The Effect of Pictures and Humor on Memory for Verbal Material in Two Extreme Scholastic Aptitude Populations.

In an effort to understand why pictures are used to supplement learning material, a study explored (1) the role of humor in recall, and (2) whether the effects of humor are independent of the individual's ability level. Subjects--students from either the high SAT (approximate average of 1200 points) population (HS), or the low SAT (approximate average of 800 points) population (LS)--were randomly assigned to one of four experimental groups that studied (1) cartoons with humorous captions, (2) cartoons with nonhumorous captions, (3) the humorous captions alone, or (4) the nonhumorous captions alone. Results indicated that humor has little or no noticeable effect, or a negative effect, on free recall for both the HS and the LS college populations. The LS participants rated the material as more humorous than did the HS population, but recalled less. Recall was lowest for the HS group that read humorous captions without pictures. For the LS group, recall was best when the nonhumorous captions were presented with pictures. Pictures relevant to the accompanying text, containing highly coherent messages, encourage superior recall for both groups. However, implications are that humor and pictures are additional stimuli competing for attention from the LS population. (Tables of findings are included.)
The Effect of Pictures and Humor
On Memory for Verbal Material
In Two Extreme
Scholastic Aptitude Populations

Sabato D. Sagaria
Capital University

Peter L. Derks
The College of William and Mary

Running head: PROCESSING OF PICTURES AND HUMOR

Presented at the American Educational Research Association
(69th, Chicago, IL, March 31 - April 4, 1985)
Pictures and Humor

Abstract

Exactly why pictures and humor are used to supplement learning material is not clear. Addressed in this research were two questions: 1) What is the role of humor in recall?; and 2) Are the effects of humor independent of the individual's ability level?

Humorous or nonhumorous captions with or without relevant pictures were presented to four experimental groups in each of two extreme SAT populations with means of 800 and 1200. Pictures enhanced recall performance across populations. It is argued that pictures provided stimulus coherence for the learner that was useful in reducing ambiguity.

The direct comparison of recall data contradicted the correlational findings between humor and recall. Humor had either no effect or a negative effect on memory. The low SAT participants rated the material more humorous than the high SAT population but recalled less. Recall was lowest for the high SAT group for humorous captions without pictures. For the low SAT group recall was best when the nonhumorous captions were presented with pictures. Implications are that humor and pictures were additional stimuli competing for attention for the low SAT population.
Pictures and Humor

The ability level of the individual, the contents of the picture and text, as well as type of humor interact to make learning material more humorous at the expense of recall, for the low SAT population. Furthermore, direct comparisons are essential to unlocking the mysteries of humor preference and benefits.
Pictures and Humor

The Effect of Pictures and Humor on Memory for Verbal Material in Two Extreme Scholastic Aptitude Population

Educators and textbook writers use both pictures and amusement to supplement reading material (Carr, 1958; Sagaria, 1980/1981). "Concrete" material has been recommended and included in text especially since Paivio (1969, 1971) reemphasized the importance of imagery as a stimulus to recall and as a general mnemonic. In addition, Zillman and Bryant (1983) observed that education has "eagerly embraced the merger with entertainment" (p. 173).

Despite anecdotal evidence and extensive research on both pictures and humor the justification for their use as aids to learning remains mixed, especially when the ability of the learner is taken into account. Consequently, the present research examined recall by two populations with very different verbal aptitudes as the presence of pictures and humor was manipulated. Specifically two questions guided this research: 1) What is the role of humor in recall?; and 2) Are the effects of humor independent of the individual's ability level?
Research on Pictures

It has been demonstrated that pictures are remembered better than words (Bevan & Steger, 1971; Derks & Dunman, 1974; Paivio & Csapo, 1973). Levin (1979) stated that "pictures in children's prose learning are positive, potent, and pervasive." In addition Levin and Lesgold's (1978) review indicated that pictures facilitated children's prose learning under most conditions.

Other researchers have found the use of pictures in text to be distracting. Samuels (1967, 1970) found that when pictures and words were presented together, the pictures were distracting stimuli and interfered with the acquisition of reading responses. Underwood (1963) and Samuels and Jeffrey (1966) also reported that pictures functioned as distracting stimuli which drew attention away from the printed words. Important to the present research is Samuels' (1967) conclusion that it was the poorer readers who were most distracted by pictures. Furthermore, Willows (1978) reported that children read words more slowly when pictures were present and that unrelated pictures produced more interference than related pictures. In agreement with Samuels, he also found distraction to be inversely related to reading ability. The present research specifically tests these findings in college populations.
Several factors influence processing of pictured information. In a meta-analysis of 16 studies, Readence and Moore (1981) found a slight but positive effect of adjunct pictures on reading comprehension. The effect was greatest for college students with traditional texts. Sagaria (1980/1981) found that the effect of adjunct aids was dependent on the amount of information overlap between sources. Information overlap refers to the ability to comprehend the pictured message independent of the verbal message, and vice versa. When information overlap was high between modes of presentation, recall of pictorial and verbal information was about the same. Study time, however, was significantly less when the information overlap between the two sources was high. The above results suggest that processing pictures and text is affected by the contents of pictures and text, student ability, and learning objective. In addition, the implications are that adjunct aids can either create or clarify ambiguity. And, knowing when each occurs allows authors to more accurately target the population best served by their publications.

Research on humor

Amusement has been endorsed as an all purpose aid to information processing. For example, humor is used by textbook authors to maintain interest and attention (Gruner, 1976), by
public speakers to emphasize a point (Jersild, 1978), by test
administrators to reduce anxiety within the individual
(McMorris, Urbach, and Connor, 1985), by educators to make
learning "fun" (Chomisky, 1979; Gilliland & Mauritsen, 1971) and
to serve as cues for recalling information (Kaplan & Pascoe,
1977). Thus, humor appear to be an effective supplement to
bring about learning. But is it?

Despite the enthusiasm for using humor in educational
material the research evidence supporting the use of humor is
often anecdotal, methodologically weak, contradictory and
relying heavily on correlational analyses. Markiewicz (1974)
reviewed 28 humor studies of which seven focused on
comprehension. Of the seven, only one study (Berlo & Kumata,
1956) resulted in beneficial effects from humor. Furthermore,
Markiewitz found that only one of 11 studies focusing on
retention of the message, produced positive effects, (Gibb,
1964). Three found negative effects, and seven demonstrated no
effects. Chapman and Crompton (1978) reported three additional
humor studies with positive results (Davies & Apter, 1980; Kauck
& Thomas, 1972; Kaplan & Pascoe, 1977). In addition, Markiewitz
also notes, "Severe methodological problems with prior research
include inadequate control messages, questionable humor
manipulations, inappropriate settings for receipt of humor,
limited subject populations, and blatant demand characteristics." (p.407) Thus support for using humor in educational materials to aid recall or comprehension appears to be tenuous.

Zillman and Bryant (1983) hypothesized a general relation between humor relevance and age. "The negative effect of the use of integrated relevant humor diminishes and vanishes with the child's advancement to improved information processing skills, this advancement making confusions in the interpretation of reality-distorting humor increasingly unlikely." (p.189). Furthermore, college students seem to be distracted by unrelated humor while related humor "is without immediate consequence for learning, but may lead to superior retention of educational information." (p.190) That is, the educational value of humor is not necessarily directly related to learning. It appears that because of methodological research problems, available research evidence fails to provide conclusive answers to some basic questions addressed by this study.

A direct way to examine the effects of humor and pictures is to compare recall of picture (cartoon) captions to recall of the captions without the picture and recall of captions that are not funny. Within this article, the use of the terms funny and not funny may be interchanged with humorous and nonhumorous,
respectively. In the present study, recall of either humorous or nonhumorous captions was compared with recall of the same captions either with or without a related picture. Usually the effect is assessed by the correlation between humor ratings and memory making causal inferences uncertain (Chapman & Crompton, 1978). Direct manipulation of humor was a departure from other research on memory for humor.

Motivation and attention is another issue related to the study of humor and memory. Humor's innocent-tendentious dichotomy as discussed by Freud (1905/1960) is based on differential motivational impact. This impact should influence attention, but how and for whom? It was hypothesized that tendentious humor (sexual, aggressive) would attract more attention and consequently be better recalled than irony and nonsense humor. To test this hypothesis the four humor types used by Eysenck and Wilson (1976), (sexual, aggressive, irony, and nonsense) as well as many of their cartoons were used.

Finally, it is not known whether recall is independent of one's cognitive abilities. Two extreme SAT populations (refer to subject section below) were sampled to assess if effects of humor and pictures were dependent upon ability. If Samuels and Jeffrey (1966), Samuels (1967; 1970), Underwood (1963) and Willows (1978) claim is correct, processing of pictured and
Pictures and Humor

humorous educational material is expected to produce variations in recall that are dependent upon the learner's ability.

Method

Subjects

The subjects for the experiment were 214 volunteers from introductory psychology classes. An additional 114 volunteers, representing the same populations, rated the funniness of the stimuli. The volunteers fulfilled a course requirement for their participation. The two extreme SAT populations will be referred to as the high SAT population (HS) and the low SAT population (LS). The HS population had an SAT average of approximately 1200 and the LS population had an SAT average of approximately 800.

Design

The 107 participants from the HS population and the 112 participants from the LS population were randomly assigned to one of the four experimental groups. These groups either studied: 1) cartoons and their humorous captions; 2) cartoons with nonhumorous captions; 3) the humorous captions alone; or 4) the nonhumorous captions alone. Not an issue in this study is the sex differences. Sex differences in humor have been diminishing over the years and are not related to gender alone (Brodzinsky, Barnet, & Aiello, 1981; McGhee, 1979).
Additionally, 41 volunteers from the HS population and 73 volunteers from the LS population were randomly assigned to rate the humor of the four sets of stimuli. Each rater used a 21 point scale with "0" meaning that the item was "not funny at all" and a rating of "20" indicated that the item was "extremely funny", (Derks, Lewis & White, 1981).

**Stimuli**

Some of Eysenck and Wilson's (1976) sexual cartoons were not approved by the ethics committees and were supplemented by milder sexual cartoons from Colell and Domino (1980). The final set of stimuli represented the four humor types (Eysenck & Wilson, 1976) and included eight each of nonsense, irony, and aggression, and twelve sexual cartoons. For analysis purposes the sexual recall score for each subject was multiplied by 2/3 yielding an adjusted maximum of eight possible correct recall score for each humor type. Consequently the contribution of sex was slightly underestimated for the sake of statistical balance. All cartoons were redrawn by a single artist to eliminate clues that might be idiosyncratic to particular pictures.

Four sets of booklets were constructed, one for each experimental condition. Set one presented the pictures and their humorous captions. In set two the captions were rewritten so that they were not humorous. The humorous and nonhumorous
Pictures and Humor

captions only were placed in booklet sets three and four. To control for order effect twelve different orders of the stimuli were used in each booklet. It should be noted that the nonsense, irony, aggression and sexuality of the cartoons was considerably altered in the various nonhumorous versions of the stimuli implying that some direct comparisons of these categories have less meaning. For example the stimulus may be significantly altered when the accompanying caption is a simple statement of the illustration. The individual may question why a picture is needed for such an obvious statement.

Procedure

Subjects were randomly placed in one of the four experimental treatments and the appropriate booklets were distributed. They studied the material in their booklets at the rate of one page every 15 seconds. After studying all learning material they were asked to recall, as accurately as possible, the caption from each page. One half hour was allowed for the recall phase of the experiment and participants could not leave until the thirty minutes had expired. To minimize rehearsal and assure long term memory recall, a five minute filler task was introduced, without prior warning, between the study and recall tasks. The filler task also allowed the experimenter time to collect booklets and thus remove any chance of participants
using them during recall. For this filler task students were asked to recall and record the first time they remembered laughing, a funny real life experience, and a practical joke they had participated in.

**Results**

**Humor ratings**

The pictures were rated funnier than the same pictures with nonhumorous captions or captions alone (refer to Table 1).

The HS individual’s average rating was 7.65; for the same pictures with the nonhumorous captions, 3.95; for the "funny" captions, 3.25; and for the nonhumorous captions, 1.61. The ratings by the LS participants were pictures, 9.15; for the same pictures with nonhumorous captions, 6.96; for the "funny" captions alone, 3.48; and for the nonhumorous captions alone, 3.51. In spite of the slightly higher ratings by the LS group, the correlations between HS and LS ratings by cartoons were: for pictures, .69; for not funny pictures, .66; and funny and not funny captions, .83 and .54 respectively. All of the above correlations were significant at \( p < .001 \) with 34 df (based on
Pictures and Humor

cartoons rather than raters). Thus, the two populations basically agreed on which items were funny. Some reversals occurred when a particular picture with a nonhumorous caption was rated funnier than the picture and its humorous caption.

Recall

Also presented in Table 1 is the percent of captions recalled for each type of humor. Since there were some reversals in the ratings of humor, two separate analyses were conducted. The first analysis included the reversals, while the second omitted two stimuli from nonsense, irony, aggression, and four from sexual type humor items. The reversed items had higher humor ratings for pictures with the not funny captions from the LS group. Only one item was given reversed ratings by the HS group. This item was also reversed by the LS group. The analysis of variance conducted with the adjusted data again yielded nonsignificant results, at the 0.05 alpha level, for humor. Since the results were similar between the analysis that contained the reversals and the analysis that deleted the reversals, further discussion is based on the data that included reversals, unless otherwise specified.

Although the LS group rated the material more humorous, they recalled less of it, $F(1,206) = 62.44, p < .001$. Sexual material was recalled best, even when it was presented as a
Pictures and Humor

relatively mundane description or comment, $F(3, 618) = 75.95, p < .001$. The four types of humor interacted with the presence or absence of pictures, $F(11.07), p < .001$, whether it was humorous or nonhumorous, $F(3, 618) = 5.81, p < .0001$, both pictures and humor, $F(3, 618) = 2.88, p < .035$, and almost with population, $F(3, 618) = 2.51, p < .058$. This pattern of interactions is not surprising, for as previously noted, the nature of the nonsense, irony, and aggression stimuli changed when the caption was not accompanied by a picture or a picture was not accompanied by a humorous caption.

Illustrated in Figure 1 is the significant triple interaction between pictures, humor and SAT populations, $F(1, 206) = 4.45, p < .036$.

No significant two way interactions resulted between SAT level and humor, $F(1, 206) = 0.10, p < .75$, nor with pictures, $F(1, 206) = 3.44, p < .065$.

The presence of pictures improved recall, $F(1, 206) = 32.81, p < .0001$, for both populations. Humor, on the other hand, was not significantly effective, $F(1, 206) = 3.63, p < .058$. In fact, the nonhumorous versions of the stimuli were recalled
better. Recall improved for the HS group when both humor and pictures were absent from the learning material. The LS group produced similar recall improvements when pictures were present but humor was absent. This evidence supports the hypothesis that humor interfered with recall in the LS population when accompanied by pictures, and in the HS population when pictures were absent. Recall by the LS group was equally low for the funny and the not funny experimental conditions when pictures were absent. Recall by the HS group was equally high regardless of humor when they studied material with pictures.

A final statistic of interest is the correlation between humor rating and recall of stimulus items. These correlations were positive. For the HS group the correlations were:
cartoons .12 (p < .10), not funny cartoons .43 (p < .01),
captions .30 (p < .10), and not funny captions .38 (p < .05).
The correlations for the LS group were: cartoons .29 (p < .10),
not funny cartoons .35 (p < .001), captions .45 (p < .10), and
not funny captions .22 (p < .10), all dfs were, 34. The correlations between humor and recall were highest for material with intermediate humor ratings.

Discussion

It is hypothesized that pictures and a sexual humor theme improves memory for text by increasing the learner's interest
and focusing attention to the stimuli. Interpreting the correlational results, we are led to conclude that aggression and sexual humor, as well as imagery contribute to making the stimuli more humorous and memorable. The correlational findings from this research are consistent with prior conclusions that humor is associated with memory processes (i.e., recall).

An important finding of the present research is that humor does not contribute to memory when direct comparisons are made between humorous and nonhumorous stimuli. Specifically, when the LS population viewed pictures and the HS population viewed verbal material, with humor, memory was reduced. The implications behind these findings are that a direct manipulation of humor is essential to understanding humor's effect on recall.

The superior performance of individuals studying pictures can be interpreted in relation to stimulus "coherence" and reduction of potential ambiguity (Kintch & Vipond, 1976). Stimulus coherence refers to how well each aspect of the message fits together to present a consistent thought. For example, high coherence exists when an author combines humor, pictures and text to convey one logically connected thought clearly and the learner, in turn, interprets the intended thought. Low coherence refers to a situation where any number of
interpretations of the author's intended message are possible. The latter increases ambiguity while the former reduces it. In the present study, the picture and its complementary caption provided a more coherent message than the caption alone. Humor played a nonsignificant role for the HS learner. As inferred from the high preference for pictures (Sagaria 1980/1981; Sagaria, Bass, & Reis, 1982) the picture is likely to be attended to before the text. If any ambiguity in the picture exists, it is reduced if and when attention is directed to the secondary source, the text. When the picture is absent, resolution of ambiguity is not possible, especially for "funny" material. Therefore the picture may be an essential complement to the caption, for when the picture is absent the HS individual cannot reduce the ambiguity in the caption, thereby rendering the material less meaningful. The end product is rote learning. Similar processing activities may be inferred for the LS group. The presence of humor adds to stimulus complexity and overall ambiguity thereby diffusing attention. Due to the fact that each stimulus component receives less overall attention and the LS individual may already have organizational difficulties, the message conveyed by the text becomes incongruous in relation to the picture. The result is high ambiguity, which in turn impedes storage and retrieval. Thus when the LS learner studies
text material that contains humor and pictures the task is more ambiguous, less coherent and/or more distracting.

Better recall is realized, by the LS individual, when a picture accompanies the text and humor is absent. The reason being that simplifying the stimulus, by the removal of humor, reduces ambiguity. The learning material becomes more coherent when a source of distraction and ambiguity is removed. When no picture was used, both the punch line and the related caption were equally meaningless and poorly remembered. The HS population seemed to be more skillful at interpreting and relating information. Relating pictures and the somewhat incongruous (i.e. funny) text was especially difficult for the LS population.

Learning material containing pictures and humor with a sexual or aggression theme improved the LS recall for not funny captions with pictures to within a few percent of HS recall for the same material. Furthermore, sex and aggression appear to attract more attention than nonsense or ironic humor and are considered to be more interesting, to the LS learner. The implications are that attention to "interesting" stimuli increases recall for the LS individual more than for the HS individual.
The present researchers are led to hypothesize that an incidental/intentional learning paradigm might help explain the discrepancy between correlational and comparative findings for humor and should be researched further. Such a paradigm may explain the positive effect due to "interest" engendered by humor (Hauck & Thomas, 1972; Zillman & Bryant, 1983). The argument is that focusing attention on particular aspects of the stimuli through the use of humor will make information that requires attention more interesting. Logically, it follows that more interest leads to more attention making the material more coherent and thus less ambiguous. The expectation is that recall will improve.

Pictures contributed to memory for captions in both populations. The superior recall is attributed to the learner's ability to develop a verbal and non-verbal memory code (an issue eloquently discussed by Paivio), and to the complementary nature of the picture and caption to reduce ambiguity. Consequently, storage of highly coherent, better organized, and less ambiguous information makes retrieval easier, at least for free recall.

Conclusions and Implications

The effects of pictures and humor on recall of printed material depends upon the content of the picture and text, the type of humor used and the ability of the learner. If the
picture is relevant to the accompanying text, and conveys a highly coherent message, it reduces ambiguity. Under these conditions superior recall occurs for both high and low ability college students.

Humor, on the other hand, has either little or no noticeable effect or a negative effect on free recall for high and low SAT college populations. For the HS population non-illustrated humor increases ambiguity and impedes recall. For the LS population, humor accompanied by a picture increases complexity and thereby disperses attentive behaviors.

The findings justify the conclusion that direct manipulation of humor provides essential data unattainable from correlational methods. Results derived from direct comparisons between recall from humorous and nonhumorous material contradict the positive correlational results between humor and recall. Embellishment of learning material with humor may be preferred, but the educational benefits on recall for the populations used in this research are negligible, at best. The challenge, therefore, is to use pictures in combination with humor to increase interest, focus attention, yet not distract the learner from intended objectives.
References


AUTHOR'S NOTES

The project reported herein was performed pursuant to a grant from the U. S. Army Research Institute for Behavioral Sciences to the first author with additional resources being provided by the College of William and Mary.

Thanks to Jana Herrman and Benton Bailey for their aid at various stages of this research. Thanks to Jim Williams who redrew the cartoons.

Requests for reprints should be sent to Peter L. Derks, Department of Psychology, College of William and Mary, Williamsburg, VA 23185.
Table 1
Recall and Humor Ratings of students from extreme S.A.T. populations

<table>
<thead>
<tr>
<th>PICTURE</th>
<th>% Recall</th>
<th>Rating</th>
<th>% Recall</th>
<th>Rating</th>
<th>% Recall</th>
<th>Rating</th>
<th>% Recall</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonsense</td>
<td>53</td>
<td>5.72</td>
<td>51</td>
<td>2.71</td>
<td>33</td>
<td>6.49</td>
<td>39</td>
<td>4.84</td>
</tr>
<tr>
<td>Irony</td>
<td>32</td>
<td>5.28</td>
<td>37</td>
<td>2.68</td>
<td>24</td>
<td>5.71</td>
<td>31</td>
<td>4.96</td>
</tr>
<tr>
<td>Aggression</td>
<td>49</td>
<td>7.87</td>
<td>44</td>
<td>3.69</td>
<td>34</td>
<td>10.83</td>
<td>41</td>
<td>8.14</td>
</tr>
<tr>
<td>Sex</td>
<td>61</td>
<td>10.38</td>
<td>60</td>
<td>5.80</td>
<td>53</td>
<td>11.43</td>
<td>58</td>
<td>8.90</td>
</tr>
<tr>
<td>Overall</td>
<td>50.3</td>
<td>7.65</td>
<td>49.2</td>
<td>3.95</td>
<td>37.8</td>
<td>9.15</td>
<td>43.9</td>
<td>6.96</td>
</tr>
</tbody>
</table>

NO PICTURE

<table>
<thead>
<tr>
<th>% Recall</th>
<th>Rating</th>
<th>% Recall</th>
<th>Rating</th>
<th>% Recall</th>
<th>Rating</th>
<th>% Recall</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonsense</td>
<td>50</td>
<td>2.29</td>
<td>46</td>
<td>1.27</td>
<td>34</td>
<td>2.61</td>
<td>27</td>
</tr>
<tr>
<td>Irony</td>
<td>27</td>
<td>3.06</td>
<td>47</td>
<td>1.06</td>
<td>18</td>
<td>2.29</td>
<td>29</td>
</tr>
<tr>
<td>Aggression</td>
<td>32</td>
<td>2.10</td>
<td>34</td>
<td>1.60</td>
<td>15</td>
<td>2.34</td>
<td>19</td>
</tr>
<tr>
<td>Sex</td>
<td>49</td>
<td>4.73</td>
<td>55</td>
<td>2.20</td>
<td>42</td>
<td>5.83</td>
<td>36</td>
</tr>
<tr>
<td>Overall</td>
<td>40.3</td>
<td>3.24</td>
<td>46.4</td>
<td>1.61</td>
<td>28.6</td>
<td>3.56</td>
<td>28.3</td>
</tr>
</tbody>
</table>

Note: Maximum rating = 20; minimum rating = 0

Overall % recall based on 36 cartoons; 12 for sex, 8 for other types
Figure Caption

Figure 1. Recall as a function of humor, pictures, and SAT population.
The graph illustrates the mean percent of captions recalled for different conditions:

- HS Pictures
- HS No Pictures
- LS Pictures
- LS No Pictures

The x-axis represents humor (Funny vs. Not Funny), and the y-axis represents the mean percent of captions recalled. The graph shows a general trend where the recall of captions increases with humor, regardless of whether pictures are present or not.