Journal of Experimental Psychology:
Learning, Memory, and Cognition
1989, Vol. 15, No. 1, 98–109

0278-7393/89/$00.75

Illusory Conjunctions of Forms, Objects, and Scenes During Rapid Serial Visual Search

Helene Intraub
University of Delaware

"Temporal migration" describes a situation in which subjects viewing rapidly presented stimuli (e.g., 9–20 items/s) confidently report a target element as having been presented in the same display as a previous or following stimulus in the sequence. Four experiments tested a short-term buffer model of this phenomenon. Experiments 1 and 4 tested the hypothesis that subjects' errors are due to the demands of the verbal report procedure rather than to perceptual integration. In Experiment 1, 12 color objects were presented at a rate of 9/s. Prior to each sequence, an object was named and subjects responded "yes" or "no" to indicate whether the target element (a black frame) occurred with that object. Consistent with the perceptual hypothesis, the yes/no procedure yielded the same results as the verbal report procedure. Experiment 2 tested the hypothesis that the direction of migration depends on "frame" detection time. Results showed that reaction time to frame detection was significantly faster in trials in which subjects reported the frame on a preceding rather than a following picture. Experiments 3 and 4 used the standard naming procedure and the yes/no procedure to test temporal migration using more complex, interrelated stimuli (objects and scenes). Implications for the use of the temporal migration effect to study visual integration within eye fixations are discussed.

A compelling phenomenon, dating back at least as far as the "prior entry" research of Wundt and his contemporaries, is the perceptual dissociation of concurrently presented stimuli (see James, 1890). Recently, there has been renewed interest in these phenomena because of their utility in studying the role of attention and memory in visual perception (e.g., Reeves & Sperling, 1986; Treisman & Gelade, 1980). Temporal illusions obtained in the visual modality have been reported in several experiments using high-speed presentation of visual displays (Gathercole & Broadbent, 1984; Intraub, 1985b; Lawrence, 1971; McLean, Broadbent, & Broadbent, 1983). In these experiments, stimuli (letters, words, or pictures) were presented in rapid succession (e.g., 9–20 items/s), in the same spatial location on a screen. The subjects were required to immediately report the stimulus presented in a particular color, in a particular letter case, or simultaneous with a black frame. The basic result was that subjects often reported not the target stimulus (the one actually bearing the searched-for feature), but a temporally adjacent stimulus in the sequence.

Intraub (1985b) has proposed a working model based primarily on picture perception (Biederman, Mezzanotte, & Rabinowitz, 1982; Biederman, Rabinowitz, Glass, & Stacy, 1974; Intraub, 1981a; Potter, 1975, 1976) and picture memory research (Intraub, 1980, 1984, 1985a; Loftus & Ginn, 1984; Potter, 1976) that provides an account of previous temporal migration research and has led to specific testable predictions about the early stages of scene processing. The basic assumptions are as follows:

1. When a well-formed picture of an object or scene is presented, after approximately 100 ms of uninterrupted viewing, the information is not likely to suffer visual masking (Loftus & Ginn, 1984; Potter, 1976). The information can then be maintained for a few hundred milliseconds in a very short-term store, during which time conceptual/visual processing continues (Avons & Phillips, 1980; Potter, 1976). The very short-term store allows for momentary identification of rapidly presented pictures that may not be remembered seconds later (Intraub, 1981a, 1981b; Potter, 1975, 1976). There is evidence to suggest that the buffer can hold more than one picture at a time (Intraub, 1984, 1985a).

2. It is hypothesized that during the early stages of processing visual components are integrated and understood as part of the same display. The assumption is that when two pictures are simultaneously processed in the very short-term store, during this integration stage, temporal migration errors between displays may occur. This assumption was tested by requiring subjects to detect the conjunction of a black outline frame and an object in a sequence of rapidly presented objects, and the conjunction of the same frame with a number in a rapidly presented sequence of numbers (Intraub, 1985b, Experiment 4). The rationale was that because numbers can be identified so much more rapidly than colored photographs, *given the same rate of presentation (9/s)*, simultaneous processing during the integration stage would be less likely for temporally adjacent numbers than for temporally adjacent pictures.

This research was supported by National Science Foundation Grants BNS-8217390 and BNS-8519417. Experiment 2 was presented at the annual meeting of the Cognitive Science Society, August, 1986, Amherst, MA. Experiments 3 and 4 were reported at the annual meeting of the Psychonomic Society, November, 1985, Boston, MA.

I thank Jim Pomerantz for his comments on an earlier draft and Alan Scott, Janice Wilhoit, and Beth Dietz for their valuable assistance in conducting these experiments.

Correspondence concerning this article should be addressed to Helene Intraub, Department of Psychology, University of Delaware, Newark, Delaware 19716.

The assumption was supported in that no frame migration was obtained when the numbers were shown at the same rate as the pictures. Yet, when presentation rate of the numbers was increased (to 18/s), thereby increasing the likelihood of simultaneous processing, frame migration reemerged. Similarly, to obtain an equal frequency of migrations, Gathercole and Broadbent (1984) reported having to slow down presentation rate when they changed their stimuli from letters to words.

3. It is hypothesized that a migration error is more likely to occur on those trials in which the searched-for component takes a relatively long time to become integrated within its display than when the component can be rapidly integrated. It is important to note that parts of objects themselves do not typically dissociate and merge with temporally adjacent objects (Intraub, 1981b), but rather dissociation and merging frequently occur when an unrelated visual component (e.g., a black outline frame or a homogeneous colored background) is photographed in the same display as one of the objects. The component does not have to be physically separate from the object in order to migrate and may even be superimposed in the center of the object itself (Intraub, 1985b, Experiment 2).

4. The assumption that subjects' high confidence errors reflect integration errors between pictures in the buffer is suggested by another set of results. In support of this view, subjects apparently perceive the host picture (the one actually presented in the frame) on trials in which the frame migrates. They will report it, for example, as the picture *following* the picture in the frame (Intraub, 1985b, Experiment 3). This argues against a masking interpretation of temporal migration, in which the host picture is obscured on certain trials as a result of visual or conceptual masking and the frame is simply reported around the most salient temporally adjacent event.

5. According to the model, the reason that the frame migrates to the preceding picture on some trials and the following picture on others (a pattern that is typical of temporal migration studies) is that because the frame is unrelated to the object (in terms of world knowledge about the objects), instead of rapidly integrating the frame and object as a single display, the subject begins processing the picture and the frame successively. On trials in which the subject begins to process the frame first and then the "host" picture, the frame may sometimes become integrated with the ongoing processing of the previous picture, which is in the very short-term store. In trials in which the subject begins processing the host picture first and then the frame, the frame may sometimes become integrated with processing of the next picture in the sequence.

The purpose of the present four experiments was to test and extend the model. Experiments 1 and 4 addressed a major methodological issue. All of the high-speed presentation experiments have used verbal report. Surprisingly, these reports have been accepted as indicative of what the subject perceived. It is important to recall the speed at which items are presented in such studies and to recognize that the monitoring task, under these conditions, is a very difficult one, requiring focal attention and considerable concentration. The requirement that subjects must identify and name the targeted object may (a) result in a report error that does not actually reflect a perceptual integration error or (b) add sufficient attentional demands in its own right to affect the task, raising concerns about how much the illusion relies on a specific response measure.

Regarding the first issue, the possibility that verbal report errors reflect a later, nonperceptual stage of processing (e.g., response organization) was a major issue in the controversy surrounding interpretation of the auditory "click migration" effect obtained in research on sentence comprehension (see Fodor, Bever, & Garrett, 1974, for a review). It is interesting that the visual literature has essentially unquestioningly accepted the verbal reports in high-speed search experiments as reflecting true illusory conjunctions. Regarding the second issue, a critical question to ask in order to understand the illusion is, "How dependent is it on the specific procedure used in these experiments?" At issue is whether advance notification of the frame's potential host picture, coupled with an elimination of the naming requirement, would yield evidence of temporal migration.

Experiment 2 tested the hypothesis that temporal migration is the result of "on-line" integration errors, rather than errors in subjects' postperceptual simultaneity decisions. A reaction time dependent measure was introduced to determine if the direction of migration (preceding versus following picture in the sequence) depends on how early or late the frame can be detected during presentation and to address the issue of whether processing of information in the very short-term buffer is conducted serially or in parallel.

Experiments 3 and 4 provided a test of the generality of the temporal migration effect with respect to stimulus organization. These experiments were designed to determine if complex interrelated scenes would dissociate during high-speed presentation. The specific purpose was to determine if one could obtain "object migration" among scenes. The more long-range goal was to explore the possibility of using temporal migration in future research on the early stages of scene perception.

Experiment 1

Subjects took part in an experiment similar to the picture experiments reported by Intraub (1985b) in which subjects searched for a black frame and reported the name of the object with which it was represented. In the present experiment, the same presentation method was used, but the naming response was eliminated by providing the name of one of the pictures in advance and requiring subjects to simply report "yes" if the frame occurred around that picture and "no" if it did not. As in the earlier experiments, subjects were familiarized with the stimulus pictures in advance. Therefore, when the experimenter named a picture prior to sequence onset, the subject knew exactly what that picture would look like.

The picture specified by the experimenter was sometimes the host picture (the one actually in the frame) and sometimes another picture from the sequence; these included the previous four pictures and the following four pictures. These positions were selected because in previous research with pictures (Intraub, 1985b), subjects' errors tended to involve the −1 and +1 pictures and rarely included the −2/+2 pictures and beyond. If affirmative responses to −1/+1 pictures are due to a "yes" guessing bias, then subjects should show the same response rate to the "catch" trials (i.e., specification of −2/+2 and beyond). According to the integrative buffer model, the yes/no procedure should yield the same pattern of results as the unconstrained procedure: (a) a relatively high hit rate and (b) an error rate to −1 and +1 pictures that exceeds the error rate for the catch trials. This outcome would also eliminate naming errors as the source of the high-confidence conjunction errors during rapid presentation.

To further test the effects of naming, one group of subjects (the naming group) was required to follow a "no" decision with a naming response to indicate which object they thought the frame had actually been presented around. The other group (the no-naming group) simply made *yes* and *no* responses to indicate if the specified picture was presented with the frame. A difference between the frequency or pattern of frame migration between these two groups would show the effects of the naming requirement on performance.

Method

Subjects. The subjects were 36 undergraduate volunteers (25 women) from the University of Delaware, who were completing an optional research requirement for an introductory psychology course.

Apparatus. Subjects were seated approximately 1.7 m from a rear projection screen. The image was projected from an adjacent room using a Visual Instrumentation Corporation Selecta-Frame 5, variable-speed projector at silent speed (18 frames/s). The size of the field was 14 × 20 cm, which subtended a visual angle of approximately 5° × 8°. The experimenter was seated behind the subject and controlled the projector with a remote control box. All films were made using a Bolex H-16 16-mm movie camera in single-frame mode. A lens and extension tube attachment were used to photograph the stimuli, which were backlighted 35-mm slides.

Stimuli. The stimuli were 12 objects that were cut out of magazines and photographed on a gray background. These were the same stimuli that were used by Intraub (1985b, Experiments 2 and 3), except that three pictures were replaced by three others from the Intraub (1979) set (the hot air balloon, car, and tractor were replaced with pictures of a pie, a man, and a snowman). On the basis of normative data obtained by Intraub (1979), these 12 pictures were ones for which there was high naming agreement among subjects, as well as relatively low visual duration thresholds, suggesting that in some sense they were equally easy to see.

Ten pictures served as potential host pictures (suitcase, man, eyes, truck, glass, flag, snowman, projector, pie, and stove). Two pictures served as fillers (chair and organ). They were presented at the beginning of the sequence on those trials that the frame was presented in serial position 8 and at the end of the sequence on those trials that the target was presented in serial position 5.

Filmed sequences. The films were created using single-frame photography (2 frames of film per picture, presented at a rate of 18 frames per second). There were 100 sequences in the experiment. Each of the 10 stimulus pictures was the target 10 times, 5 times in serial position 5 and 5 times in serial position 8. For each target picture, there were two orders of picture presentation, one the reverse of the other, so that the same pictures that preceded the target in one sequence order followed the target in the other sequence order. Each of the 10 stimuli was specified (named by the experimenter prior to onset of a sequence) 10 times, twice when it was the target and one time when it was in each of the nontarget specification positions (i.e., −4, −3, −2, −1, 1, 2, 3, 4). Each picture appeared equally often in each of the nontarget specification positions. Twenty-two additional sequences were filmed to serve as practice sequences.

Design and procedure. Subjects were randomly assigned to either the naming group or the no-naming group and were individually tested. Subjects in both groups received the same familiarization procedure. This included (a) showing each picture for 4 s and providing the subject with a name for each, (b) presenting the subjects with six naming sequences in which they named each picture as quickly as possible and were immediately presented with the next picture, and (c) presenting each picture with the black frame around it for 4 s each so the subject was familiar with each potential conjunction. All subjects were told that the experimenter would specify a picture by name prior to each sequence and that their task was to immediately respond "yes" if that picture was the one in the frame and "no" if it was not, along with a confidence rating of "sure," "pretty sure," "not sure," or "guess." They were instructed to focus on the fixation point when the experimenter gave the ready warning and to maintain fixation for the 1.3-s duration of each sequence. The instructions to the two groups differed in that the naming group was asked to follow a "no" response with the name of the picture that they saw with the frame, along with a second confidence rating. An example of a "yes" and a "no" response in the naming group is "yes—pretty sure" and "no—sure—truck—pretty sure." For the no-naming group, the latter would be expressed "no—sure." All subjects were instructed to not wait until the end of the sequence, but to respond as soon as they saw the frame.

After being familiarized with the pictures and receiving instructions, subjects were presented with 22 practice sequences. After answering any questions, the experimenter then proceeded with the experiment. Following each third of the experiment, the subjects received a brief break and were presented with the pictures for a few seconds each. They were asked to quickly provide the name associated with each picture. Subjects showed no difficulty in quickly naming the pictures.

Results and Discussion

The pattern of results was strikingly similar to that obtained using the standard unconstrained naming procedure with objects (Intraub, 1985b) and with alphanumeric stimuli (Gathercole & Broadbent, 1984; Intraub, 1985b; Lawrence, 1971; McLean et al., 1983). Although the experimenter had provided the subjects with the name in advance, so that they knew which picture to look for, they frequently made high-confidence conjunction errors involving the −1 and +1 pictures. This occurred both in the primary yes/no task and in the secondary unconstrained naming task required of subjects in the naming group. The results for each task shall be addressed separately.

The yes/no task. The percentages of "sure," "pretty sure," "not sure," and "guess" confidence ratings in the naming group were 42%, 44%, 11%, and 2%, respectively, and in the no-naming group were 53%, 34%, 10%, 3%, and 0, respectively. The percentage of misses (times the subject did not see the frame) was 1% in the naming group, and misses never occurred in the no-naming group.

Table 1 shows the percentage of high-confidence responses ("sure" and "pretty sure" responses) that were "yes" responses as a function of the position of the specified picture. (Note that "yes" responses to anything but the host picture are

Table 1
Mean Percentage of High-Confidence ("Sure" and "Pretty Sure") Responses and "Sure" Responses That Were Also "Yes" Responses in Each of the Specification Conditions

Group	−4	−3	−2	1	0	+1	+2	+3	+4
	Position of the specified picture								
	High confidence (*ns* = 18)								
Naming	0	1	1	26	86	33	3	4	3
No naming	9	9	12	34	85	42	15	6	8
	Sure confidence								
Naming	0	0	0	15	85	18	3	1	2
n	16	16	14	13	13	13	15	16	17
No naming	7	7	6	23	79	26	10	3	11
n	17	17	17	15	17	14	16	16	18

errors.) A two-way mixed analysis of variance (ANOVA) (Group × Specification Position) was performed on the number of high-confidence "yes" responses. In the analysis, the symmetrical positions were collapsed, yielding the following five specification conditions: 0, −/+1, −/+2, −/+3, and −/+4. (Each subject had received 20 trials in each of these five conditions). The results yielded a highly significant decrease in the number of "yes" responses the further the specified picture was from the target, $F(4, 134) = 246.20$, $MS_e = 645.56$, $p < .001$. The difference between groups (naming vs. no-naming) was not significant, $F(1, 34) = 3.06$, $p < .10$; nor was there any interaction of group and position ($F < 1$).

It is important to note that the magnitude of the position effect is attributable to the frequent high-confidence "yes" responses to the 0 and −/+1 specification conditions. As was the case in the unconstrained procedure, subjects tended to report the frame around the −1, 0, and + 1 pictures. The possibility that migration errors to the −1/+1 pictures reflects a "yes" guessing bias was not supported. A Neuman-Keuls test showed that the number of "yes" responses did not differ among the −/+2, −/+3, and −/+4 conditions, but that there were significantly more "yes" responses in the −/+1 condition ($p < .01$) and a significantly greater number in the 0 condition ($p < .01$). If subjects' error responses reflected a "yes" guessing bias, one would not expect to find a difference in "yes" responses between the −/+1 and −/+2 positions and no difference among the positions that were −/+2 and further removed in the sequence.

An analysis of the confidence ratings, collapsing over groups, yields the same pattern of results. The percentages of high-confidence responses that were obtained, irrespective of whether the subject responded "yes" or "no," for the 0, −/+1, −/+2, −/+3, and −/+4 specification conditions were 82%, 78%, 89%, 91%, and 93%, respectively. A one-way, repeated-measures ANOVA showed that subjects were more confident of their responses the further removed the specified picture was from the host picture, $F(4, 140) = 16.75$, $MS_e = 13{,}230.00$, $p < .001$. A Neuman-Keuls test showed that the difference was due to a lower percentage of high-confidence responses in the 0 and −/+1 conditions than in the −/+2, −/+3, and −/+4 conditions ($p < .01$). The 0 and −/+1 conditions did not differ; nor did the −/+2, −/+3, and −/+4 conditions differ from one another.

In sum, when the host picture, or the −1/+1 pictures were specified, subjects sometimes rejected the host picture and sometimes accepted the −1 or +1 picture as the picture in the frame with high confidence. Most subjects rarely accepted specified pictures that were −2/+2 or further removed from the frame. As a temporal reference, recall that onset of a −2 picture precedes onset of the host picture by only 222 ms. The difficulty in discriminating among the host picture and the −1/+1 pictures is reflected also in the lower frequency of high-confidence ratings in those three conditions compared with cases in which the −/+2 pictures and beyond were specified.

Unconstrained responses. Recall that although both groups' primary task was the yes/no task, subjects in the naming group had the additional task of following a "no" response with the name of the picture that they had seen in the frame. Across the five specification conditions, there were a total of 555 "no" responses associated with high-confidence object reports. The reported object was the host picture 73% of the time and was the −1/+1 picture 21% of the time. Subjects reported a picture falling beyond the −1/+1 position on the remaining 6% of the trials. The results parallel those obtained with the standard unconstrained procedure. When the subjects rejected the picture that had been specified by the experimenter, they tended to report the frame around the host picture, the −1, or the +1 picture in the sequence.

Experiment 2

Experiment 1 demonstrated that prior warning regarding the potential conjunction of the frame with a specified object and elimination of the naming requirement did not eradicate the temporal migration effect. To determine if the effect reflects integration errors during the early stages of scene processing, Experiment 2 focused on the pattern of errors that has been repeatedly obtained. Of particular interest was whether the direction of migration (preceding vs. following picture) is a result of on-line integration processes or later errors in memory.

According to the working model described earlier, if the pattern of errors reflects ongoing integration processes, direction of migration may depend in part on how soon processing of the frame is initiated relative to its host picture. In those trials in which the subject rapidly begins to process the frame, it may become integrated with the previous picture, which is still being processed in the very short-term buffer. In those trials in which processing of the frame is slightly delayed, it may become integrated with processing of the following picture in the sequence. Thus direction of migration would be expected to be diagnostic of rapid versus delayed initiation of processing of the frame. This assumption runs counter to the memory hypothesis.

According to the memory hypothesis, direction of migration is not an important feature of the phenomenon. Subjects are reporting which of two or three temporally adjacent pictures (now in memory, without strong temporal order tags) had been presented simultaneously with the frame. Error responses in this case would randomly involve an earlier picture on some trials and a later picture on others. The error, in this view, is not related to on-line processing, but to a review of events in memory. This position is consistent with some subjective reports in which there is a subjective sense of having to "piece together" the events in memory and determine the conjunction.

The rationale adopted in Experiment 2 was to use the same unconstrained report procedure as in previous research and to include a second dependent variable—one that would provide an assessment of processing onset. This was accomplished through the addition of a reaction time detection task. Subjects were required to (a) press a key as soon as they saw the frame, (b) report the object the frame had been presented with, and (c) provide a confidence rating.

A clear-cut prediction of the integrative buffer model is that given high confidence, frame detection should be faster when subjects make a −1 migration error than a +1 migration error. In addition, according to the model, when the subject correctly reports the frame around the host picture, frame detection should not be faster than in the −1 condition or slower than in the +1 condition. The latter prediction is as important as the former. It runs counter to a plausible alternate prediction. According to the alternate view, correct responses would be expected to yield the fastest reaction times because trials on which subjects detect the correct conjunction are ones in which the frame has been seen most clearly and unequivocally. The errors occur on trials where the frame is not well perceived.

According to the post hoc memory hypothesis, frame detection time should not differ as a function of the direction of migration. Detection time during presentation would have no bearing on the direction of migration because the error is made at a later stage of processing, not during perceptual integration. This view would predict either equivalent reaction times for all three responses types (−1, +1, and 0) or faster times to correct responses (as described earlier) with equivalent slower times for −1 and +1 responses. Considering these two possible response patterns, the latter would be more informative, because no difference in reaction times at all could be interpreted either as supporting the post hoc memory view or as indicating that the reaction time task is not sensitive enough to measure differences in processing onset.

Method

Subjects. Subjects were 27 undergraduate volunteers (21 women) from an introductory psychology course who were paid $3.00 each for their participation.

Apparatus. The apparatus was the same as in Experiment 1, except that reaction times were measured in the following way. An interface was made between the control box of the projector (which contained a digital frame counter) and an Apple II Plus computer. The experimenter started each sequence from frame 1. The clock started when the experimenter switched the projector into forward cine mode and was stopped when the subject depressed a hand-held response key. The time between frame 1 and the frame containing the target picture (i.e., the picture with the frame), was subtracted from the total time so that the time from the onset of the target picture to the key press was determined. Reaction times were accurate to 1 ms.

Stimuli. The stimuli were the same 12 objects used by Intraub (1985b). These included a car, a hot air balloon, a suitcase, an organ, a chair, a tractor, a goblet, an American flag, a stove, a pair of eyes, a movie projector, and a truck.

Filmed sequences. Each sequence contained all 12 pictures, with one of the 12 photographed with the black frame around it. Each picture was presented in the frame six times, yielding 72 sequences. On the six occasions that a picture was the host picture, it appeared with three different pairs of flanking pictures (i.e., −1 and +1 pictures), such that on one occasion, the order was ABC and on the other it was CBA (B is the host picture). Twelve additional sequences (in which each picture was the host picture once) were included, so that the film contained 84 sequences. The film was presented twice to subjects, yielding 168 experimental trials.

Procedure. The procedure was basically the same as the unconstrained procedure used by Intraub (1985b). The only difference was that subjects were instructed to press a hand-held key as soon as they saw the frame. They were told to follow the key press with a report of the picture the frame had appeared around and their confidence in this response ("sure," "pretty sure," "not sure," or "guess"). As in the previous experiment, subjects were familiarized with the pictures in advance and received naming practice. They then received 24 practice sequences followed by the 168 experimental sequences. After approximately every 28 sequences, subjects received a brief break filled with a naming trial in which they rapidly named each picture in the sequence.

Results

The pattern of reaction times predicted by the model was obtained, supporting the hypothesis that the subject's experience regarding the location of the frame is predicated, at least in part, on how soon frame processing is initiated during presentation. The results shall be presented in the following order. First the temporal migration data are discussed, because it is important to show that the basic phenomenon was replicated using this slightly modified unconstrained procedure. Then the reaction times associated with these reports are presented.

Detecting the conjunction of the picture and the frame. The subjects' accuracy in detecting the conjunction of the picture and the frame as a function of confidence rating is presented in Table 2. Addition of the reaction time task did not alter the pattern of results that had been obtained with the same 12 pictures, using the large frame in Intraub's experiments (1985b, Experiments 2 and 3). All subjects showed the migration effect, and all but one had their highest two error rates in the −1 and +1 positions (the "odd" subject showed the highest error rate to −1 pictures with relatively few +1 errors).

Reaction time. The mean reaction time associated with correct reports and with +1 and −1 error reports under conditions of high confidence was calculated for each subject. Any reaction times that were faster than 200 ms or longer than 2.5 standard deviations from each subject's grand mean were excluded from the analysis. Of the total number of responses, 12% were excluded because either they were misses (i.e., subject did not see the frame) or the subject failed to press the key, 1% were excluded because the reaction times were too long, and 4% were excluded because the reaction

Table 2
Percentages of Responses Reporting the Conjunction of the Picture and the Frame as a Function of the Picture's Position and the Confidence Rating

Confidence	%T[b]	Position of the reported picture[a] −3	−2	−1	0	+1	+2	+3
"Sure"	12	0	1	8	81	9	2	3
"Pretty sure"	46	2	2	19	58	15	1	3
"Not sure"	27	7	6	23	36	18	3	6
"Guess"	3	12	7	20	28	15	6	13
Misses	12							

Note. The −3/+3 columns contain responses to the −3/+3 positions and further (e.g., −4/+4).
[a] Host picture = 0, preceding pictures = −, and following pictures = +.
[b] Percentage of total responses falling into each confidence category.

times were too short. The average reaction time window was 200–567 ms. The overall mean reaction time (mean of subjects' grand means) of the acceptable responses was 337 ms ($SD = 42$ ms).

Mean reaction times as a function of the reported position of the frame are shown in Table 3. The percentages of high-confidence −1, 0, and +1 responses were 17%, 63%, and 13%, respectively. As may be seen in the table, the reaction times follow the prediction of the integrative buffer model. A one-way repeated-measures ANOVA showed a significant effect of position, with the fastest reaction time occurring with the −1 reports and the slowest occurring with the +1 reports, $F(2, 52) = 19.35$, $MS_e = 269.08$, $p < .0001$. The critical comparison between times associated with −1 and +1 reports was highly significant, $t(26) = 5.33$, $p < .005$. The difference in reaction time was also significant by sign test ($p < .01$), with 22 of 27 subjects responding more quickly when the frame was reported around the preceding picture in the sequence than around the following picture.

The number of reaction times contributing to each mean varied with subject and condition. For the −1 condition, the number of contributing scores ranged from 2 to 28, with a median score of 10. For the +1 condition, the range was 4 to 38, with a median score of 13. When subjects were correct, the range was 16 to 104, with a median score of 54. Because subjects sometimes had a large imbalance in the number of contributing scores in the −1 and +1 conditions, a second analysis was run including only those subjects who had at least 9 reaction times contributing to each mean. This subset included approximately one half of the subjects ($n = 13$). The mean of the means for this subset is also shown in Table 3. The more stable scores showed the same pattern. A one-way repeated-measures ANOVA yielded a significant main effect of position, $F(2, 24) = 8.56$, $MS_e = 227.01$, $p < .002$. The critical comparison between the −1 and +1 conditions was significant, both by t test, $t(12) = 3.50$, $p < .005$, and by sign test ($p < .05$), with 11 of 13 subjects showing the effect.

Picture analysis. The migration effect was obtained when each of the 12 pictures was the host picture. The hit rate ranged from 30% to 89% (median = 65%, $M = 61\%$, $SD = 19\%$). For all host pictures, the characteristic migration pattern was obtained in that the largest number of errors were reported to the −1 and +1 pictures in the sequence.

Table 3
Mean Frame Detection Times Associated With High-Confidence Hits (0) and High-Confidence −1/+1 Errors in Experiment 2

Subjects	Position of reported picture −1	0	+1
Total ($n = 27$)			
M	327	332	353
SD	42	40	46
Subset ($n = 13$)			
M	342	343	364
SD	48	45	52

Discussion

As predicted by the integration model, in those trials in which a −1 migration error occurred, frame detection time was faster than on those trials on which a +1 migration error occurred. This pattern of differential response times was upheld across subjects. Also consistent with the model were the reaction times to correct conjunction responses (no migration), which were no faster than the response times associated with −1 migration errors.

In judging the psychological significance of the 26-ms difference between the −1 and +1 conditions, it is important to recall that the duration of the display itself (host picture and frame) was only 111 ms. Proportionately speaking, the magnitude is large. It is not as large, however, as would be expected if processing of the picture and the frame were conducted serially. According to a serial view, the frame would be processed either before the host picture (in which case it might migrate to the −1 picture) or after the host picture (in which case it might migrate to the +1 picture). If this were the case one would expect to find a larger frame detection difference (perhaps closer to 100 ms) because of the estimated time it takes to process a picture.

The model, however, does not posit a simple serial view. Instead, the assumption is that several types of visual and conceptual processes are conducted in parallel. Experiment 2 indicates that as these ongoing processes take place, differences in frame detection time can influence the direction of a temporal migration error. In terms of the present model, although subjects usually integrate the frame with its host picture (the majority of high-confidence responses are correct), −1 errors are more likely to occur when detection of the frame is relatively rapid, in which case frame onset would be associated with the final stages of processing of the previous picture (which is still in the buffer), and +1 errors are more likely to occur when detection of the frame is relatively delayed, in which case frame onset becomes associated with the early stages of processing of the next picture in the sequence.

The process that determines frame detection time is not clear. One likely candidate is attention. That is, attention may be directed either to the frame first or to the host picture first on a given trial. If we consider that attention is under the control of a subject's strategy, then we would expect that tasks that bias the subject to attend to the frame first or to the host picture first would affect direction of migration. Interestingly, the present task—one requiring a frame detection reaction time response prior to the description of location—did not yield results with a greater negative shift than previous similar experiments without the reaction time task (Intraub, 1985b, Experiments 2 and 3). In a recent experiment in which a gap was placed in the frame and subjects were required to indicate gap orientation (up, down, left, or right) as well as which picture the frame had appeared on, no shift in the direction of migration was obtained, compared with a control condition in which the subjects saw the same sequences but reported the picture in the frame without having to report gap orientation. Another comparison that might have revealed a shift in the direction of migration is the comparison between the unconstrained and yes/no tasks, because the latter actually specifies a picture in advance and thus might have resulted in a processing bias favoring the host picture (in trials in which the host picture was specified), as opposed to the former, which explicitly requires the subject to search for the frame and report the picture. A direct comparison of these conditions for the stimuli in the present Experiments 3 and 4 did not yield such a shift.

It may be that none of these tasks vary the salience of the frame or bias the subject to process it first. The task requirements may not have been sufficient to overshadow the fact that subjects are ultimately required to report the conjunction of the frame and the picture and that this may take precedence even when the conjunction report is the secondary task in the instruction. It may be that attention is an important factor in the direction of migration and that these tasks have not been successful in altering the allocation of attention; or it may be that the relatively symmetrical pattern of +1 and −1 errors reflects varying sensitivity to the target frame that randomly changes from trial to trial. Should the latter be the case, the basic hypothesis of the model would be unchanged regarding the reason for the migration; that is, migration is most likely to occur when the identification/integration time of the display is long relative to the presentation rate (see Assumption 2 in the Introduction).

It should be noted that Gathercole and Broadbent (1984) have reported slight shifts in the frequency of positive and negative frame migration, depending on how a target item (a digit or a word) was specified (identity or category). Future research will be directed at clarifying this controversy. In terms of the present argument, however, regardless of whether detection time proves to be a randomly fluctuating factor or to be dictated by the subject's attentional strategy, Experiment 2 shows that on-line frame detection time is strongly associated with the direction of migration. This supports a major assumption of the developing integration model.

Experiment 3

The purpose of Experiment 3 was to determine if a migration effect similar to that obtained with objects and unrelated frames would occur for meaningful objects presented in real-world scenes. The reason for this question is that expected contextual relations of various types may or may not influence integration of visual information. Previous temporal migration studies in which subjects were required to report the conjunction of two configurations have used unrelated, discrete stimuli (e.g., Gathercole & Broadbent, 1984; Intraub, 1985b; Sperling & Reeves, 1980; Experiments 1 and 2 of the present study). For example, subjects have been required to monitor (a) concurrently presented streams of alphanumerics, (b) frames and alphanumerics or words, and (c) frames and objects. The purpose of Experiment 3 was to determine if migration would occur under conditions in which the display is made up of interrelated components, rather than discrete, unrelated ones such as these.

Scenes are particularly interesting because of numerous experiments demonstrating the rapidity of scene comprehension in single tachistoscopic flashes of 100 ms and less (Biederman et al., 1974; Biederman et al., 1982) and during high-speed sequential search (Intraub, 1981a, 1981b; Potter, 1975, 1976). As discussed earlier, no evidence of dissociation or faulty integration of components of successively presented photographs of objects was obtained in search tasks (Intraub, 1981a, 1981b). It was only when an unrelated stimulus (the frame) was added to the display that a striking migration effect emerged.

In Experiment 3, subjects viewed outline drawings of scenes that were presented at a rate of 9/s and searched for a specified object (e.g., a bull or a chair). The object always appeared in the same location on the screen. Subjects were required to report the scene the object appeared in as soon as they saw it. All potential host scenes appeared in each sequence, although only one scene per sequence contained the target object.

Method

Subjects. Subjects were 20 undergraduate volunteers (13 women) from the University of Delaware, enrolled in a laboratory course in cognition. Subjects in the additional coat rack/tree condition were 7 undergraduates from the introductory psychology subject pool at the same university.

Materials. The stimuli were 12 outline drawings of indoor and outdoor background scenes. Initial drawings were tested using a visual duration threshold procedure, and a procedure in which pictures were presented (masked) for 100 ms and had to be identified. Changes in the drawings were made to make them easier to identify, and the cycle was repeated several times. Visual duration thresholds for the final set were more than 100 ms for all pictures; however, once subjects were familiarized with the pictures, all pictures could be identified when presented for 100 ms, preceded and followed by a mask that was made up of unidentifiable pieces of other outline drawings. All subjects agreed on the identity of each scene, providing the names listed below or synonyms.

Six of the scenes (the pasture, the barnyard, the bullfight arena, the bedroom, the living room, and the study) served as host pictures in the experiment (i.e., scenes in which the target could appear). The remaining six (the playground, the front yard of a house, the garden, the office, the waiting room, and the hallway) never served as host pictures or in the two positions surrounding the host picture but were always presented in the sequence.

The pictures were drawn in black India ink on white paper. The target objects (a bull and an easy chair) were drawn with the same pen and ink on acetate so that they could be copied onto the scenes using a 3M copier machine. The acetate drawings were filled in with a white paint so that they would copy as solid objects, occluding the background. An example of one of the scenes, with and without a target object, is shown in Figure 1. Copies of the scenes, with and without objects, were photographed on 35-mm slides. The slides were then backlighted and photographed on 16-mm movie film using single-frame photography. The camera was a Bolex H-16 16-mm movie camera with an extension tube. Each picture was photographed on two frames of movie film.

Apparatus. The same apparatus, display size, and film speed were used as in Experiments 1 and 2.

Design. Each individually tested subject took part in two conditions—one in which the bull was the target and one in which the chair was the target. Order of presentation was counterbalanced across subjects. Each sequence began with a fixation point (a small triangle in the center of the field). After a verbal warning to the subject, a double row of Xs was flashed on the screen for approximately ¼ s. This was followed by a repetition of the fixation point for approximately ½ s and the 12 pictures which were presented at a rate of 9/s. The fixation point reappeared immediately following the last picture of each sequence.

There were 72 sequences in each condition. In a given condition, each of the six potential host scenes was presented with the target object 12 times (4 times in each of three serial positions: 4, 6, or 8). When the target was presented in one of the three outdoor pictures (the barnyard, the pasture, or the bullfight arena), the other two outdoor pictures were used to fill the immediately preceding and following positions in the sequence. The order of these flanking scenes was alternated so that each one preceded the target twice and followed the target twice at each of the three serial positions (the two orders can be described as ABC and CBA, with B referring to the target picture). The other three host pictures (in this case the bedroom, the living room, and the study) were randomly mixed with the six filler pictures with the constraint that at least 1 picture would be presented somewhere in the sequence prior to the target and one would be presented somewhere in the sequence, following the target. On approximately 60% of the trials in each condition, at least one of the three pictures appeared in the −2 or +2 position. The same constraints applied when the target appeared in one of the indoor scenes.

Since each object was presented in all six potential host scenes, this design ensured that the −1 and +1 pictures were high-probability scenes when the target was in a high-probability host scene (e.g., chair in indoor scene) and were low-probability scenes when the target was in a low-probability host scene (e.g., chair in outdoor scene). This was done to eliminate response bias in those positions that might occur if the target was in a low-probability scene and was flanked by high-probability scenes in the −1 and +1 positions, and vice versa.

Procedure. Subjects were tested individually. To familiarize them with the stimuli, subjects were presented with the 12 outline drawings for about 5 s each. The experimenter provided a name for each scene, but told the subject that if he or she felt more comfortable with another label, to use it. The subject then practiced naming the scenes. This was accomplished by presenting the pictures one at a time and advancing to the next picture as soon as the subject provided a name. Four naming sequences were used. By the end of the trials, subjects could name the pictures rapidly with no errors.

To familiarize themselves with the rapid presentation rate, subjects then viewed sequences of the scenes once at each of the following presentation rates: 4/s, 6/s, and finally 9/s—the rate used in the experiment. The final step in the familiarization procedure was to show the subject the target (bull or chair) alone on a white background, and to show it in each of the six host scenes. It should be noted that subjects were not explicitly told that these were the only scenes that the object would appear in. They were told "I will show you what the bull [chair] looks like. Then, I will show you what it looks like in some of the scenes." They were told not to worry about the number of times they reported a scene and were warned that some scenes may actually never be presented with the object.

After the familiarization procedures, subjects were instructed to search for the object in each sequence and to immediately report the name of the scene in which it appeared. They were told to not wait until the end of the sequence, but to respond as soon as they saw the target object and to follow each response with a confidence rating of "sure," "pretty sure," "not sure," or "guess." After the first and second thirds of each condition, subjects took part in a naming trial including all 12 scenes (without the target), to be sure that they had not become confused about the names of the scenes, as well as to provide a rest interval. Subjects had no difficulty rapidly naming each scene. Subjects were presented with 12 practice sequences and 72 experimental sequences.

After completing one condition, subjects were then familiarized with the new target and were presented with 12 practice sequences and the 72 experimental sequences of that condition. The experiment required approximately 40–45 min.

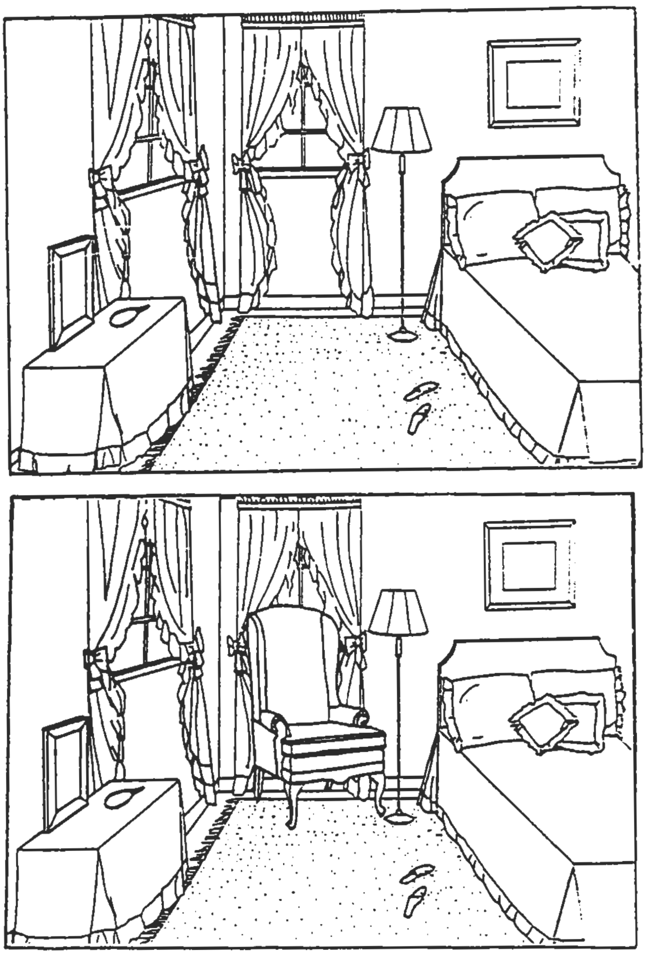

Figure 1. The bedroom scene with and without the chair, used in Experiments 3 and 4.

Results and Discussion

Although the scene and the target object were presented simultaneously, they were often not interpreted as a single visual event. The pattern of errors obtained with the objects and scenes was the same as that obtained with the objects and unrelated frames in Experiments 1 and 2 and in the Intraub

(1985b) study. Accuracy in detecting the correct conjunction of the object and its background scene as a function of confidence rating and target object is shown in Table 4. The percentages of "sure," "pretty sure," "not sure," and "guess" responses were 10%, 46%, 32%, and 6%, respectively, in the bull condition and 10%, 48%, 31%, and 6%, respectively, in the chair condition. Subjects were unwilling to guess, claiming, for example, a lapse in attention, on 6% and 5%, respectively, of the trials in each condition.

All 20 subjects frequently reported the object in the wrong scene. Once again, when subjects expressed high confidence ("sure" and "pretty sure" ratings), their errors primarily involved the immediately preceding or following outline scene in the sequence. Contrary to a simple guessing hypothesis, subjects rarely reported a picture that was further away than the −1 or +1 position in the sequence, even though all six potential targets were presented in each sequence. When they did report a peripheral picture, they sometimes reported potential target and sometimes reported filler pictures.

Consistent with this pattern of results, subjects who remained after the experiment and took part in some additional trials with feedback were surprised when they were informed that a high-confidence response was wrong. For example, a typical response would be "I thought I saw the frilly curtain separately and saw the chair next to the bookcase" in a case where the chair actually was presented in the bedroom in front of the frilly curtain, and the study (the scene with the bookcase) was in the −1 or +1 position.

Object migration as a function of individual host scenes. The data presented to this point reflect the subjects' accuracy in detecting the conjunction of objects and scenes, collapsing over the six different scenes. It is important to determine if the migration effect occurs for all the pictures or if some pictures are always correctly reported and some never correctly reported. The migration effect was obtained for each host scene. Although hit rates varied, the general pattern of results was clearly upheld. The hit rates for the six scenes for high-confidence responses were 45%, 69%, and 30% for the indoor scenes and 69%, 39%, and 67% for the outdoor scenes. An ANOVA showed no main effect of object or interaction of object with indoor versus outdoor scenes ($F < 1$ for both).

Replication with new objects. The same procedure was conducted using six different host pictures and two different target objects. The stimuli were the same 12 pictures as in the bull-chair experiment, except that the six potential host pictures in this experiment were the six pictures that were never host pictures in the previous experiment. They were the waiting room, office, hallway, playground, front yard, and garden. The remaining six pictures served as fillers. The target objects were a coat rack and a tree that, like the bull and chair, were drawn to be approximately the same size. Seven new subjects took part in the experiment. The design and procedure were identical to the bull-chair experiment in all other respects.

Given a high confidence rating, subjects reported the correct scene 43% and 47% of the time for the tree and coat rack, respectively. They reported the −1 scene on 21% of the trials for each object and the +1 scene on 33% and 30% of the trials, respectively. A notable outcome obtained in the replication argues further against a guessing hypothesis. Although the subjects reportedly named all six host pictures during naming trials and saw the targets in each, one picture (the house front) was very difficult to see during rapid presentation. Although it was one of the three potential outside host scenes, it was very rarely reported by the subjects (less than 3% of the high-confidence trials). Subjects apparently report what they see in the sequence. They were guessing from the small set of potential host scenes.

Table 4
Percentages of Responses Reporting the Object in the Correct Scene (0), Preceding Scenes (−), or Following Scenes (+) as a Function of Confidence Level

Confidence	Position of the reported scene in the sequence −3	−2	−1	0	+1	+2	+3
Bull target							
"Sure"	0	0	13	74	13	0	0
"Pretty sure"	0	0	23	51	25	0	0
"Not sure"	1	2	25	41	30	0	0
"Guess"	5	5	19	40	25	2	5
High confidence[a]	0	0	21	57	21	0	1
Chair target							
"Sure"	0	0	15	64	21	0	0
"Pretty sure"	0	0	20	52	27	0	0
"Not sure"	4	1	19	40	32	1	2
"Guess"	2	1	21	32	41	1	1
High confidence[a]	0	1	20	52	25	0	1

Note. The −3 and +3 positions contain all responses made to scenes that were three or more pictures removed from the host picture.
[a] Composed of the "sure" and "pretty sure" responses.

Experiment 4

In Experiment 4, the yes/no procedure used in Experiment 1 (naming condition) was applied to the object/scene sequences to test object migration under conditions in which the scene was specified in advance by the experimenter.

Method

Subjects. The subjects were 32 undergraduates (15 men) from the University of Delaware, taking part in the optional introductory psychology subject pool. One half of the subjects searched for the bull target and the remaining subjects searched for the chair target.

Stimuli. The stimuli were the same 12 scenes and two target objects described in Experiment 3. The same two films were used—one in which the bull was the target object and one in which the chair was the target object.

Design. The 72 sequences in each film were organized in the following way. The host picture was specified in one half of the sequences, with each of the six pictures specified six times (three times in the ABC pattern and three times in the CBA pattern). In the remaining 36 sequences either the −1 picture or the +1 picture was specified (each of the potential host pictures was specified when it was in the −1 position three times and when it was in the +1 condition three times). Two different orders of specification were created for each film so that those sequences that had the host picture specified for half the subjects had the −1 or +1 pictures specified for the other half and vice versa.

In both films, 12 additional sequences were presented to serve as "catch trials" for the "yes" responses. In these, the specified scene was in the −3 or −2 flanking positions (−F) on one half of the trials and in the +3 or +2 flanking positions (+F) on the other half. (Sequence constraints prevented equal representation of each peripheral position. Across the entire experiment, specification of the −3, −2, +3, and +2 pictures occurred 47%, 4%, 18%, and 31% of the time, respectively.) Each of the six scenes was specified two times in these peripheral positions. Those pictures that were specified in −F positions in one order of specification were specified in +F positions in the other order. In each order, specifications were assigned so that no more than three host specifications (true "yes" responses) or three nonhost specifications (true "no" responses) could appear consecutively.

Procedure. The familiarization procedure was the same as that used for these stimuli in Experiment 3. The yes/no report procedure was the same as in the naming condition in Experiment 1. Once again subjects were frequently reminded to fixate their eyes on the center of the screen. The yes/no procedure took about twice as long as the standard procedure used in Experiment 3. Subjects were only tested in one condition (bull target or chair target). The experiment took approximately 45 min.

Results and Discussion

Subjects persisted in reporting erroneous conjunctions of objects and scenes under conditions in which the primary task was a simple "yes" or "no" response. The results mirrored those obtained with the objects and frame in Experiment 1. Collapsing over specification conditions, the percentages of responses at each confidence level from "sure" to "guess" were 26%, 47%, 20%, and 4%, respectively. On average, subjects were unwilling to guess, claiming, for example, a lapse of attention on 4% of the trials. Table 5 shows the percentage of high-confidence responses in each specification condition that were "yes" responses. All 32 subjects showed the same pattern of responses. Once again, subjects rarely responded "yes" to specified pictures in the −2/+2 positions and beyond. In all specification conditions, when subjects responded "no" they tended to correctly report the host picture or report the immediately preceding or following picture in the sequence.

Object migration as a function of individual host scenes. Another important similarity to the results of Experiment 1 is that the same pattern of results was obtained for each of the six scenes when it contained the object. Collapsing over the host-specified and −1/+1-specified conditions, the high-confidence hit rates for the three indoor scenes were 49%, 54%, and 57%; the hit rates for the three outdoor scenes were 48%, 69%, and 79%. Once again, although hit rates vary, the migration effect occurred for each scene when it contained the object.

Table 5
Percentage of High-Confidence "Yes" Responses to Each Specified Scene as a Function of Its Position in the Sequence for Each Target Object

	Position of the specified scene				
Target	−F	−1	0	+1	+F
Bull	2	41	75	48	11
Chair	0	36	82	44	0

Context and object migration. The results of Experiments 3 and 4 clearly demonstrate object migration. This observation is very encouraging in that it raises the possibility that object migration will provide a new way to explore the organization of complex scenes. Biederman and his colleagues (Biederman, 1981; Biederman et al., 1982) have suggested a set of five relational rules that determine a well-formed scene. These rules include expectations about the size of an object in a scene, the probability that it would normally occur in such a scene, the likelihood that one object will occlude parts of another if placed in front of it, the need for physical support, and the likelihood of objects occurring in particular locations in a scene. It has been demonstrated that if an object violates one or more of these relational expectations in a briefly presented scene (e.g., 150 ms) that is preceded and followed by a visual noise mask, subjects take longer to identify it and make more errors in identification (Biederman, 1981; Biederman et al., 1982).

An interesting hypothesis derived from the integration model is that integration time may also be affected by the general knowledge and expectation inherent in these rules. For example, with respect to single objects, Intraub (1985b) suggested that the unrelated frame may take longer to integrate with a face than a pair of eyes would. As a result, all things being equal (presentation rate, picture order, etc.), the frame would migrate to temporally adjacent scenes more frequently than the eyes would. In terms of scene perception, following the same logic, if any or all of the relational rules described earlier play a role during the integration phase, then objects undergoing a violation of one or more rules should migrate more frequently than when it does not violate the rule(s).

Although Experiments 3 and 4 clearly demonstrated temporal migration of objects, no context effect was obtained (i.e., the bull did not migrate more frequently among indoor scenes than outdoor scenes and vice versa for the chair). This is not surprising in view of a number of mitigating factors that caused these stimuli to fall short of providing an adequate test of the effects of context on object migration.

The target object was always presented in the same spatial location on the screen, and the subject was familiarized with the targets in each of the potential host scenes. These factors were necessary to control in the present experiments, because their primary purpose was to determine if object migration would occur. To test this it was necessary to use the same procedures as in the "frame" migration research. In terms of contextual issues, however, these factors would be expected to reduce the natural differences in expectancy inherent in the various object/scene combinations.

More important, the outline scenes seemed to fragment at high speed and did not clearly impart their gist. For example, it seemed that rather than picking up the whole bedroom when claiming to see the chair in that scene, subjects would see the chair against the frilly curtain and, knowing that the frilly curtain was in the bedroom scene. would report that scene. While subjects reported integration of nonsimultaneous target objects and background objects, the gist of the scene as a bedroom or a study may not have been available. Both the lack of texture and shading in the simple outline drawings (which might have facilitated identification and integration of the background) and the predictable location of object placement (which may have focused spatial attention too tightly within the central portion of the scene) will be eliminated in future research. We are currently developing stimulus sets composed of photographs of staged scenes to determine if high-level relational rules play a role early enough in scene processing to affect on-line integration.

General Discussion

Temporal migration refers to a situation in which concurrently presented visual information dissociates, and components of one display become integrated with components of preceding or following displays in a rapidly presented sequence (Gathercole & Broadbent, 1984; Intraub, 1985b; Lawrence, 1971; McLean et al., 1983). This phenomenon raises very interesting issues about the role of attention and memory in visual perception. The present four experiments provide additional support for a working model proposed by Intraub (1985b) in which the key element is a very short-term integrative buffer.

Three major findings shall be summarized and discussed. First, it was found that a yes/no procedure yields the same pattern of migration errors as the standard unconstrained naming procedure typically used in temporal migration research (Experiments 1 and 4). This shows that the migration effect is remarkably robust. The yes/no procedure eliminates naming as a response, and because a potential host picture is named in advance, the subject has the benefit of expectancy. He or she must search for that particular familiar picture and indicate "yes" or "no" whether that picture is the one in the frame. The pattern of high-confidence errors shows that the target element (the frame or the outline object, depending on experiment) frequently became integrated with the −1 and +1 pictures but was rarely reported as occurring with any specified picture falling beyond the −1 and +1 positions. Previous recognition memory research with the colored objects used in Experiments 1 and 2, and with colored photographs of scenes, has provided independent evidence that the very short-term buffer can hold more than one picture at a time (Intraub, 1985a). The −1 and +1 pictures would be the ones most likely to be undergoing integration in the very short-term buffer at the time that the target element and its host picture are presented.

The second major finding was that temporal migration errors are apparently the result of on-line integration errors in the short-term buffer. Consistent with the prediction of the model, a reaction time detection task in conjunction with the standard unconstrained naming task showed that frame detection time is significantly faster when subjects make a −1 error than a +1 error (Experiment 2). This is consistent with the assumption that when processing of the target element occurs relatively early, it may become integrated with the final stages of integration associated with the previous picture in the sequence. When processing of the target element occurs relatively late, it may become integrated with the initial stages of integration associated with the following picture in the sequence. It is interesting to note that conscious knowledge of serial order of rapidly presented visual information is nonveridical (e.g., Reeves & Sperling, 1986) and that serial order information for pictures in particular is poor (cf. Paivio, 1971). Informal experimentation in our lab with experienced experimenters as subjects supports this position. Although subjects may not be consciously aware of which is a −1 or a +1 picture, their reaction times show that the rapidity with which they can detect the frame is strongly associated with which flanking picture is integrated with the frame when a migration error occurs.

Finally, the results of Experiments 3 and 4 show that the migration effect is not limited to discrete unrelated stimuli but can be obtained with interrelated objects and scenes. Subjects frequently reported target objects in temporally adjacent scenes in the sequence. For example, although on a given trial the chair might have been in the bedroom, occluding part of the curtains in the background, a subject might report the chair to be in the study, occluding part of the bookcase. Future research will determine whether the degrees to which an object is related to its host scene will affect migration in predictable ways. This raises interesting possibilities for research on the decomposition of component parts of meaningful scenes.

References

Avons, S. E., & Phillips, W. A. (1980). Visualization and memorization as a function of display time and poststimulus processing time. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 407–420.

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kobovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 213–253). Hillsdale, NJ: Erlbaum.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14,* 143–177.

Biederman, I., Rabinowitz, J., Glass, A., & Stacy, E. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology, 103,* 594–600.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language: An introduction to psycholinguistics and generative grammar.* New York: McGraw-Hill.

Gathercole, S. E., & Broadbent, D. E. (1984). Combining attributes in specified and categorized target search: Further evidence for strategic differences. *Memory and Cognition, 12,* 329–337.

Intraub, H. (1979). The role of implicit naming in pictorial encoding. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 78–87.

Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 1–12.

Intraub, H. (1981a). Identification and processing of briefly glimpsed visual scenes. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 181–190). Hillsdale, NJ: Erlbaum.

Intraub, H. (1981b). Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 604–610.

Intraub, H. (1984). Conceptual masking: The effects of subsequent visual events on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 115–125.

Intraub, H. (1985a). Conceptual masking of briefly glimpsed photographs. In R. Groner, G. W. McConkie, & C. Menz (Eds.), *Eye movements and psychological processes* (pp. 59–69). Amsterdam: North Holland.

Intraub, H. (1985b). Visual dissociation: An illusory conjunction of pictures and forms. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 431–442.

James, W. (1890). *Principles of psychology* (Vol. 1, Chap. 11). New York: Holt.

Lawrence, D. H. (1971). Two studies of visual search for word targets with controlled rates of presentation. *Perception and Psychophysics, 10,* 85–89.

Loftus, G. R., & Ginn, M. (1984). Perceptual and conceptual masking of pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 435–441.

McLean, J. P., Broadbent, D. E., & Broadbent, M. H. P. (1983). Combining attributes in rapid serial visual presentation tasks. *Quarterly Journal of Experimental Psychology, 35,* 171–186.

Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart & Winston.

Potter, M. C. (1975). Meaning in visual search. *Science, 187,* 965–966.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 509–522.

Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review, 93,* 180–206.

Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Sperling, G., & Reeves, A. (1980). Measuring the reaction time of a shift of visual attention. In R. Nickerson (Ed.), *Attention and performance VIII* (pp. 347–360). Hillsdale, NJ: Erlbaum.

Received August 21, 1987
Revision received February 9, 1988
Accepted February 15, 1988 ■