Visual and semantic organization in picture recall

ELIZABETH LOTZ STINE
Brandeis University, Waltham, Massachusetts

and

ANGELINE E. BENHAM and ANDERSON D. SMITH
Georgia Institute of Technology, Atlanta, Georgia

Frost (1971, 1972) demonstrated that subjects can organize pictures by using either visual features or semantic features of the items presented. For the present study, a set of pictures was constructed that could be organized visually and semantically. The subjects viewed these pictures at either a fast rate (1.5 sec/item) or a slow rate (5 sec/item) of presentation. The subjects expected a recognition task, which would enhance attention to visual detail. Clustering by visual (spatial orientation) and semantic (taxonomic category) categories on an unexpected recall task was measured. At the fast presentation rate, semantic and visual clustering occurred at equal levels, but when presentation rate was decreased (longer presentation times), visual clustering decreased and semantic clustering increased. These results are discussed in terms of the changing representation of picture information as a function of study time.

The superior memory for pictures over words is a well-replicated finding, implying that visual features can play a role in the encoding of information (Winograd, Smith, & Simon, 1982). One method for assessing the relative value of visual and semantic features at encoding is to determine the extent to which semantic and visual features are used in organization at recall. The study described here was an investigation into the use of verbal and visual features as a function of encoding time during presentation. Using a 1.5-sec presentation rate, Frost demonstrated that subjects expecting a recognition task will cluster in recall using visual attributes (Frost, 1971), and that when visual and semantic features are available, subjects will use both types of features in clustering (Frost, 1972). Recently, Smith (1983) had difficulty in replicating Frost's findings, but used presentation rates somewhat slower than those used by Frost.

It is possible that slower presentation rates promote semantic clustering and that, because of the advantage of semantic information for adequate retrieval, semantic organization is the preferred organizational strategy. In other words, at different presentation rates, different types of information are better retained.

At longer presentation rates, there is some evidence that some kinds of picture information are better retained than others. Mandler and Johnson (1976) presented pictures of organized or disorganized collections of objects at 5, 20, or 60 sec. Distractors in the recognition test were varied with respect to kind of transformation. In organized scenes, relative spatial location information was retained as well as object type at all presentation rates. Subjects more often chose distractors containing deletions, changes in object detail, or changes in object position that did not seriously alter relative spatial location. These results suggest that visual spatial codes are established earlier than visual pattern codes or figure-ground relationships. Subjects also became more accurate in rejecting distractors with more study time, and this improvement varied with the type of distractor and organization. The Mandler and Johnson study clearly demonstrates that all kinds of picture information are not processed at the same rate.

METHOD

Subjects

Thirty-eight students from introductory psychology classes at Georgia Institute of Technology volunteered to participate in the experiment for class credit. The subjects were divided into two groups, one of which viewed pictures at a fast presentation rate ($n = 20$) and one of which viewed the pictures at a slow presentation rate ($n = 18$).

Stimulus Materials

The stimulus pictures were simple white-line drawings on a black background and were projected onto a screen using a Kodak carousel projector. The pictures were of 32 objects from four semantic categories (articles of clothing, vehicles, animals, and pieces of furniture). Each picture was drawn in one of four orientations (vertical, horizontal, left-slant, or right-slant). There were also four items, all slanted to the right, that were used for a practice trial. In order to reduce confounding of semantic and visual categories, the categories were partially overlapped as described in Frost (1972). Some of the pictures were from Frost's original collection; others were added to increase list length. Target pictures were

randomly ordered with the restriction that successive items share neither semantic category nor visual orientation. All subjects viewed the pictures in this single random order.

Procedure

In order to maximize the use of visual clustering (Frost, 1972), subjects were first shown a short series of slides and were told that the task was a recognition task. They viewed four practice slides at either a fast rate (1.5 sec/item) or a slow rate (5 sec/item). They were then shown a transparency of target pictures and an equal number of distractor pictures that were lettered A, B, C, and so on. Subjects were asked to write down the letters corresponding to the items they had seen before.

Following this practice trial, subjects were told that they would now see a similar, but longer, series of pictures and that afterward they would be asked to pick out the pictures they had seen as they had done before in the practice trial. The target slides were then presented at the same rate as the practice slides. After the slide presentation, all subjects spent 3 min performing a distractor task of crossing out numbers divisible by three. Subjects were then given 3 min for an unexpected recall task; they were asked to write down the names of all the pictures they had seen in the target slides. Finally, the experimenter explained the rationale of the experiment and the necessity of the deception involved.

Clustering by semantic and visual categories was measured using the difference score (D) described by Dalrymple-Alford (1970). This clustering measure was used because it takes into account the amount of clustering that would be expected to occur by chance and controls for the total number of words recalled by different subjects. In the D measure, a score of zero indicates clustering at chance level, and positive numbers less than one indicate the presence of clustering. Total recall was also measured. Intrusions were not counted in computing total recall. When intrusions occurred between two items that would normally be considered repetitions (used in computing the D score), the interrupted repetition was not counted. Occasionally, subjects wrote down the same word twice in the course of recall; the second occurrence was treated as an intrusion.

The effects of clustering were analyzed in a 2 (presentation rate: fast, slow) × 2 (clustering type: semantic, visual) analysis of variance. Presentation rate was varied between subjects; clustering type, within subjects.

RESULTS

Subjects who saw the pictures at a 1.5-sec rate recalled the names of fewer pictures (mean = 9.85) than those who saw them at a 5-sec rate (mean = 14.39) $[F(1,36) = 11.50, p < .002]$.

The grand mean of the D scores was significantly different from zero $[F(1,36) = 119.40, p < .001]$, indicating that clustering did occur in the recall protocols. Average scores and their standard deviations for the four conditions are found in Table 1.

There was significantly more semantic clustering than visual clustering $[F(1,36) = 13.67, p < .001]$, but this was modified by a presentation rate × clustering type interaction $[F(1,36) = 12.53, p < .001]$. At the fast

presentation rate, semantic and visual clustering were at about equal levels. With a slower presentation rate, however, semantic clustering increased and visual clustering decreased.

Results of a t test to determine whether group means were significantly different from zero are also reported in Table 1. At both fast and slow presentation rates, there were significant amounts of both semantic and visual clustering.

DISCUSSION

The use of semantic and visual information as organizational cues in picture memory has been found to vary as a function of study time. In the present study, when little time was allotted for picture encoding, items were clustered by both visual and semantic features. With more study time there was more semantic clustering than visual clustering, and the use of visual characteristics in organization at recall was reduced to very low levels. Our finding that visual coding was essential to efficient encoding in the early stages of picture processing is consistent with Nelson, Reed, and Walling (1976), who found that at short presentation rates (1.1 sec and 2.2 sec), reducing distinctiveness of visual attributes reduced or even reversed the advantage of pictures over words (picture superiority effect).

These results suggest that early in the process of studying pictures for later recognition, visual characteristics play a more important role than they do later in this process. With more study time, semantic features of the pictures become more important and visual information becomes less important. As semantic characteristics are used more in encoding, superficial visual orientation information is used less.

A similar kind of processing has been proposed for letter perception by McClelland and Rumelhart (1981). In their interactive activation model, it is assumed that "perceptual processing takes place within a system in which there are several levels of processing, each concerned with forming a representation of the input at a different level of abstraction" (p. 377). It is also assumed that visual perception involves parallel processing in a spatial sense and parallel processing between levels, and that visual processing involves bottom-up and top-down processing between those levels.

Applied to picture processing in the present study, levels of abstraction would consist of (1) extraction of visual features, (2) object identification, and (3) categorization. At fast presentation rates, there may not have been enough time to process the stimuli to higher levels of abstraction. The recall task ensured that items considered were at least identified, but with minimal higher order semantic processing, certainly not the level of processing required of interitem organizational encoding (Winograd & Smith, 1978). At the fast presentation rate, the similarity of lower order visual features was more salient; hence, visual clustering was more probable. Attention to these visual characteristics was probably enhanced by the expected recognition requirement (Frost, 1972), although this is only speculation because it was not manipulated. With longer presentation times, however, processing of the higher order similarity, the taxonomic categories, was more likely. This processing led to semantic clustering of the items. As higher levels were activated, lower levels were inhibited, making visual clustering less likely.

These data also suggest that Frost's (1971) initial explanation of the phenomenon of visual clustering may not have been complete. She demonstrated that the expectation of a recognition test was an important factor in making visual codes salient for item accessibility. The presence of visual clustering, however, is also contingent on a very fast presentation rate that minimizes access to semantic codes, at least those used in organizational encoding.

Table 1
Means and Standard Deviations for Visual and Semantic Clustering (D) as a Function of Presentation Rate

	Fast (1.5 sec/item)			Slow (5.0 sec/item)		
	Mean	SD	t	Mean	SD	t
Semantic	.257	.157	5.00*	.419	.290	7.74*
Visual	.251	.177	7.83*	.239	.094	4.08*

Note—The t test was performed to determine whether the means were greater than zero. *p < .001.

REFERENCES

DALRYMPLE-ALFORD, E. C. (1970). Measurement of clustering in free recall. *Psychological Bulletin*, **70**, 32-34.

FROST, N. (1971). Clustering by visual shape in the free recall of visual stimuli. *Journal of Experimental Psychology*, **88**, 409-413.

FROST, N. (1972). Encoding and retrieval in visual memory tasks. *Journal of Experimental Psychology*, **9**, 317-326.

MANDLER, J. M., & JOHNSON, N. S. (1976). Some of the thousand words a picture is worth. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 529-540.

McCLELLAND, J. L., & RUMELHART, D. E. (1981). An interactive model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, **88**, 375-407.

NELSON, D., REED, V., & WALLING, J. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 523-528.

SMITH, A. D. (1983). *Interaction between aging and memory* (Tech. Rep. No. 3). Atlanta: Georgia Institute of Technology.

WINOGRAD, E., & SMITH, A. D. (1978). When do semantic orienting tasks hinder recall? *Bulletin of the Psychonomic Society*, **11**, 165-167.

WINOGRAD, E., SMITH, A. D., & SIMON, E. (1982). Age-related differences in the picture-superiority effect. *Journal of Gerontology*, **37**, 70-75.