Recognition memory for pictorial events

JULIE LEIBRICH and K. GEOFFREY WHITE
Victoria University of Wellington, Wellington, New Zealand

Subjects viewed sequences of slides depicting everyday events, and in later recognition tests, they correctly rejected distractors that were inconsistent with some invariant of the event but falsely accepted consistent distractors. J. J. Jenkins has accounted for the differential recognition of consistent and inconsistent slides of pictorial sequences in terms of fusion, the abstraction of an event from a series of temporally related items. The conditions under which event fusion was likely to occur were manipulated in three experiments: degree of ordering of the action-sequence slides, semantic vs. nonsemantic orienting tasks during acquisition, and duration of retention interval. Recognition performance was generally more accurate under conditions of semantic processing, ordered acquisition sequences (for the semantic task), and shorter retention intervals. However, these variables did not affect differential recognition of consistent vs. inconsistent slides. A further experiment showed that the absence of an effect of disordering acquisition sequences on differential recognition could not be attributed to subjects' reordering the disordered sequences in a way consistent with the original event. The data did not support the hypothesis that recognition performance was based on fusion of events depicted by temporally ordered slide sequences. It was suggested that memory for featural detail is a relevant factor in performance.

Biederman's (1972, 1977, 1981) extensive analysis of the comprehension of real-world scenes has established the importance of contextual organization in the recognition of individual objects. Spatially disorganizing a briefly presented scene violates the semantic cohesion or "relational congruity" of an object in that scene and results in impaired recognition of the object (Biederman, 1972, 1981). This and related findings on the perception or recognition of pictorial stimuli (Antes, 1977; Hock, Romanski, Galie, & Williams, 1978; J. M. Mandler, & Ritchey, 1977; Palmer, 1977) are consistent with the proposal of a scene schema as the basis for perception and recognition.

Extension of the scene-schema notion to account for recognition of individual scenes comprising a related sequence (Allen, Siegel, & Rosinski, 1978; Baggett, 1975; Park & Whitten, 1977; Pezdek, 1978) appears relatively straightforward. One significant development has been advanced by Jenkins (1974), who described fusion as the abstraction of an event from a series of related items. In a set of experiments examining the nature of visual events, Jenkins, Wald, & Pittenger (1978) attempted to demonstrate fusion of pictorial action sequences by showing that "specific memory for individual pictures

[was] outweighed by the abstract or general memory for the event experienced" (p. 139). In this and related studies (Kraft & Jenkins, 1977; Pittenger & Jenkins, 1979), subjects were shown an ordered sequence of slides describing a pictorial event. Jenkins et al. (1978, p. 137) posed their research questions as follows: "Would [the subject] falsely recognize pictures of the event that he had not seen before? Would he be able to reject pictures that were highly similar to the pictures he had seen but which violated some invariant of the event or some detail of the observation?" The affirmative answers to both questions led Jenkins et al. (1978) to conclude that recognition was indeed based on fusion of the originally presented slides to form a representation of the overall event. That is, in recognition testing, originally presented slides ("originals") and also distractor slides that were consistent with the event but were not previously presented ("belongings") were frequently recognized as having been presented, whereas slides inconsistent with the event by virtue of a specific violation ("transformations") were correctly rejected.

The present experiments sought to replicate and extend the main finding reported by Jenkins et al. (1978). The general procedure was refined in the following ways. First, each subject in Experiments 1, 2, and 3 was shown six short action sequences, all of the same length and standardized with respect to photographic conditions and construction. The degree to which material could be disorganized could therefore be varied by disordering slides within or between sequences. Subjects in Experiment 4 saw one long action sequence designed to be difficult to reconstruct when disordered. Second, transformations were constructed by changing

This research was based on a doctoral dissertation submitted by the first author to Victoria University of Wellington; it was supported by a University Grants Committee postgraduate scholarship to the first author and by Grants 46/77 from Victoria University of Wellington and 77/162 from the University Grants Committee. We are grateful to I. Biederman and M. Corballis for helpful comments. Reprint requests should be addressed to J. Leibrich, Department of Justice, Private Bag, Wellington, New Zealand.

one of four specified elements within a picture along the same lines as in J. M. Mandler and Johnson's (1976) study of schematic pictures. Third, whereas the test-slide categories in Jenkins' (1974; Jenkins et al., 1978) studies were restricted to specific slides, in the present experiments slides were sampled in such a way that they served as both originals and belongings. Fourth, recognizability was measured in terms of d' , whereas Jenkins et al. (1978) reported hit and false alarm rates. Jenkins et al.'s main result was that false alarm rates for belongings were much higher than those for transformations, as predicted by the fusion hypothesis. Expressing performance in terms of d' has the advantage of making explicit the fact that recognition involves two discriminations, one between originals and belongings ($d' O:B$) and one between originals and transformations ($d' O:T$), along with the added advantage of describing performance relatively independently of response bias. In these terms, high probabilities of recognizing belongings because of their consistency with the fused event would be indicated by the $d' O:B$ measure being smaller than the $d' O:T$ measure.

Experiment 1 (and subsequent experiments) replicated the Jenkins et al. (1978) finding by confirming that when subjects were shown action sequences depicting everyday events, $d' O:B$ was significantly lower than $d' O:T$ for recognition-test slides. Experiment 1 also examined the effects of disordering the slides within and between action sequences. Since an event is fused only from an ordered sequence of slides, disordering should result in higher values of $d' O:B$ relative to $d' O:T$ than should ordered sequences. Disorderer sequences offers a crucial test of the fusion hypothesis, since it results in "destroying" the event (Jenkins et al., 1978). The effect of disordering in some of Jenkins' studies was unclear (Kraft & Jenkins, 1977; Pittenger & Jenkins, 1979), and its fundamental bearing on the fusion hypothesis demanded its closer examination.

Experiment 2 also incorporated the disordering variable and examined performance following semantic and nonsemantic orienting tasks in acquisition (Craik & Lockhart, 1972). For fusion to occur, subjects must process material semantically, whereas nonsemantic processing should impair fusion. The fusion hypothesis would therefore be supported if the difference between $d' O:B$ and $d' O:T$ was greater in the semantic condition than in the nonsemantic condition.

Experiment 3 varied the duration of the retention interval between acquisition and recognition. Since it appears that memory for semantic information is relatively durable whereas memory for details decays rapidly (Bartlett, 1932; Mandler & Ritchey, 1977; Sachs, 1967, 1974), the extent to which a belonging slide was perceived to be consistent with the event should increase with longer retention intervals. Transformations, however, should not become any less inconsistent with the event. According to the fusion hypothesis, therefore, lengthening the retention interval should result in $d' O:B$ and $d' O:T$ measures becoming increasingly divergent.

In summary, Experiments 1, 2, and 3 attempted to provide support for the fusion hypothesis by examining recognition performance under conditions in which fusion should be more or less likely to occur. Since some conditions would have resulted in general increases or decreases in recognizability, the recognizability of transformations ($d' O:T$) served as a control against which changes in the recognizability of belongings ($d' O:B$) could be contrasted (cf. Jenkins et al., 1978). Variations in recognition of belongings predicted by the fusion hypothesis were therefore assessed relative to recognition of transformations.

EXPERIMENTS 1, 2, AND 3

Method

Subjects. Three groups of eight subjects (16 women, 8 men) and six groups of eight subjects (23 women, 25 men) served in Experiments 1 and 2, respectively. There were three groups of 16 subjects (22 women, 26 men) in Experiment 3. Mean ages (and standard deviations) of subjects in the three experiments were 22.8 (7.1), 21.0 (4.1), and 23.8 (6.6) years, respectively. Subjects (mainly university students) took part in the experiments individually, and all instructions were written.

Stimulus Materials. A Nikkormat camera fitted with an F2 50-mm lens and ultraviolet filter was used to make six sets of slides. The camera was set on a tripod so that each of the six sets of slides was taken from a fixed stationary point. Some sequences were taken by natural light, whereas electronic flash lighting was used for others. Kodachrome ASA 64 film was used throughout.

Each set consisted of 20 slides, 16 of which depicted an action sequence in a consistent and progressive manner and 4 of which changed some aspect of the sequence. Each action sequence was a distinct activity performed by one person. The activities were as follows: (1) a woman paying a bill, (2) a woman ironing a tablecloth, (3) a woman hanging out washing, (4) a man chopping wood, (5) a man making a sandwich, and (6) a man lighting a fire.

Before photographing each activity, it was rehearsed by the actor and timed. The real-time duration of activities varied between 80 and 120 sec. When the time for an activity had been established, it was divided by 16 so that each of the 16 photographs depicting the activity could be taken on cue at equal intervals throughout the action. Thus, depending on the duration of the activity, a photograph was taken every 5, 6, or 7 sec. In this way, six sets of 16-frame sequences were made, each one depicting an everyday activity.

The four transformation slides changed some aspect of the actual activity in such a way that although each was similar to the 16 slides that depicted the activity, it differed from them according to one of four specific rules: (1) Addition—An object that might reasonably appear in such a setting but had not been included was added. For example, a hearth-brush was added to the "man lighting a fire" sequence. (2) Token—Some change was made in the appearance of the actor. For example, the man making a sandwich wore glasses. (3) Perspective—A wide-angle lens (F2 28 mm) was substituted for the F2 50-mm lens. (4) Orientation—A duplicate of 1 of the 16 slides was made and reversed with respect to left-right orientation. In photographing the scenes, the four transformation slides were spaced at approximately equal intervals over the course of the whole activity. Addition, token, and perspective slides were made at the same time as the 16 slides depicting the activity. The action was "frozen" at the appropriate moment, the change was made, and the photograph was taken, and then the original scene was restored.

Apparatus. Slides were projected onto a white wall by means of two Kodak Carousel SAV 2000 projectors that operated in

direct succession. The projected image was .45 x .69 m for both acquisition and test stimuli, and subjects sat approximately 1.5 m from the screen. Each slide was shown for 5 sec in Experiments 1 and 2 and for 4 sec in Experiment 3, with no inter-stimulus interval. Exposure duration was controlled by Gerbrands tachistoscopic shutters mounted in front of the projector lenses. Experimental events were controlled by relay and solid state programming apparatus.

Procedure. Eight random orders of sets for presentation of the six sets of slides were generated. The first subject in each condition saw one of these orders, the second subject saw another, and so on. Four groups of evenly spaced belonging slides could be withdrawn from any sequence of 16 slides depicting an activity: that is, slide numbers 1, 5, 9, 13; 2, 6, 10, 14; 3, 7, 11, 15; or 4, 8, 12, 16. The group of belonging slides withdrawn from a sequence was randomly determined for each of the six sequences seen by the first subject in each condition, and then again for each of the six sequences seen by the second subject in each condition, and so on. The remaining 12 original slides were shown during acquisition. Four of these were selected at random to be shown during testing. All first subjects in each condition were shown the same original slides during testing, as were second subjects, and so on. In this way, eight different selections of 72 slides were presented during acquisition, one for each of eight subjects within a group.

Subjects in the ordered conditions of Experiments 1 and 2 saw the slides within each sequence in their actual chronological order, as did all subjects in Experiment 3. In Experiments 1 and 2, subjects in the disordered-within condition were presented with randomly ordered slides within sequence, but the sequences were shown successively. Subjects in the disordered-between condition saw all 72 slides completely randomized across all six sequences of 12 slides.

The recognition material consisted of four original slides, four belonging slides, and four transformation slides from each of the six sets. The order of these 72 test slides was determined randomly, with no distinction being made between the sets. An order was generated for all first subjects in each condition, another order was generated for all second subjects, and so on.

In the acquisition phase of Experiments 1 and 3, subjects were told they were taking part in a memory experiment and were asked to try to remember the slides that were about to be shown to them. The 72 acquisition slides were then presented. In Experiment 1, and in the immediate condition of Experiment 3, the recognition test was conducted in the same way for all subjects approximately 3 min after acquisition. In the delayed conditions of Experiment 3, subjects returned either 24 h later or 1 week later for the recognition test phase. During the interval between the acquisition and recognition phases, subjects read instructions asking them to indicate whether they had seen the slides before. They were not told about the prior probabilities of distractor slides.

In the acquisition phase of Experiment 2, subjects were told they were taking part in a "perception" experiment. Subjects in the semantic conditions were asked to judge the logical progression of the slides. They had to complete a form as they viewed the acquisition material, by marking "Yes" or "No" for the question "Does this slide follow on logically from the previous slide?" for every slide except the first. Subjects in the nonsemantic conditions were asked to judge the relative brightness levels of the slides. They had to complete a similar form by marking "Yes" or "No" for the question "Is this slide brighter than the previous slide?" for every slide except the first.

Results and Discussion

Experiment 1. Hit rates for originals and false alarm rates for belongings and transformations (Table 1) confirmed the results of the Jenkins et al. (1978) study,

Table 1
Mean Hit and False Alarm Rates for Target and Distractor Slides as a Function of Presentation Order in Experiment 1

	Original		Belonging		Transformation	
	Mean	SD	Mean	SD	Mean	SD
Ordered	.76	.11	.51	.13	.29	.11
Disordered-Within	.84	.10	.58	.13	.30	.09
Disordered-Between	.79	.11	.43	.13	.30	.09

Table 2
Mean d' Measures as a Function of Presentation Order in Experiment 2

	$d'O:B$		$d'O:T$	
	Mean	SD	Mean	SD
Ordered	.73	.38	1.35	.68
Disordered-Within	.85	.40	1.59	.43
Disordered-Between	1.02	.35	1.37	.46

with consistently higher false alarm rates for belongings than for transformations. Hit and false alarm rates were converted to d' measures (Green & Swets, 1966) describing recognizability of originals relative to false recognitions of belongings ($d'O:B$) or transformations ($d'O:T$). The mean (and standard deviation) d' measures for each type of distractor are given in Table 2. There was no main effect of order of presentation ($F < 1$). The within-subjects variable of distractor type had a significant effect, with $d'O:B$ lower than $d'O:T$, as predicted by the fusion hypothesis [$F(1,21) = 35.18$, $MSe = .11$, $p < .001$]. Although $d'O:B$ measures for the three groups of subjects appeared to increase from ordered through disordered-within to disordered-between conditions of presentation (Table 2), the interaction between presentation condition and distractor type was not significant [$F(2,21) = 1.45$, $MSe = .16$, $p > .05$]. While replicating the main effect reported by Jenkins et al. (1978), data from this experiment failed to show an increase in $d'O:B$ relative to $d'O:T$, although the pictorial event had been disordered.

Experiment 2. Six groups of subjects in Experiment 2 combined semantic or nonsemantic orienting tasks with three levels of presentation order in acquisition. As in Experiment 1, there was within-subjects manipulation of distractor category. False alarm rates were higher overall for belongings than for transformations (Table 3), again confirming the main effect of the Jenkins et al. (1978) study. For the d' measures given in Table 4, presentation order had no effect [$F(2,42) = 2.56$, $MSe = .32$, $p > .05$], but d' values were higher overall for the semantic orienting task than for the nonsemantic task [$F(1,42) = 11.34$, $MSe = .32$, $p < .01$]. The $d'O:B$ measures were much lower than $d'O:T$ [$F(1,42) =$

Table 3
Mean Hit and False Alarm Rates for Target and
Distractor Slides as a Function of Presentation
Order and Orienting Task in Experiment 2

	Original		Belonging		Trans- formation	
	Mean	SD	Mean	SD	Mean	SD
Semantic Task						
Ordered	.84	.12	.57	.12	.38	.80
Disordered-Within	.81	.11	.59	.13	.37	.10
Disordered-Between	.71	.17	.61	.13	.37	.10
Nonsemantic Task						
Ordered	.65	.21	.56	.21	.33	.17
Disordered-Within	.59	.15	.55	.17	.26	.11
Disordered-Between	.55	.19	.48	.17	.30	.14

Table 4
Mean d' Measures as a Function of Presentation
Order and Orienting Task in Experiment 2

	d' O:B		d' O:T	
	Mean	SD	Mean	SD
Semantic Task				
Ordered	.93	.64	1.42	.51
Disordered-Within	.68	.29	1.25	.33
Disordered-Between	.32	.23	.95	.44
Nonsemantic Task				
Ordered	.19	.48	.98	.42
Disordered-Within	.13	.39	.94	.48
Disordered-Between	.24	.43	.74	.51

137.14, $MSe = .07$, $p < .001$]. No interactions were significant.

The absence of interactions and no effect of presentation order are consistent with the data of Experiment 1 in indicating a failure to show that disordering action sequences enhanced d' O:B relative to d' O:T, particularly for the semantic orienting task, in which fusion was most likely to occur. The absence of an interaction between orienting task and distractor category also failed to demonstrate stronger event fusion under the semantic condition.

The overall increase in d' measures in the semantic condition, however, is consistent with an enormous literature documenting the effects of differential encoding on memory performance (Baddeley, 1978; Craik & Tulving, 1975). Responses in the semantic orienting task to the question of whether slides were logically progressive gave a measure of perceived degree of order of the acquisition slides for the different presentation conditions. Mean frequencies of "No" responses were 10.75, 34.13, and 55.38 for the ordered, disordered-within, and disordered-between conditions, respectively [$F(2,21) = 123.93$, $MSe = 37.52$, $p < .001$]. In the ordered condition, a minimum of five "No" responses was expected, since six action sequences were presented successively.

The different presentation orders were therefore perceived to differ in their degree of orderliness. The differential encoding effect is sensitive to different degrees or "depths" of semantic processing (Craik, 1973; Klein & Saltz, 1976; Marslen-Wilson & Tyler, 1976; Mistler-Lachman, 1974). We would therefore have expected an effect of presentation order on d' measures for the semantic group, and yet the Presentation by Orienting task interaction was not significant. However, it was decided to make two separate post hoc examinations of the effect of order. For the semantic group alone, d' values were significantly influenced by both presentation order [$F(2,21) = 3.55$, $MSe = .33$, $p < .05$] and distractor category [$F(1,21) = 96.25$, $MSe = .04$, $p < .001$], although there was no significant interaction. A subsequent Scheffé test showed that d' measures were higher in the ordered condition than in the disordered-between condition ($p < .001$). For the nonsemantic group alone, distractor category had a significant effect on d' [$F(1,21) = 58.60$, $MSe = .10$, $p < .001$], but there was no effect of presentation order and no interaction.

These separate analyses revealed that disordering acquisition material affects recognition accuracy in the semantic group but not in the nonsemantic group, consistent with the literature showing that semantic encoding effects are sensitive to the degree to which material can be processed semantically.

Experiment 3. Recognition tests were conducted for three groups of 16 subjects, immediately, 24 h, or 1 week following acquisition. Hit rates were found to be roughly stable over the three retention intervals, whereas false alarm rates for both belongings and transformations increased (Table 5). Table 6 shows that d' O:B and d' O:T did not change from immediate test-

Table 5
Mean Hit and False Alarm Rates for Target and Distractor
Slides as a Function of Retention Interval in Experiment 3

	Original		Belonging		Trans- formation	
	Mean	SD	Mean	SD	Mean	SD
Immediate	.77	.11	.59	.09	.32	.12
24 Hour	.81	.14	.63	.14	.38	.10
One Week	.77	.09	.67	.13	.44	.17

Table 6
Mean d' Measures as a Function of Retention
Interval in Experiment 3

	d' O:B		d' O:T	
	Mean	SD	Mean	SD
Immediate	.59	.44	1.25	.64
24 Hour	.62	.47	1.26	.47
One Week	.27	.39	.89	.52

ing to testing after a 24-h delay, but after a 1-week delay both measures showed a considerable decrease [$F(2,45) = 3.36$, $MSe = .39$, $p < .05$]. At all times, $d'O:B$ was lower than $d'O:T$ [$F(1,45) = 19.92$, $MSe = .49$, $p < .001$]. There was, however, no interaction between distractor category and retention interval ($F < 1$), as reflected in the virtually identical differences between $d'O:B$ and $d'O:T$ at the different retention levels. The persistence of memory over 24 h is consistent with previous studies that have examined retention interval effects on memory for line drawings (Baggett, 1975; J. M. Mandler & Ritchey, 1977). The absence of a differential change in $d'O:B$ and $d'O:T$ with increasing retention interval, however, was not the pattern of response expected if fusion of the event was to increase over time.

EXPERIMENT 4

In the previous experiments, the absence of an effect of disordering the action sequences was interpreted as a failure of support for the fusion hypothesis. However, a criticism of this test of fusion is that subjects may reconstruct the original event from the disordered sequence. Indeed, this argument has been offered in previous studies in which disordering effects failed to support the fusion hypothesis (Jenkins et al., 1978; Kraft & Jenkins, 1977; Pittenger & Jenkins, 1979). For verbally presented stories, subjects are likely to reorder disorganized stories into their natural schematic order (Bower, Black, & Turner, 1979; Kintsch, Mandel, & Kozminsky, 1977; J. M. Mandler, 1978; Stein & Glenn, 1978), although Baker (1978) found no evidence that subjects reconstructed story order when shown disordered stories. In Experiment 4, the reordering criticism was examined by producing a pictorial action sequence that, when disordered, could not be reconstructed into its original event. A long slide sequence was devised in which the event depended on the exact chronological order of a series of encounters between people for its definition. Different groups of subjects saw ordered or disordered presentations of the long sequence before being tested for recognition of the slides under conditions similar to those of the previous experiments. In particular, subjects in the reconstruction groups were asked to write an account of the story prior to the recognition test. The written accounts were judged by another group of subjects for their degree of correspondence to the original sequence.

Method

Subjects. Sixty subjects (49 women and 11 men) with a mean age of 21.9 years ($SD = 5.3$) participated individually. Four groups of 12 subjects served in the ordered or disordered by recognition or reconstruction conditions, and another 12 subjects served as judges.

Stimulus Materials. A set of stimulus materials was produced by the same method as in previous experiments, except that

one set of 120 slides was made. The real-time duration of the sequence was approximately 11 min, and each of the 96 slides depicting the action was taken at 7-sec intervals. Twenty-four transformations were made, six according to each of the specific rules: addition, perspective, token, and orientation.

The action sequence involved four actors, and the setting was a park bench. The story was a series of encounters between men and women that resulted in new friendships. The outline of the story is as follows: A young man seated on a park bench opens a bag of fruit and eats. An older man arrives and sits on the other end of the bench, and the two talk before the older man starts to read a magazine. A young woman appears, sits on the bench, and lights a cigarette. She talks with the young man, who offers her some fruit, and they leave together. The older man watches and then continues his reading. An older woman arrives, sits on the bench, and feeds some birds while talking to the older man. After a while, the man leaves but forgets his magazine, which the woman is reading when he returns. He sits down again and they talk, both feed the birds, and finally leave together.

Procedure. Each of the 24 subjects in the ordered conditions and 24 subjects in the disordered conditions were shown slides in the acquisition phase under the same conditions as in Experiment 3. Four sets of 72 acquisition slides (originals) included every 4th slide from the initial sequence of 96 slides, with the 1st slide differing for the different sets. Each set was presented to six subjects in each presentation condition. In the ordered conditions, the 72-frame action sequences were presented in the correct chronological order. In the disordered conditions, the order of slides was randomized.

For 12 subjects from each of the ordered and disordered presentation conditions, the acquisition phase was followed immediately by a recognition test ("recognition" groups). For remaining subjects ("reconstruction" groups) the story-writing phase described below preceded the recognition test.

In recognition testing, subjects indicated whether they had seen each slide before. The 72 randomly ordered test slides comprised 24 each of originals, belongings, and transformations. The slides were chosen in such a way (because of the design of the acquisition sequences) that over subjects within a group, each original and each belonging were sampled three times and each transformation was sampled 12 times.

For subjects in the reconstruction groups, a period not exceeding 14 min intervened between acquisition and recognition testing, during which subjects were asked to write a description of the story they had just seen. Reconstruction-group subjects in the ordered condition were told that the slides they had just seen "depicted a simple story" and that they "saw them in their correct order." Subjects in the disordered conditions were told that the slides they had just seen "depicted a simple story" but they "saw them jumbled up." These subjects were asked to describe the story depicted by the slides, "as if you had been shown them in their correct order."

Typewritten copies of the 24 stories produced by subjects in the reconstruction groups were read by 12 independent judges who were shown the original 96-frame action sequence twice. The judges were asked to rank order the 24 stories according to the extent to which they corresponded to the original action sequence. Following the ranking task, the judges were asked to rate each story on a scale of 1 to 10 according to its correspondence to the original sequence, where 10 indicated extremely close correspondence. The tasks required about 2 h for each judge.

Results and Discussion

Written descriptions of the ordered acquisition sequence were judged to have greater correspondence to the original 96-frame action sequence than were descriptions of the disordered acquisition sequence. Rank

orders assigned by the 12 judges were averaged over stories within a presentation condition. The mean rank order for stories produced by subjects in the ordered group (7.83) was notably higher than that for stories produced by subjects in the disordered group (17.17). The actual mean ranks obtained were close to those associated with exact separation (6.5 compared to 18.5). The degree of interjudge agreement was reasonably high, according to Kendall's coefficient of concordance, which was $W = .72$. The mean rating for "ordered" stories (6.27) was greater than that for "disordered" stories (2.97) [$t(23) = 66.33$, $p < .001$]. We were therefore confident that subjects instructed to reorder a disordered sequence were not able to generate a description consistent with the original event.

The hit and false alarm rates for originals, belongings, and transformations under the two conditions of presentation order are given in Table 7, which again shows the main effect reported by Jenkins et al. (1978): higher false alarm rates for belongings than for transformations. The d' measures confirmed this difference (Table 8), with $d' O:B$ smaller than $d' O:T$, for both reconstruction [$F(1,22) = 141.04$, $MSe = .23$, $p < .001$] and recognition [$F(1,22) = 85.05$, $MSe = .39$, $p < .001$] groups. But in neither case was the effect of presentation order significant, nor were there significant interactions ($F < 1$ in all cases).

Table 9 gives proportions of correct rejections made by the recognition groups for each of the four types of

Table 9
Mean Proportions of Correct Rejections for Transformation Slides for Recognition Groups in Experiment 4

	Addition		Orientation		Perspective		Token	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Ordered	.85	.13	.93	.13	.94	.11	.75	.26
Disordered	.76	.24	.98	.06	.90	.25	.85	.18

transformations. Presentation order had no effect on the proportion of correct rejections ($F < 1$), but correct rejections did differ across transformations [$F(3,66) = 5.90$, $MSe = 268$, $p < .01$]. There was no interaction between presentation order and type of transformation [$F(3,66) = 1.50$, $p > .05$]. Scheffé tests showed that orientation transformations were better detected than token and addition ($p < .001$) and perspective slides were better detected than token and addition ($p < .01$). Kraft and Jenkins (1977) also found that orientation transformations were the most frequently detected. The absence of an interaction between presentation order and transformation type suggests that our other results are general to all categories of transformation employed in the present experiments.

An analysis including d' measures for all groups (Table 8) confirmed the absence of a presentation effect ($F < 1$), the presence of a strong distractor effect [$F(1,44) = 211.65$, $MSe = .31$, $p < .001$], and no interactions. In addition, however, there was a reliable effect of recognition vs. reconstruction condition [$F(1,44) = 7.39$, $MSe = .57$, $p < .01$], reflecting the generally higher d' values for the recognition groups than for the reconstruction groups that were required to perform the intervening story-writing task (Table 8). The story-writing task was predominantly semantic in nature, whereas the subsequent recognition test examined memory for discrete pictorial items. The lower d' measures for the reconstruction groups are therefore consistent with Hunt and Elliot's (1980) conclusion that memory for surface form is impaired when subjects are given an intervening nonfeatural task to perform.

In summary, disordering long action sequences in Experiment 4 failed to generate the increase in $d' O:B$ relative to $d' O:T$ predicted by the fusion hypothesis. Further, our data suggest that this failure cannot be accounted for in terms of subjects' reordering the disordered sequence to form the original event. Whereas it is possible that the subjects could have reordered the disordered slides to form some event other than the original, belonging slides that were recognized consistent with fusion belonged to the original event depicted by the ordered sequence and not to some other event. It appears that variables other than event fusion are effective in determining recognition performance and, in particular, the difference between $d' O:B$ and $d' O:T$ measures.

Table 7
Mean Hit and False Alarm Rates for Target and Distractor Slides as a Function of Presentation Order in Experiment 4

	Original		Belonging		Transformation	
	Mean	SD	Mean	SD	Mean	SD
Reconstruction						
Ordered	.80	.10	.68	.15	.15	.08
Disordered	.76	.12	.65	.16	.13	.08
Recognition						
Ordered	.86	.13	.67	.20	.14	.13
Disordered	.79	.13	.58	.11	.12	.12

Table 8
Mean d' Measures as a Function of Presentation Order in Experiment 4

	$d' O:B$		$d' O:T$	
	Mean	SD	Mean	SD
Reconstruction				
Ordered	.40	.52	1.99	.43
Disordered	.33	.29	2.03	.74
Recognition				
Ordered	.83	.39	2.59	1.05
Disordered	.72	.53	2.29	.94

GENERAL DISCUSSION

Jenkins et al. (1978) had found that after viewing temporally ordered pictorial sequences, subjects discriminated well between original slides and transformations that were inconsistent with the sequence, compared to their ability to discriminate between originals and belongings that were consistent with the event depicted by the sequence. Jenkins et al. interpreted the differential response to the distractor types as evidence of fusion when recognition performance was based on the comparison of test slides to the abstracted schema of the total event in acquisition. In each of the four experiments we reported here, the main effect described by Jenkins et al. (1978) was replicated for pictorial action sequences constructed along the same lines as those used by Jenkins et al.; that is, d' values for the discrimination between originals and belongings were smaller than those for the discrimination between originals and transformations.

A crucial test of the fusion hypothesis was made by disordering the pictorial sequence presented in acquisition. In the present experiments, however, disordering had an effect only in the semantic condition of Experiment 2, and then only in a way consistent with the general improvement of memory by organization (Kintsch, 1977; G. Mandler, 1972; Tulving, 1972), and not in a way that supported the fusion hypothesis (Jenkins et al., 1978). That is, d' measures were equally decreased by disordering the sequence, rather than differentially for the different distractor categories, as predicted by the fusion hypothesis. Experiment 4 demonstrated that absence of a differential effect of disordering on recognition performance was not due to subjects' reconstruction of the disordered sequences. Manipulations of other variables expected to enhance or attenuate the likelihood of the events being fused, that is, the nature of the orienting task and retention interval duration, also failed to produce the effects on recognition performance expected according to the fusion hypothesis. It might be argued that since many of the effects we expected to find would have been manifested in significant interactions, failure to detect interactions may be due to the use of unequal-interval scales for the dependent measures. However, the particular interactions we expected according to the fusion hypothesis were of the interpretable variety (Loftus, 1978), in that scaling changes should not mask their appearance.

The general failure of the present experiments to provide support for the fusion hypothesis suggests that factors other than fusion may account for the main effect (the d' O:B vs. d' O:T difference) that the fusion hypothesis was originally intended to explain.

An alternative explanation is that transformations were featurally less similar to originals than were belongings and that subjects' performance was not in fact evidence of fusion of the acquisition event, but rather,

of confusion with respect to detail in the recognition task. Although Jenkins et al. (1978) rejected this alternative account of performance, the present set of experiments found no evidence that it should be dismissed. However, assessment of a featural account of the data would require an independent measure of featural similarity.

Although relative featural dissimilarity may contribute to the performance pattern obtained, event fusion may also be occurring, but in a way unrelated to the temporal order of the sequenced slides. Consistent with our data, Allen et al. (1978) found that disordering the scenes from a walk had no effect on differential recognition of scenes belonging to the walk vs. scenes from other walks (transformations). Allen et al. concluded that the temporal order of the scenes was not necessary for acquisition of knowledge of the walk's route. In that and the present studies, the variables of features or spatial relations within slides and meaning or relations between slides may interact in determining recognition performance. Biederman (1977) demonstrated that object identification and scene recognition were more accurate when scenes were presented in a meaningful organized way. He also demonstrated that in labeling briefly presented scenes, identification of discrete objects in the scene influenced perceptual decisions. Both overall meaning and the specific features of objects in scenes or sequences may be relevant to perceptual processing and recognition performance. However, procedures designed to assess the influence of either of these two main variables are likely to involve a confounding with the other.

REFERENCES

- ALLEN, G. L., SIEGEL, A. W., & ROSINSKI, R. R. The role of perceptual context in structuring spatial knowledge. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, 4, 617-630.
- ANTES, J. R. Recognizing and localizing features in brief picture presentations. *Memory & Cognition*, 1977, 5, 155-161.
- BADDELEY, A. D. The trouble with levels: A re-examination of Craik and Lockhart's framework for memory research. *Psychological Review*, 1978, 85, 139-152.
- BAGGETT, P. Memory for explicit and implicit information in picture stories. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 538-548.
- BAKER, L. Processing temporal relationships in simple stories: Effects of input sequence. *Journal of Verbal Learning and Verbal Behavior*, 1978, 17, 559-572.
- BARTLETT, F. C. *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press, 1932.
- BIEDERMAN, I. Perceiving real-world scenes. *Science*, 1972, 177, 77-80.
- BIEDERMAN, I. On processing information from a glance at a scene: Some implications for a syntax and semantics of visual processing. In S. Treu (Ed.), *User-oriented design of interactive graphics systems*. New York: Association for Computing Machinery, 1977.
- BIEDERMAN, I. On the semantics of a glance at a scene. In

- M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, N.J: Erlbaum, 1981.
- BOWER, G. H., BLACK, J. B., & TURNER, T. J. Scripts in memory for text. *Cognitive Psychology*, 1979, 11, 177-220.
- CRAIK, F. I. M. A "levels of analysis" view of memory. In P. Pliner, L. Krames, & T. M. Alloway (Eds.), *Communication and affect: Language and thought*. New York: Academic Press, 1973.
- CRAIK, F. I. M., & LOCKHART, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 671-684.
- CRAIK, F. I. M., & TULVING, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 1975, 104, 268-294.
- GREEN, D. M., & SWETS, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.
- HOCK, H. S., ROMANSKI, L., GALIE, A., & WILLIAMS, C. S. Real-world schemata and scene recognition in adults and children. *Memory & Cognition*, 1978, 6, 423-431.
- HUNT, R. R., & ELLIOTT, J. M. The role of non-semantic information in memory: Orthographic distinctiveness effects on retention. *Journal of Experimental Psychology: General*, 1980, 109, 49-74.
- JENKINS, J. J. Remember that old theory of memory? Well, forget it! *American Psychologist*, 1974, 29, 785-795.
- JENKINS, J. J., WALD, J., & PITTENGER, J. B. Apprehending pictorial events: An instance of psychological cohesion. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science* (Vol. 9). Minneapolis: University of Minnesota Press, 1978.
- KINTSCH, W. *Memory and cognition*. New York: Wiley, 1977.
- KINTSCH, W., MANDEL, T. S., & KOZMINSKY, E. Summarizing scrambled stories. *Memory & Cognition*, 1977, 5, 547-552.
- KLEIN, K., & SALTZ, E. Specifying the mechanisms in a levels-of-processing approach to memory. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, 2, 671-679.
- KRAFT, R. N., & JENKINS, J. J. Memory for lateral orientation of slides in picture stories. *Memory & Cognition*, 1977, 5, 397-403.
- LOFTUS, G. On interpretation on interactions. *Memory & Cognition*, 1978, 6, 312-319.
- MANDLER, G. Organization and recognition. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- MANDLER, J. M. A code in the node. *Discourse Processes*, 1978, 1, 14-35.
- MANDLER, J. M., & JOHNSON, N. S. Some of the thousand words a picture is worth. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, 2, 529-540.
- MANDLER, J. M., & RITCHEY, G. H. Long-term memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, 3, 386-396.
- MARSLER-WILSON, W., & TYLER, L. K. Memory and levels of processing in a psycholinguistic context. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, 2, 112-119.
- MISTLER-LACHMAN, J. L. Depth of comprehension and sentence memory. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 98-106.
- PALMER, S. E. Hierarchical structure in perceptual representation. *Cognitive Psychology*, 1977, 9, 441-475.
- PARK, C. D., & WHITTEN, W. B. Abstraction of linguistic, imaginal, and pictorial ideas. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, 3, 525-538.
- PEZDEZ, K. Recognition memory for related pictures. *Memory & Cognition*, 1978, 6, 64-69.
- PITTENGER, J. B., & JENKINS, J. J. Apprehension of pictorial events: The case of a moving observer in a static environment. *Bulletin of the Psychonomic Society*, 1979, 13, 117-120.
- SACHS, J. Recognition memory of syntactic and semantic aspects of connected discourse. *Perception & Psychophysics*, 1967, 2, 437-444.
- SACHS, J. Memory in reading and listening to discourse. *Memory & Cognition*, 1974, 2, 95-100.
- STEIN, N. L., & GLEN, C. G. An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *Multi-disciplinary perspectives in discourse comprehension*. Hillsdale, N.J: Erlbaum, 1978.
- TULVING, E. Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization in memory*. New York: Academic Press, 1972.

(Received for publication July 6, 1982;
revision accepted November 15, 1982.)