were evenly distributed to view the four sequences of the experiment: eight boys and eight girls saw each story once and each media format condition once.

Materials

Formats of Presentation

Each story was produced in four formats of presentation: talking head, voice-under-text, voice-under-video and multiformat, which is a combination of the three. The talking head was taped in a studio with a noninterfering blue backdrop. A female announcer read from cue cards placed near the camera lens to replicate the style of most video announcers. She did not appear to be reading the story from a text and she looked straight into the lens to increase believability and influence recall (Tankard, 1970). The pace of the presentation was between 160 and 180 words per minute, within the ideal range for memory (Schlater, 1970). The soundtrack from this initial taping was used for the other three presentation formats. Thus, the audio, including vocal inflections and timing, was consistent across all four experimental conditions.

The voice-under-text used in the experiment was designed to suggest an interactive computer screen. The text was shown in serif font with white letters on a blue background. Each “page” was numbered at the bottom of the screen and there were ten lines of text on each “page”. This was in keeping with the recent literature on screen design (Faiola and DeBlois, 1988).

The voice-under-video format was produced by first laying down the soundtrack on a 3/4" tape. Video which was redundant with the audio of the story was then copied onto the soundtrack. For example, if the text described a man at work on a sewing machine, a video showing a man working at a sewing machine was shown.

The final format used in the experiment was produced by combining talking head with the voice-under-video and then enhancing the film with key points of the story by using captions over the video. These captions were chosen from the main headings of a conceptual map developed from the text.
Each of the testing materials designed and used for this study were unique: each of the formats of presentation had their own individual characteristics, and each of the narratives were organized in a different structure or with a different purpose. The formats of presentation were categorized as being either primarily audio or primarily visual. The talking head and voice-under-text were primarily audio because the main focus of the viewer of a presentation in either of these formats was on the audio portion of the narration. In other words, the visual representation of the story through text on the screen provided additional information, but was not the primary focus of the viewer’s attention. Similarly, the eye contact of the narrator may have provided an additional incentive for the viewer to listen to the story, but no additional information about the narration was projected.

The primarily visual formats identified in this study were the voice-under-video and multiformat presentations. In each of these formats, the primary focus of the viewer was the image on the screen. Admittedly, the audio portion of the film provided a structure to the images being projected on the screen, but it was the visual image which seemed to enhance cue recognition for information processing and retrieval. In the voice-under-video presentation there was no break in the visual images being projected over the audio narration. This was not the case in the multiformat presentation because the video portion was interspersed with clips of the talking head and captions of text. Theoretically, the two formats should compliment each other by directing and maintaining the attention of the viewer.

The Stories

In general, understanding and remembering the events of a dramatic story utilize the same processes people empower to understand and recall other events and presentations (Bower and Black, 1980). Since narrative, dramatic stories have proven to be valuable in testing children’s understanding of media presentations (Meringoff, et al., 1983). Four original narratives were written for this experiment. The stories were intentionally diverse to eliminate any confusion when subjects were asked to recall the presentation. The four narratives were: “A Child is Taken”, “The Strength of Memories”, “Return to the Sea” and “The Sound of Hatred”.

“A Child is Taken”, was about a young child being kidnapped from her home in the middle of the night. Time and space were key factors in understanding the inferences presented in this narrative which sequentially presented the events surrounding the kidnapping and vividly described the scene of the crime. A second story, entitled “The Strength of Memories”, was about a prisoner in Auschwitz during WWII. It was presented as a series of abstract recollections by a man confined to a German prison cell. There was nothing sequential or predictable about the story as it described the man’s girlfriend, his family, and his past career as a champion boxer. This was the most abstract of the stories and the one which required the most concentration on the part of the viewer. A third narrative, entitled “Return to the Sea”, was about a beached baby whale. Though time and space were important to this story, the primary structure of the narrative was a problem to be resolved. This was the most positive of the four narratives. A fourth story, entitled “The Sound of Hatred”, was about Apartheid. It was written to be an emotion-evoking depiction of a powerful event: the attack of a South African village by white military. In this narrative,
an ominous tone was set at the beginning of the story when the narrator asked: “What was that sound?”. The subject did not know what to expect, but the scenario implied something negative. The “why” of the event was never given, however inferences were clearly presented about the occurrence to suggest racial prejudice and injustice. It should be noted that three of the stories (“A Child is Taken”, “Return to the Sea” and “The Sound of Hatred”) were more schematically predictable than the other story (“The Strength of Memories”). Also, “A Child is Taken” and “The Sound of Hatred” ended on a very negative emotional tone, whereas “The Strength of Memories” and “Return to the Sea” ended with positive connotations and inferences about what had happened and was expected to happen next.

Procedure

Since each story was produced in four media format conditions, there were sixteen segments to be manipulated for the experiment. Four sequences were made and each story in each format was used once. The order of the format was manipulated in the four sequences of the experiment so that each format was in every position once. No format was in the same position more than once; however, a story may have been in a position more than once. The four presentation sequences used in the experiment are listed in Table 1.
Table 1 — Order of Video Presentation for Each Sequence

<table>
<thead>
<tr>
<th>Ord</th>
<th>Sequence I</th>
<th>Sequence II</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Multiformat A Child is Taken</td>
<td>Talking Head Sound of Hatred</td>
</tr>
<tr>
<td>2</td>
<td>V-U Text Sound of Hatred</td>
<td>Multiformat Return to the Sea</td>
</tr>
<tr>
<td>3</td>
<td>V-U Video Return to the Sea</td>
<td>V-U Text Strength of Memories</td>
</tr>
<tr>
<td>4</td>
<td>Talking Head Strength of Memories</td>
<td>V-U Video A Child is Taken</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ord</th>
<th>Sequence III</th>
<th>Sequence IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>V-U Video Sound of Hatred</td>
<td>V-U Text Return to the Sea</td>
</tr>
<tr>
<td>2</td>
<td>Talking Head Return to the Sea</td>
<td>V-U Video Strength of Memories</td>
</tr>
<tr>
<td>3</td>
<td>Multiformat Strength of Memories</td>
<td>Talking Head A Child is Taken</td>
</tr>
<tr>
<td>4</td>
<td>V-U Text A Child is Taken</td>
<td>Multiformat Sound of Hatred</td>
</tr>
</tbody>
</table>

The experiment was designed to show each student a different story under four different media format of presentation conditions. As a result, each format/story combination was seen by 16 students. The students viewed a video tape of the four stories, completed the personal media information sheet, ranked the titles of the stories and performed four tasks for each story — verbal recall, spatial task, forced-choice recognition and continuation. The students were shown the tape and interviewed individually. The total time to view the tape was 18 minutes. The total time for the experiment was 1 hour and 15 minutes. The dependent variables used in this study were: the proportion of propositions recalled verbally; the proportion of forced choice recognition questions answered correctly; the proportion of propositions recalled in the spatial task; and the type of elaborations made in the continuation of the story. As stated previously, this paper focuses on the findings of the verbal recall and spatial task activity.

In the verbal recall analysis the interviewer asked each subject the following question and audio taped the response for later transcription and analysis: "The title of the first story you saw was: (Title of Presentation) Suppose I didn't see or hear this
presentation. Tell me everything you can remember about this story." There were four different tasks to determine spatial recall. The task for "A Child is Taken" asked the students to imagine they were investigating the Armstrong kidnapping and needed a map of the scene of the crime which would show the route taken by the kidnappers as well as clues which may have been left. In the task for "The Strength of Memories" students were asked to imagine they were writing history books about boxers and wanted to include a map of the town where Salamo was raised. The spatial task for "The Sound of Hatred" asked subjects to imagine they were newspaper reporters and were covering the raid to the Crosstown Settlement. They needed to draw a map of the settlement for the article. In "Return to the Sea" students were asked to imagine they were movie directors and needed to draw a story board of the scene when the baby whale was saved by the villagers.

Coding Procedures

Each story was broken down by propositions (Kintsch & VanDyke, 1975) for analysis and coding. Subject recall of the narrative was coded for each proposition to reflect if the proposition was recalled and whether that recall was an exact or synonymous replication of what was stated in the text or whether it was recalled erroneously by the subject. A coding sheet was also designed to record subject recall of the spatial aspects of the presentation. This coding sheet served as a way to evaluate a drawing made by the subject in accordance with specific propositions from the presentation. For example in "The Sound of Hatred" certain items were mentioned in the narrative such as shelters, a road, a bus stop and a raw sewage ditch. The spatial task drawings by the students were evaluated to show whether the students drew the items, textually represented the items on the spatial map or neglected to represent the items in any manner for the task.

Results

Overview

The influence that format of presentation had on a subject's verbal recall of a narrative was the strongest finding in this study. Not only did the overall verbal recall of the narrative differ significantly among groups of students who saw the talking head, voice-under-text, voice-under-video or multiformat presentation condition, but also, when the variance due to the influence of the other independent variables (story, age and gender) was removed in an analysis of covariance, the variance among the different format of presentation groups continued to be significant. Furthermore, whether the narratives were analyzed separately or together, the formats that generally showed the highest mean proportion of propositions recalled were the primarily visual formats (multiformat and voice-under-video). Specifically, the proportion of propositions verbally recalled by the students differed significantly among the four formats of presentation: students remembered the most about stories shown in the more visual formats (multiformat and voice-under-video); and students remembered the least from the primarily audio formats (voice-under-text and talking head). The greatest proportion of propositions recalled was from narratives shown in a multiformat presentation and the least proportion of propositions was recalled from stories seen in the talking head format. The spatial task findings were basically inconclusive and will be discussed later.
Verbal Recall Results

The proportion of propositions recalled verbally were computed for each student according to the coding procedure described previously. Analysis of this data showed significant main effects for three (format, story and age) of the four independent variables used in the study. Specifically, the proportion of propositions verbally recalled by the students differed significantly among the presentation formats: talking head, voice-under-text, voice-under-video and multiformat \((F,(3.252)=11.29,p<.001)\). Furthermore, when the nuisance effects of the other independent variables (story, age and gender) were removed in an analysis of covariance, the proportion of propositions verbally recalled still differed significantly among the talking head, voice-under-text, voice-under-video and multiformat presentation formats \((F,(3.249)=11.32,p<.001)\). The highest mean proportion of propositions recalled was by students who saw multiformat presentations (.29) and to a lesser extent, voice-under-video presentations (.27). Students who saw the stories in the voice-under-text (.23) or talking head (.20) formats recalled fewer propositions from the narratives.

Table 2 shows the individual and overall mean and standard deviation of the proportion of propositions recalled from presentations shown as a talking head, voice-under-text, voice-under-video and multiformat.

<table>
<thead>
<tr>
<th>Format of Presentation</th>
<th>Entire Population</th>
<th>Multiformat</th>
<th>V-u-Video</th>
<th>V-u-Text</th>
<th>Talking Head</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>.2491</td>
<td>.2923</td>
<td>.2728</td>
<td>.2320</td>
<td>.1991</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>.1052</td>
<td>.1070</td>
<td>.1076</td>
<td>.1032</td>
<td>.0764</td>
</tr>
</tbody>
</table>

The second independent variable showing a strong main effect significance was story. The proportion of propositions verbally recalled by the students differed significantly among “A Child is Taken”, “Return to the Sea”, “The Sound of Hatred” and “The Strength of Memories” \((F,(3.252)=5.80,p=.001)\). In general, it was found that students recalled the most from the two positive narratives, “Return to the Sea” (.28) and “The Strength of Memories” (.27). The most predictable story, “A Child is Taken” (.23), ranked third in the proportion of propositions verbally recalled by students in the study, and “The Sound of Hatred” (.22) was recalled the least among the students. This last story most resembles a news documentary.

Spatial Task Results

The proportion of propositionally relevant items drawn on the “spatial map” were analyzed for each student. Table 3 shows the mean and standard deviation of that recall by format of presentation.
Table 3 -- Proportion of Spatial Items Recalled by Format of Presentation

<table>
<thead>
<tr>
<th>Format</th>
<th>Overall</th>
<th>Multiformat</th>
<th>V-u-Video</th>
<th>V-u-Text</th>
<th>Talking Head</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.416</td>
<td>.4272</td>
<td>.4292</td>
<td>.4206</td>
<td>.4008</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td>.1790</td>
<td>.1584</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td></td>
<td></td>
<td></td>
<td>.1729</td>
<td>.1676</td>
</tr>
</tbody>
</table>

Analysis of variance of the proportion of items recalled by students during the spatial task did not show independent variable effects for format, age or gender. However, the proportion of spatially recalled items did show a main effect story significance among “A Child is Taken”, “Return to the Sea”, “The Sound of Hatred” and “The Strength of Memories” ($F_{3,224}=21.50$, $p<.001$). Students recalled the most from “Return to the Sea” (.52) and progressively less from “The Strength of Memories” (.44), “A Child is Taken” (.39) and finally “The Sound of Hatred” (.32). This was consistent with the verbal recall findings in the study. The two-way interaction of format and story was also significant ($F_{9,209}=2.06$, $p<.05$), but none of the other interactions were significant.

Since story was the only independent variable that showed a main influence on recall for spatial task, the stories were investigated further to see if there was an individual story-format relationship. Table 4 shows the mean proportion of spatial items recalled when the stories were analyzed separately. Only “The Sound of Hatred” differed significantly in the proportion of items recalled by students who saw the voice-under-text (.40), talking head (.35), voice-under-video (.26) or multiformat (.26) presentations ($F_{3,60}=3.75$, $p<.05$).

Table 4 -- Mean Proportion of Spatial Items Recalled by Format and Story

<table>
<thead>
<tr>
<th>Story</th>
<th>Talking Head</th>
<th>V-u-Text</th>
<th>V-u-Video</th>
<th>Multiformat</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHILD</td>
<td>.35</td>
<td>.34</td>
<td>.39</td>
<td>.46</td>
</tr>
<tr>
<td>RETURN</td>
<td>.50</td>
<td>.51</td>
<td>.52</td>
<td>.55</td>
</tr>
<tr>
<td>SOUND</td>
<td>.35</td>
<td>.40</td>
<td>.26</td>
<td>.26</td>
</tr>
<tr>
<td>STRENGTH</td>
<td>.40</td>
<td>.44</td>
<td>.48</td>
<td>.44</td>
</tr>
</tbody>
</table>

Conclusions

The findings reported here suggest that format of presentation does influence the memory for a narrative. Significant differences in the proportion of propositions verbally recalled were found among stories presented as talking head, voice-under-text, voice-under-video and multiformat presentations. Students who viewed the primarily visual
formats (multiformat and voice-under-video) showed the highest recall and students who viewed the primarily audio formats (voice-under-text and talking head) recalled fewer propositions. Also, more was recalled from the more positive narratives than the stories with a relatively negative slant.

Since the verbal recall findings were so strong, it showed that there were memory differences associated with presentation formats, but the coding procedures used to evaluate the spatial tasks of the students were so rough they were unable to adequately reflect the differences in the drawings. The coder was instructed to look for certain items on the spatial task sheet and mark only whether they were represented in some manner by the student. As a result, this static evaluation lost the creativity of the drawings as well as their uniqueness. For example, as can be seen in the attached drawings, subjects 4 and 46 saw the same version of the experiment. However, each subject had a different interpretation of the tasks. Some drawings are very "busy" and the student completely covered the space provided for the spatial task with incidental scenery, people or other information about the story. Another drawing by that same student is a very verbal, textual representation the task for the story. This contrast in drawings by the same student was neither recognized nor evaluated with the current coding scheme. More sensitive coding procedures could be developed to identify unique drawing patterns within each format of presentation for comparison and analysis.

Furthermore, the fact that the task for each story was different confounded the problem of accurately interpreting a student's spatial memory for a story. For example, the most structured task was for "Return to the Sea" and included a story board for students to use to draw certain scenes from the narrative. Students were able to draw the most propositionally relevant items for this story. On the other hand, the tasks associated with the other three stories gave students about the same amount of structure to stimulate their response in the spatial drawing. It is interesting that the proportion of spatial items recalled from these stories were about the same.

Finally, a coding scheme could be developed to see if there was a positive-verbal and negative-visual relationship among the presentation formats, the stories, and the drawings. For example, did students who saw the primarily visual presentation formats (voice-under-video and multiformat) display more recall in the spatial task for the more negative stories ("The Sound of Hatred" and "A Child is Taken") than the positive stories? Similarly, did students who viewed the primarily audio presentation formats (talking head and voice-under-text) display less recall of the negative stories than the more positive stories ("Return to the Sea" and "Strength of Memories")? The coding procedure to determine this relationship would need to reflect a sensitivity to the media formats, the emotional valence of the stories, the spatial task of the drawing and individual differences in a subject's drawing ability. That information could also then be compared to the students verbal recall and forced recognition performance of this study.

Much more analysis is needed to further advance the implications of the findings in this particular portion of the study. However, if one conclusion can be drawn from the data provided in this study, it is that each format of media presentation has unique attributes which interact with, and enhance the memory for, a variety of information.
References


