

Recognition memory for pictures: Dynamic vs. static stimuli

ALVIN G. GOLDSTEIN, JUNE E. CHANCE, MARGO HOISINGTON,
and KEITH BUESCHER

University of Missouri, Columbia, Missouri 65211

In prior research, pictorial recognition memory typically has been measured by having subjects encode static visual stimuli, such as photographs and projected photoslides. An experiment explored short-term recognition memory for the content of a 10-min segment of film. In one testing condition, 8-sec film segments derived from the study film segment and from other sections of the same film were used as targets and foils. In another testing condition, still pictures ("freeze frames") derived from the film segment and from other portions of the film were presented as targets and foils. Control subjects saw still pictures and were tested on still targets and foils, all derived from the films. Pictorial recognition memory for dynamically presented material was significantly better than memory for the same information presented statically.

Much recent research on human memory has demonstrated a vast capacity to encode and store complex pictorial information in a form that permits later recognition memory performance to reach levels of excellence approximating perfection, even when hundreds or thousands of pictures are encoded (e.g., Nickerson, 1965; Shepard, 1967; Standing, Conezio, & Haber, 1970).

The human's formidable pictorial memory capacity raises an interesting question. Why are we so proficient in processing static information when almost all of our real-world visual information is dynamic? Our normal visual world is almost never static. For instance, if the objects in the visual world are not themselves moving (which they often are) against a stationary background, then the "background" itself is moving by virtue of our natural visual scanning strategies (e.g., slow eye movements), or our own body movement (e.g., walking), or because we are moving in a vehicle. This line of reasoning led us to ask, could recognition memory for dynamic events surpass our performance with static stimuli? Curiously, as far as we have been able to ascertain, studies of picture recognition memory invariably have used static stimuli as the material to be remembered. That is, still photographs, photoslides, or similar pictorial material served as stimuli in these studies.

In this preliminary exploration, we compared recognition memory for nearly identical stimulus information in either a dynamic or static mode.

Portions of this research were presented at the Midwestern Psychological Association meeting, Minneapolis, Minnesota, May 1982. The authors want to thank Carrie Elliott for her dedication to the onerous job of transcribing hundreds of film segments from the master videotapes to the videotapes used in the experiment.

METHOD

Subjects

One hundred and thirteen (61 men and 52 women) college students served as subjects in return for bonus credit in an introductory psychology course.

Stimuli and Apparatus

Two master films served as sources of dynamic and static visual stimuli (*Ivanhoe*, about 70 min in length, released by Metro-Goldwyn-Mayer about 1953, featuring Robert Taylor and Elizabeth Taylor; *Secrets of Sleep*, about 50 min long, a didactic film depicting various aspects of sleep research, published by Time-Life Multimedia). Sections of both master films used in the experiment were transcribed onto .5-in. open-reel videotape using a Sony (Model AV 3600) reel-to-reel videotape recorder (VTR). The Sony VTR, in conjunction with a Sanyo black-and-white television monitor (13-in. screen), was used to present the visual stimuli to the subjects. "Freeze frames" (stopped video action) were transcribed from master videotapes onto working videotapes by first halting the master tape with the pause control of the playback VTR and then electronically recording the resultant still picture on the Sony reel-to-reel VTR. The duration of the freeze frame, which determined the length of time subjects would subsequently be exposed to each still picture, was controlled by hand; a stopwatch was used to measure the elapsed time interval. With the present electronic equipment, freeze frames had to be adjusted to lessen the effects of electronic interference (e.g., an irregular line across the video screen). In all cases, even though some residual distortion was apparent, either at the top or bottom edge of the video display, the content of the picture was clearly visible. The series of 8-sec videotape excerpts, used in the test phase of two experimental conditions, was electronically recorded from the master tapes in a straightforward manner. All stimuli were recorded without an audio signal.

Procedure

Picture recognition memory performance was measured in a standard recognition memory paradigm. Visual stimuli were presented during the study phase of the experiment. The test phase followed a few minutes after the last target of the study phase was presented. The test series consisted of the targets, now randomly distributed among a large number of similar-

looking distractor stimuli. The subjects' task was to decide, for each stimulus in the test series, whether it was an old stimulus (one seen before in the study session) or a new stimulus (one not seen earlier in the session).

As diagrammed in Table 1, there were three different treatment conditions, each replicated once with two films. Subjects were assigned randomly to one of the six conditions. In the dynamic-dynamic (d-d) condition, subjects studied a 10-min videotape clip from either *Secrets of Sleep* (Sd-d) or *Ivanhoe* (Id-d). A recognition memory test, consisting of 80 8-sec (approximate) videotape clips separated by 5-sec blank intervals, immediately followed the study film. Twenty of these test segments were excerpts from the 10-min study film (targets), and 60 were randomly selected from other parts of the master tape (distractors).

Subjects in the static-static (s-s) condition studied 60 video freeze frames for approximately 5 sec each. These target stimuli were systematically culled from the entire 10-min study film clip used in the d-d condition. Recognition tests for subjects in the s-s condition consisted of 120 freeze frames, 30 targets and 90 foils. The foils were obtained from approximately the middle of the 8-sec segments used as foils in the d-d condition. The s-s condition in this study is comparable in all respects to the typical picture recognition memory experiment in which static photoslides are used as stimuli.

Subjects in the dynamic-static (d-s) condition studied the 10-min film clip used in Condition d-d and were tested with 120 freeze frames of Condition s-s. In other words, the subjects in this group, who viewed a moving picture in the study session and a series of still pictures in the test session, had to decide whether a still picture had been "incorporated" in a moving picture. In all conditions, the order of appearance of both the video segments and the freeze frames was identical to the order in which they appeared in the 10-min film.

In all treatment conditions, subjects were instructed to pay attention to and try to remember the contents of the study film or freeze frames. Thus, subjects were informed about a subsequent memory test but the exact form of the test was never specified. During the test session, subjects responded to each film segment or freeze frame by marking their answer sheets in the appropriate space under columns headed "seen before" or "not seen before." To assist subjects in keeping pace with the test trials and thereby reduce errors of scoring, a tape-recorded voice, synchronized with the appearance of each test item,

announced the number of every item in the series. One-minute rest intervals were systematically introduced after Items 30 and 60 in Condition d-d, and also following Test Item 90 in Conditions s-s and d-s.

RESULTS AND DISCUSSION

The results of the experiment, summarized in Table 1, were analyzed in a 2 by 3 ANOVA: The factors were films (*Secrets of Sleep* and *Ivanhoe*) and study-test conditions (d-d, s-s, and d-s). Separate ANOVAs were run for the three performance measures: hits, false alarms, and d' . Only the results of the d' analysis will be presented in this report.

Film Analysis

Accuracy of recognition memory for the two films did not differ significantly [$F(1,106) = 1.53, p > .05$]. Also, the interaction between film and experimental condition was not significant [$F(2,106) = 2.51, p > .05$]. Although analysis of d' data revealed no difference in recognition memory for the two films, the additional data contained in Table 1 concerning patterns of hits and false alarms suggest that film content did influence subjects' performance. Taken together, the results indicate that in future research on recognition memory for dynamic stimuli, appropriate sampling of various perceptually different films should be included in the design of the study.

Static vs. Dynamic Stimuli

The primary aim of this study was to compare recognition memory for stimuli presented statically with memory for equivalent stimuli presented dynamically. The d' ANOVA revealed a significant main effect of experimental conditions [$F(2,106) = 9.61, p < .01$]. A Duncan's multiple-range test indicated that mean performance accuracy of subjects in the d-d condition ($d' = 2.61$) was significantly superior to the performance of subjects in the s-s condition ($d' = 2.22$), and also superior to that of subjects in the s-d condition ($d' = 2.08$). The interaction between film and experimental condition was not significant. This evidence suggests that memory for dynamic visual material is slightly but reliably more efficient than memory for static material. However, considering the nature of the task d-d subjects faced, their performance deserves additional comment and further research.

Recall that after d-d subjects viewed a 10-min film segment, they were shown a series of 8-sec film clips, some that were taken from the 10-min study film and some that were not in the study film but came from other portions of the master film. By design, the distractor film clips resembled the target film segment in many aspects. For example, Elizabeth Taylor was present in the study film and also in some of the distractor segments. Intuitively, the task confronting the d-d subjects seems more difficult than the task given to

Table 1
Experimental Conditions and Three Measures of Recognition Memory for Visual Stimuli Dynamically and Statically Presented

Study	Test	N	% Hits	% FAs	d'
<i>Secrets of Sleep</i> (S)					
Dynamic	Dynamic (Sd-d)	21	92	19	2.53
Static	Static (Ss-s)	20	87	15	2.39
Dynamic	Static (Sd-s)	23	89	19	2.17
Mean			89	18	2.36
<i>Ivanhoe</i> (I)					
Dynamic	Dynamic (Id-d)	15	79	4	2.74
Static	Static (Is-s)	15	78	14	1.99
Dynamic	Static (Id-s)	19	80	16	1.96
Mean			79	12	2.19
Films Combined					
Dynamic	Dynamic	36	87	13	2.61
Static	Static	35	83	14	2.22
Dynamic	Static	42	85	18	2.08

the s-s subjects, yet the results cited above do not support this subjective judgment. It would seem that d-d subjects had to first encode the visual input contained in 10 min of continuous, changing action and later compare the memory for the activity contained in the study film to the action contained in the 8-sec film excerpts. In contrast, s-s subjects needed to encode much less visual information during study, inasmuch as static stimuli represent single, almost infinitely brief, slices of the ongoing sequence. On a priori grounds, the recognition memory test for s-s subjects should be easier than that for d-d subjects, presumably because static pictures contain less information than do the continuous visual stimuli used in this experiment. If that is not enough reason to expect that d-d subjects were at a disadvantage compared with s-s subjects, then there is the additional interference that d-d subjects presumably experienced as a function of the actual movement within the study and test stimuli. How does one "capture" action in memory?

It could be argued that, strictly speaking, subjects in the s-s and d-d conditions might not have received exactly equal treatment. Thus, the tasks for the d-d subjects might have differed from the task for the s-s subjects not only by virtue of the experimental manipulation (moving stimuli compared with static stimuli) but also as a function of other confounded variables. For example, one task could have required many more discrete encodings than the other task, a situation comparable to differences in list length in a verbal learning experiment. Inasmuch as we wanted to contrast memory performance for continuously moving stimuli with memory for static stimuli, the study tasks (and the test tasks) necessarily differed. How could we "equate" the movement study task with the static study task? On what variable should they be "equated?" There are a variety of reasons for the ambiguity. For instance, any decision regarding the length of the dynamic study film clearly affects the number and length of the dynamic test segments that can be culled from the study film. A long study film segment would afford many relatively long test segments, and a similar but opposite statement can be made about a short study film. In other words, the decision to use a 10-min study film placed constraints on the number of trials, length of film segments, and so on, of other study and test conditions within the experiment. With these considerations in mind, we decided for all three conditions to roughly equate the total time subjects spent in both the study and test conditions. Thus d-d subjects viewed a 10-min study film, followed by approximately 20-min in the testing session (80 8-sec film segments with about 5-sec intervals between segments), for a total of 30 min. Subjects in the s-s condition completed their study task in about 10 min (60 freeze frames for 5 sec each plus 5-sec intervals between stimuli) and their test trial in 20 min (120 5-sec freeze frames plus 5-sec interstimulus intervals), for a total of 30 min. The d-s subjects completed their

study in 10 min and their test in 20 min. Perhaps most important, for all groups, the ratio of target to foils was equated at 1:3. In contrast, control of the content of the stimuli seen by subjects in d-d and s-s conditions was less troublesome. As explained earlier, static stimuli were obtained by systematically sampling freeze frames from the 10-min study film and from the 8-sec film segments, a method that insured that the pictorial stimuli seen by subjects in the dynamic and static conditions were culled from the same original universe of stimuli.

Clearly, the question whether all variables in the study session other than the dynamic-static variable were controlled still remains open at the end of this preliminary investigation. But, from the foregoing analysis, it appears reasonable to assume that d-d subjects did not have an easier task than s-s subjects had. If this assumption is shown by subsequent research to be valid, an explanation for the obtained performance difference that does not implicate task inequality will have to be sought.

Analysis of d-s Condition

Perhaps some of the most interesting data were obtained from subjects in the d-s condition, the only mixed-mode condition in the study. Recall that after viewing a 10-min film segment, subjects were tested with statically presented targets and foils. Because the encoded stimuli were so different from the test stimuli in the d-s task, performance should have been appreciably worse than in the other two conditions. Although subjects in the d-s condition performed relatively poorly (and the main reason for their reduced performance can be traced to the higher frequency of false alarms), a d' difference of only .70 separated the worst performance (Id-s) from the best performance (Id-d), a fact that can be interpreted to mean that even when subjects have to look at a stopped-action picture and "compare" it with their memory of the stream of images they presumably encoded during their perception of the original continuous film, they can make this comparison with remarkable accuracy. In fact, subjects' overall performance ($d' = 2.08$) in the mixed-mode condition did not differ significantly from the overall performance of the subjects in the s-s condition ($d' = 2.22$).

Finally, one last speculative comment. Based on the pictorial recognition memory literature, we expected the performance of subjects in the s-s condition to approximate the performance of subjects in earlier picture recognition studies, because s-s subjects viewed static pictures for exposure durations that did not differ much from exposures employed by other investigators. Curiously, s-s subjects' performance did not approach the levels of excellence reported by earlier researchers. If, as suggested by the depressed scores of the s-s subjects, the pictures used were difficult either to remember or to discriminate, then it is reasonable to assert that the dynamic pictures were also difficult to

remember. In other words, subjects in the d-d condition not only performed better than subjects in the s-s condition but also accomplished this memory feat with intrinsically difficult material.

The outcome of the study, assuming that the pictorial material presented was difficult in the absolute sense and assuming that d-d subjects had more complex encoding and retrieval tasks than subjects in other conditions had, strongly suggests that memory for portions of continuously changing visual material may be much more efficient than memory for static material.

REFERENCES

- NICKERSON, R. S. Short-term memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology*, 1965, 19, 155-160.
- SHEPARD, R. N. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, 1967, 6, 156-163.
- STANDING, L., CONEZIO, J., & HABER, R. N. Perception and memory for pictures. Single trial learning of 2500 visual stimuli. *Psychonomic Science*, 1970, 19, 73-74.

(Received for publication May 13, 1982.)