Visual Dissociation: An Illusory Conjunction of Pictures and Forms

Helene Intrabu
University of Delaware

Undergraduates viewed rapidly presented series of color photographs (9/s) and were required to indicate which photograph appeared within a black outline rectangle (the "frame"). Experiment 1 demonstrated that subjects were often confident and wrong, reporting the immediately preceding or following picture in the sequence. Experiment 2 showed that migration of the frame to other pictures cannot be attributed to spatial separation, because the same effect occurred when a small frame was presented in the center of the picture itself. Experiment 3 ruled out masking of the "framed" picture as the cause of the illusion by showing that the framed picture is indeed identified on those trials where the frame appears to be elsewhere. Experiment 4 showed that when simpler, more familiar stimuli (numbers) were presented, a more rapid presentation rate (18/s) was required to obtain the effect. It is proposed that the illusion reflects the action of integrative processes in a very short-term buffer and that it may provide a new tool with which to study the integration of features within scenes.

Several theories have addressed the role of cognition in visual perception. According to one model, features such as lines, angles, and colors are registered by the visual system, and these features are then integrated into our perceptual experience by a process that makes reference to our expectations about the visual environment (Treisman & Gelade, 1980; Treisman, Sylves, & Gelade, 1977). If this is so, these conditions should exist under which the integrative process can be disrupted, resulting in a faulty integration of features. For example, consider the anecdotal report of a viewer who thought he glimpsed a bald-headed, bearded friend, but on further scrutiny he saw that the bald head belonged to one man and the beard to another (Treisman & Schmidt, 1982). Anecdotal errors such as these are certainly rare and difficult to interpret. The purpose of this research was to explore a high-speed presentation technique for creating an environment within which perceptual errors of this type could be reliably induced and studied. This type of error will be referred to as visual dissociation because it describes a situation in which concurrently presented visual components are perceived as temporarily distinct. Under conditions such as these, dissociated components of one display may become integrated with other visual displays presented close in time. The dissociation phenomenon will be discussed in the context of a model of the early stages of scene perception. The results of four experiments will be reported that are consistent with the model and suggest that the dissociation effect will provide a new method with which to study the cognitive processes involved in perceptual integration of complex scenes.

Visual Dissociation and Pictures

When studying a scene, we frequently shift the direction of our gaze, making an average of three eye fixations per second (Yarbus, 1967). There is convincing evidence to suggest that 1/2 of a second is usually sufficient to allow us to understand what we are looking at (Biederman, Mezzanotto, & Rabinowitz, 1982; Biederman, Rabinowitz, Glass, & Stacy, 1974). Visual search experiments using high-

The results of Experiments 1 and 2 were presented at the Psychonomic Society Meeting, Philadelphia, Pennsylvania, November, 1981. The results of Experiments 3 and 4 were presented at the Psychonomic Society Meeting, San Antonio, Texas, November, 1984. The author would like to thank Mary C. Potter and Geoffrey R. Loftus for their comments on the manuscript and H. Rabinstein, M. Proctor, and A. Scott for their assistance in data collection and scoring of results.

Requests for reprints should be sent to Helene Intrabu, Department of Psychology, University of Delaware, Newark, Delaware 19716.
speed presentation of pictures have demonstrated that unrelated scenes can be understood at presentation rates that surpass the average fixation frequency of the eye (Intraub, 1981a, 1981b; Potter, 1975, 1976). For example, subjects were good at categorizing pictures presented at rates ranging from 114 to 258 ms per picture and could often detect and describe a particular picture in the sequence based upon relevant visual information such as "a picture that does not depict a means of transportation" (Intraub, 1981b).

The results of these experiments are consistent with the subjective impression one has while viewing photographs at high speeds: one is not assaulted with an assortment of blurred swatches of color and line, but instead experiences a succession of meaningful scenes. Furthermore, subjects' descriptions of pictures presented at a rate of 9/s that were obtained during search tasks (Intraub, 1981b) or during free recall (Intraub, 1979) do not suggest that subjects are experiencing gross perceptual errors. As in the case of normal visual activity when viewing a common object or scene, visual dissociation, if it occurs at all, must occur quite infrequently. A striking and robust dissociation effect did occur, however, when an unrelated visual component was added to the pictorial sequence. This unexpected dissociation effect emerged under the following conditions.

Viewers watched a sequence of color photographs of common objects (e.g., utensils, machines, foods, animals, etc.) that were cut out of magazines and rephotographed on a gray field. For selected pictures scattered throughout the sequence, a black frame (which was to serve as a visual marker in an experiment) was placed around the periphery of the gray field with the picture in the center. On several occasions it appeared to the viewers that the black frame had been mistakenly photographed around the wrong picture. The error, however, turned out to be the viewers', who were sometimes "seeing" the frame around preceding or following pictures in the sequence. The illusion was so compelling that to determine which picture the frame actually appeared around, it was necessary to inspect the film frame by frame.

The temporal dissociation and faulty integration of the picture and frame is similar to errors reported when alphanumeric stimuli are presented in rapid succession. In these experiments illusory conjunctions of color and form, letter case and word, and errors in noting the simultaneity of concurrently presented characters were obtained (e.g., Lawrence, 1971; McLean, Broadbent, & Broadbent, 1983; Spelten & Reeves, 1980). What is particularly interesting in the present case with pictures is that features of the pictures themselves (color or form) do not appear to dissociate at these speeds. The frequent dissociation of the frame in contrast to this raises interesting questions about the early stages of picture processing. A model of these stages, based upon other picture research, provides a possible explanation of why visual dissociation occurs under these conditions.

A Model of the Early Stages of Picture Processing

It will be proposed, based on previous research, that briefly glimpsed pictures are rapidly understood and are maintained for a few hundred milliseconds in a short-term conceptual store (Potter, 1970). This store, which maintains information at a conscious level, can hold up to three pictures at a time when pictures are presented at high speeds (Intraub, 1984, in press). The argument will be made that it is while more than one picture is in the store that visual dissociation errors such as the one briefly described above can occur. A possible reason for dissociation of the frame, as opposed to other features of the display, can be found in research that studied the effects of context on object perception.

Biederman and his colleagues (e.g., Biederman, 1981; Biederman et al., 1982; Biederman, Kanwoswit, Glass, & Stasty, 1974) demonstrated that the speed and accuracy of object identification is affected by the context of the scene in which the object appears even when exposure duration is as brief as 100 ms. Object identification takes longer and is

---

1 Eye movement recordings show that at rapid presentation rates such as these, subjects suppress major eye movements and fixate the center of the screen (Potter & Levy, 1969).
less accurate when the object appears in an unexpected context. Eye movement experiments, showing longer fixation times on unexpected objects, provide converging evidence that longer processing time is required when an object occurs in a low-probability context (e.g., Antes & Penland, 1986; Friedman, 1979; Loftus & Mackworth, 1978).

According to this model, the reason that the frame dissociates is that, whereas features of the objects themselves do not, is that the frame is not conceptually related to any of the pictures. At high speeds, as pictures momentarily enter the short-term conceptual store, features of the objects are rapidly identified and integrated. The frame, however, is not conceptually related to any of the pictures. As a result, the frame requires more processing time than do the other features in the display. Depending on whether the subject happened to attend to the frame first or to the picture first, the frame may sometimes be interpreted as part of the previous picture's processing or the following picture's processing in the short-term conceptual store.

Four experiments were designed to test this interpretation of the dissociation effect. Experiment 1 was conducted to verify the reliability of this informally observed effect and to obtain a measure of its frequency.

**Experiment 1**

**Method**

Subjects. Subjects were 15 male and female undergraduate volunteers from Bucknell University.

Stimuli. The stimuli were 156 color magazine photographs (e.g., food, animals, machines, plants, people, musical instruments, etc.). The pictures had been cut out in such a way that little or none of the original background remained, and they were rephotographed in the center of a gray field. Average visual duration thresholds had been obtained for these pictures in previous research (Lubar, 1979).

Apparatus. The stimuli were presented using an L-W variable-speed 16-mm cine projector. The image was rear-projected from an outer room onto a screen in the room housing the subject and the experimenter. This served to minimize projector noise. The experimenter controlled the projector from inside the subject's room.

The size of the gray field was approximately 36 cm × 27 cm, and the size of the outline frame (measured from its outer edges) was 33 cm × 25 cm. The pictures were different shapes and sizes. They filled areas ranging from approximately 17 cm × 12 cm to 23 cm × 22 cm. Subjects sat approximately 2.4 m from the screen.

**Filmed sequences.** The pictures were on 35-mm slides. The slides were backlit and photographed using a Bell & Howell 16-mm camera with an extension tube attachment. The frame, which was on clear acetate, was placed directly on the slide when it was photographed. This procedure was used in all four experiments.

Design. The stimuli were divided into 12 sets of 12 pictures each. Each set of 12 was photographed four times using single-frame photography on 16-mm film, yielding four independent sequences per set. In each of the four sequences, one of three pictures was designated to be the target picture, that is, the picture to be photographed with the frame around it. In each of the four sequences, the order of the pictures was designated with the constraint that the three potential targets always appeared next to one another. The order of the three across the four sequences was ABC, CBA, CAB, and BCA, with the central picture serving as the target. (In those sequences where the target was the first or last picture in the sequence, the pattern was the same except that the first or last picture, rather than the central picture of the triplet, was the target.) Those three pictures and the picture immediately preceding and following them were all chosen to have approximately equal visual duration thresholds (VDTs), ranging from 5 to 11 ms.

This was done in an attempt to equate the relevant pictures with respect to ease of identification and to exclude those pictures known to be more difficult to identify. The frame always appeared in the same serial position for the four sequences of a given set. All 12 positions were tested over the course of the experiment; each of the 12 sets showed the target in a different serial position. Thus, each subject viewed 48 experimental sequences. An additional set of 12 pictures was used for practice by the subject.

Procedure. Subjects were individually run. They were seated in the dim light in a projection room. The experimenter sat so that he or she could be seen only if the subject turned away from the screen. Prior to viewing each sequence, subjects were given two sets, the first set of which was familiarized with the 12 pictures constituting that set. This was done by presenting each picture for 5 s and asking the subject to provide a name. Subjects were then shown the four sequences one at a time at a rate of 111 frames per second. Subjects were asked to look at the lower right-hand point and to keep their eyes on the center of the screen while viewing each sequence. They were told not to wait until the end of the sequence but to respond immediately when they saw the frame. They were instructed to report which picture was in the frame and to indicate whether they were "very sure," "pretty sure," "not sure," or "guessing.

Results and Discussion

Although the picture and the frame were presented simultaneously, they were often not

2 Note that these visual duration thresholds were taken without a mask (Lubar, 1979) and should be interpreted as retinal measures, allowing for a relative measure not as an absolute measure of the exposure time needed to identify pictures in a rapidly presented sequence. Mean exposure times ranged from 5 to 55 ms for individual pictures in the stimulus pool.
perceived as an integrated visual event. All 15 subjects frequently reported the frame as appearing around the wrong picture, usually the immediately preceding or following picture in the sequence. Accuracy in detecting the correct location of picture and frame with respect to confidence level is shown in Table 1. Subjects reported being “sure” on 75% of the trials, and “not sure” on 25% of the trials. They reported guessing on 9% of the trials and could offer no response on 2% of the trials. Even when subjects were highly confident of their response, they were often wrong; they made 30% errors and 42% errors in the top two confidence conditions. All subjects exhibited the dissociation effect. Overall, the mean hit rate (proportion of trials in which subjects correctly reported the target picture) was .46 (SD = .17), and the median was .44. Of the 48 experimental sequences, only one sequence did not exhibit the dissociation effect (across subjects). This sequence was one of the four sequences in which the target was in the first serial position. Those four sequences and the four sequences in which the target was in the final serial position yielded the highest average hit rates. These were 73.3 and 78.5, respectively.

Temporal dissociation of simultaneously presented events is certainly not a new discovery. James (1890) discusses the well-known experiments of Wundt and others that demon-strated temporal displacement of simultaneously presented stimuli. These experiments were usually conducted across modalities. Visual, auditory, and/or tactile stimuli were simultaneously presented to a subject. Subjects often reported the stimuli as occurring consecutively. More recent examples of temporal dissociation within the same modality are the chick migration experiments in which subjects were required to monitor a sentence and indicate the onset of a click (cf. Fodor, Revere, & Garrett, 1974) and experiments in which subjects were required to monitor a visually presented stream of letters, find a target letter, and report a number from a simultaneously presented but spatially disjointed stream of numbers (Sperling & Reeves, 1980).

In experiments closer to the type reported in the present article, Lawrence (1971) and

<table>
<thead>
<tr>
<th>Confidence</th>
<th>Hit</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
<td>0</td>
<td>1</td>
<td>19</td>
<td>70</td>
</tr>
<tr>
<td>-2</td>
<td>3</td>
<td>8</td>
<td>28</td>
<td>16</td>
</tr>
<tr>
<td>-1</td>
<td>0</td>
<td>7</td>
<td>26</td>
<td>14</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>19</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

Note. Numbers indicate the position of the reported picture with respect to the picture that actually appears with the frame (e.g., +1 indicates the immediately preceding picture, +2 the second following picture and +3 the third following picture or further).

McLean et al. (1983) presented alphaquamic stimuli at high speeds in a single location and asked subjects to search for a particular attribute. Lawrence (1971) presented words and required subjects to indicate which one of the words was in capital letters. McLean et al. (1983) presented letters or numbers and required subjects to indicate which one was in a specified color. Both studies reported migration of the attribute to temporarily adjacent stimuli. Experiment 1 shows that visual dissociation of this type is not limited to cases in which the subject must read symbols (alphaquamic characters or words) at high speeds. In addition, as mentioned previously, the "migrating frame" effect is particularly interesting because features of the color photogra-phy themselves do not seem to dissociate and migrate to other pictures (Intraub, 1981a).

According to the picture processing model described earlier, the freer movement of the frame reflects the role of expectation and meaning in scene perception. An alternate explanation of the phenomenon is that it is simply an artifact of the spatial separation and lack of contiguity between the picture and the frame. The shift in spatial attention from one part of the scene to another takes time (e.g., Sperling & Reeves, 1983), and the shift, in conjunction with the rapid rate of presentation, results in the illusion. It should
be noted that both approaches would account for the direction of the migration in the same way, suggesting that it is determined by whether the subject first attends to the picture or to the frame.

The purpose of Experiment 2 was to determine if dissociation of the frame reflects the spatial separation of the picture and the frame or may instead reflect the action of integration processes that rely to a great extent on meaning and expectancy.

Experiment 2

To test the spatial separation hypothesis, Experiment 2 compared integration accuracy between two conditions. In one condition, referred to as the large-frame condition, a smaller black outline rectangle, placed directly in the center, was presented. The rationale was that if spatial separation and the lack of contiguity between the large frame and the picture caused the dissociation to occur, then no dissociation should be obtained with the small frame.

The films differed from those used in Experiment 1 in that (a) a single set of 12 pictures were used throughout the experiment so that subjects could become more familiar with the stimuli; (b) the subject was provided with no advance indication of where in the sequence the frame would appear (e.g., early or late); and (c) the first and last pictures in the sequence were never presented with a frame. The large-frame condition provided a test of the replicability of the dissociation effect obtained in Experiment 1 under these conditions.

Method

Subjects. The subjects were 16 male and female undergraduates from Bucknell University, reporting normal or corrected vision.

Stimuli. The experimental stimuli were 12 pictures from the stimulus pool described in Experiment 1. These included, a car, a flag, a hot air balloon, a truck, a golfer, a movie projector, a pair of eyes, a snail, an organ, a chat, a minivan, and a tractor. As in Experiment 1, they were selected on the basis of VDT and their dissociation with respect to naming. These pictures were among those exhibiting the lowest VDTs in the stimulus pool (all had average VDTs of 3 ms). Another 12 pictures were selected to be used as practice sequences for the subject.

Apparatus. The apparatus was the same as in Experiment 1. The size of the small frame was 2.5 cm x 2.9 cm (measured from the outer edges).

Framed sequences. Each of the two conditions (large frame and small frame) had 12 practice sequences (using the 12 practice pictures) and 24 experimental sequences (using the 12 experimental pictures). Across sequences, targets were presented equally often in Serial Positions 2, 5, 6, or 11. Serial position varied randomly from one sequence to the next, unlike Experiment 1, where the same serial position was used in each set of four sequences. In the experimental sequences each of the 12 pictures served as target twice. Because the subject's reaction might be affected by the appearance of a particular striking picture (in spite of equal VDTs), the same two photographs flanked the target both times if appeared, although their order was reversed the second time to see if this would affect the patterns of errors. The order of all other pictures was held constant in both sequences.

Procedure. All 16 subjects took part in both the large-frame and small-frame conditions, with order of condition counterbalanced across subjects so that half the subjects saw the large-frame condition first and half saw the small-frame condition first. Each condition was preceded by the 12 practice sequences. In the practice portion and experimental portion of each condition, prior to viewing the rapidly presented sequences, subjects were shown each of the 12 pictures for 10 s and were provided with a name. The subjects were then shown what each picture would look like with the frame (large or small depending on condition) and received practice naming the pictures. As in Experiment 1, subjects were instructed to (a) fixate the center of the screen, (b) respond immediately with the name of the particular picture, and (c) indicate their confidence, this time by responding "sure," "pretty sure," "not sure," or "guess.

Results and Discussion

Contrary to the spatial separation hypothesis, subjects often saw the small frame in the center of the wrong picture. Again, errors usually involved the immediately preceding or following picture in the sequence. In the large frame condition, subjects were "sure" on 7% of the trials, "pretty sure" on 52% of the trials, "not sure" on 37% of the trials, and were "guessing" on 4% of the trials. Confidence ratings were similar in the small-frame condition, where the breakdown of ratings was 10%, 47%, 35%, and 8%, respectively. Accuracy in determining the correct conjunct of picture and frame for the combination of the two highest confidence ratings for each condition is shown in Table 2. There was not a significant difference in the proportion of hits between the large- and small-frame conditions, t(15) = 1.63, ns (combina-
Table 2
Percentage of Responses Reporting the Frame on the Correct Picture (Hit), Preceding Pictures (−), and the Following Pictures (+), in the Large-Frame and Small-Frame Conditions for the Combination of the Two Highest Confidence Ratings (Rounded to the Nearest Whole Number)

<table>
<thead>
<tr>
<th>Condition</th>
<th>−3</th>
<th>−2</th>
<th>−1</th>
<th>Hit</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large</td>
<td>1</td>
<td>0</td>
<td>38</td>
<td>44</td>
<td>16</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Small</td>
<td>2</td>
<td>1</td>
<td>21</td>
<td>55</td>
<td>22</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

**Position of reported picture in sequence**

- Large: 1 0 38 44 16 1 1
- Small: 2 1 21 55 22 0 2

**Rationale and Discussion**

The data presented in Table 2 indicate that the effect of the frame on the recognition of a picture depends on the position of the picture in the sequence. The large-frame condition resulted in a higher percentage of correct responses when the picture was preceded or followed by other pictures, whereas in the small-frame condition, the effect was less pronounced. This suggests that the presence of the frame affects the recognition of pictures, and this effect is more evident when the picture is not isolated within the sequence.

**Implications**

1. **Single and Double-Frame Effects**
   - The effect of the frame on the recognition of pictures was more pronounced when the frame was not isolated. This indicates that the frame may have a more significant impact when it is part of a larger context, such as a sequence of pictures.

2. **Interactions Between Frames**
   - There were no significant interactions between the large and small frames, indicating that the effect of the frame was consistent across the two conditions.

3. **Methodological Considerations**
   - The use of confidence ratings provides a more nuanced understanding of the effect of the frame on picture recognition. This method allows for a more detailed analysis of the data, highlighting the varying degrees of confidence in the responses.

**Further Research**

Future research could explore the role of other factors, such as the temporal proximity of the frames, the number of pictures in the sequence, and the individual differences in recognition patterns, to better understand the impact of frames on picture recognition.

**Conclusion**

The results suggest that the presence of the frame in a sequence of pictures can influence the recognition of individual pictures, particularly when the frame is not isolated. This effect is more pronounced in the large-frame condition, highlighting the importance of considering the context in which pictures are presented.
The results of Experiment 2 are consistent with the model of picture processing described earlier. As in Experiment 1, when confidence is high, the frame rarely migrates more than one picture in either direction. This suggests that the short-term buffer holds two pictures simultaneously—the picture with the frame and either the preceding or following picture in the sequence. If the subject attends to the frame, the target should be integrated with the preceding picture (which is still in the buffer at that point). If the subject attends to the picture first, the frame may become integrated with the following picture (which has just entered the buffer). According to this view the short-term buffer plays a role in integrating the features in a display. All the features are present in the buffer, but the integration process can break down and thus result in misperceptions. There is a strong alternate explanation of the results, however, that minimizes the integrative role of the short-term store.

According to this alternate view, the information from both pictures is not always available in the short-term store. While the subject is searching for the relatively simple set of features composing the frame, at these rapid presentation rates, the flanking pictures sometimes mask major areas of the target picture due to visual masking and/or conceptal masking (Intraub, 1984; Potter, 1976). The subject detects the frame, but the target picture itself is obscured. On those trials where dissociation of the target and the frame occurs, subjects do not see the target. Search experiments with pictures have shown that although the ability to identify briefly presented pictures is very good, detection accuracy is not perfect at speeds of about 9/s. Depending on the type of cue, detection accuracy ranges from about 35% to 60% correct (Intraub, 1981a; Potter, 1975, 1976).

Therefore, it may be the case that on those trials in which migration occurs, the target picture itself is masked. This is different from the perceptual integration hypothesis, which suggests that the target picture is perceived and held in the short-term buffer along with one other picture and that the frame becomes integrated with the wrong one. Experiment 3 provides a test of these two hypotheses.

Experiment 3

To determine if the target picture is perceived by the subject when the frame is integrated with another picture, subjects took part in two conditions, using the same large-frame sequences as in Experiment 2. In the first condition, the subject received the same instructions as in the previous experiments. They were instructed to report which picture had the frame around it (the standard condition). In the second condition the subjects were instructed to report the picture that appeared immediately following the picture in the frame. According to the masking hypothesis, the target picture should never be reported as the picture following the frame. It will either be seen as the "picture with the frame" or it will be masked. According to the short-term buffer model, the subjects should frequently report the target picture as being the picture following the frame. In principle, they should report it as frequently as they report the frame to be around the preceding picture in the standard condition.

Method

Subjects. Subjects were 10 male and female undergraduates from the University of Delaware who were fulfilling a research requirement in introductory psychology.

Stimuli. The same filmed sequences were used as in Experiment 2, large-frame condition.

Apparatus. A new lab was used with a similar rear-projection arrangement. The projector was a Visual Instrumentation Corporation variable speed 16-mm cine projector. The image sizes were approximately the same as in Experiments 1 and 2, but subjects were seated closer to the screen (1.1 m from the screen to the middle of the chair).

Procedure. The procedure was the same as in the large-frame condition, Experiment 2. The only difference in procedure was that following the large-frame condition, the subject was shown the same 12 practice sequences and 24 experimental sequences, with the instruction to report the picture immediately following the picture with the frame. The same confidence ratings were used. Subjects were not aware that they were seeing the same sequences a second time.

Results and Discussion

Contrary to the masking hypothesis, subjects frequently reported the target picture as the picture immediately following the frame (35% and 33% of the time with the "sure"
and “pretty sure” confidence ratings, respectively. Accuracy in detecting the correct picture in each condition for the combination of the two highest confidence ratings is shown in Table 3. These results clearly replicate the dissociation effect (between conditions) and show that it cannot be attributed simply to masking of the target picture. Subjects reported the target picture as the picture following the frame as frequently as they had reported the preceding picture as being in the frame (standard condition). The mean number of responses (with the two highest confidence ratings) for the standard condition and the picture following condition, respectively, was 4.9 (SD = 1.85) and 4.4 (SD = 2.27), t(15) = 6.4, n.s. All 10 subjects showed both types of responses with the high confidence ratings. Apparently, the frame can migrate to other pictures, leaving perception of the target picture intact. The subject identifies the target picture as a picture without a frame. This is consistent with comments occasionally made by subjects in the standard condition (across all the experiments) in which they name the target picture and say that the frame was around the picture that just preceded or followed it.

The table shows that subjects performed comparably on the two detection tasks. Accuracy in detecting the correct picture was about the same in both conditions, with 38% correct detections in the standard large frame condition and 35% correct in the picture following-the-frame condition. The breakdown of responses falling into each of the four confidence levels from highest to lowest was 7%, 42%, 37%, and 12% in the standard condition and 11%, 41%, 32%, and 13% in the other condition. In both conditions subjects were unable to respond (reporting a lapse of attention or missing the frame) on only 3% of the trials.

Experiment 4
If the dissociation effect is caused by a combination of long integration times for unrelated visual features and concurrent processing of items in the short-term buffer, then it should be affected by the complexity of the stimuli in the sequence. In other words, if the color pictures are replaced by more readily identifiable stimuli, then given the same presentation rate, the dissociation effect should be eliminated or at least reduced. Even though the frame would still be an unrelated visual feature, the reduced processing time required by the stimuli would decrease the likelihood of the frame’s becoming integrated with the previous or following stimulus in the sequence. For example, under these conditions, identification of the preceding item may be completed before the target stimulus and frame are presented, and integration of those two forms may be completed before the following item enters the buffer. According to this reasoning, although dissociation should be eliminated or reduced by using more readily identifiable stimuli, a robust dissociation effect should be obtained with these stimuli when presentation rate is increased. The stimuli chosen to test this hypothesis were the numerals 0–9.

The hypothesis that a more rapid presentation rate would be required to obtain frame migration with digits was supported. In part, by experiments in which letters or words were presented at high speeds. In these experiments much more rapid rates were used than in Experiments 1–3. For example, Lawrence (1971) asked subjects to indicate which word in a rapidly presented series was in capital letters, and found dissociation errors at rates of 16–20 items/s. At a rate of 9 items/s (the same rate used in Experiments 1–3), no dissociation errors occurred. McLean et al. (1983) reported visual dissociation of color and form under conditions where subjects had to report which letter in a rapidly presented stream was printed in a specified
color. They used a presentation rate of 15 items/s. These studies suggest that the prediction regarding visual dissociation and presentation rate may hold. But because neither task required a simultaneity judgment about forms (as in the case of the present experiments) and because they used different presentation procedures, Experiment 4 was conducted. Experiment 4 used the same procedure as in Experiments 2 and 3 but replaced the pictures with numbers.

Method

Subjects: The subjects were 10 male and 10 female University of Delaware undergraduate volunteers reporting normal or corrected-to-normal vision.
Stimuli and fanned sequences: The stimuli were black, single-digit numbers from 0-9. They were photographs in the same way as the pictures in the previous experiments. In order to have 12 items in a sequence, two numbers were repeated. In no case did the same two numbers surround the target number (e.g., 2-4-2, when 4 is the target). The same black frame was used as in the previous large frame experiments. Each digit served as the target in two of the sequences. The target serial positions were the same as in Experiments 2 and 3.
Apparatus: The same apparatus was used as in Experiment 3. The numbers were approximately 5 cm × 1 cm (height and width of numeral 0).
Procedure: The procedure was the same except that there were 10 sample sequences and 20 experimental sequences (because there were 10 different stimuli, and each served as target twice). The 20 experimental scenes were shown twice, once forward and once in reverse, so that subjects saw a total of 40 sequences.

Results

When subjects saw the number sequences at 9/16, no dissociations occurred, so the next highest speed available on our equipment was tested. Subjects viewed the sequences at about 18 frames per second (a speed within the range used by Lawrence, 1971, and McLean et al., 1983). At this rate of presentation, dissociation frequently occurred. Table 4 shows the proportion of correct and incorrect conjunctions of numbers and frames at each confluence level.

Using the procedures described in this article, the dissociation effect, which occurs frequently for pictorial stimuli, does not occur for number stimuli presented at the same rate. When presentation speed is increased, however, a large dissociation effect emerges. This difference, which may be due to processing time differences for color pictures and digits, will be discussed further in the next section.

Another difference between these results and those obtained in Experiments 1-3, is the pattern of errors. In the first three experiments the errors included both the preceding item and the following item in the sequence, although there was sometimes a greater number of errors involving the immediately preceding picture. In the present experiment, like the other experiments with alphanumeric stimuli, the errors tended to occur on the immediately following picture and some beyond that. Errors involving the preceding pictures were relatively rare. It's not clear how to interpret the difference. Before attributing it to factors related to stimulus type, it is important to note that James (1980) reported that with the same cross modality stimuli, at slower rates errors tended to occur more with preceding items, and as rate was increased, errors tended to occur more with following items.

General Discussion

Pictures were rapidly presented (approximately 9/16), and subjects were required to detect the presence of a frame (a black outline rectangle) and indicate the picture within which it appeared. Experiment 1 demonstrated that simultaneous presentation of a picture and a surrounding frame did not ensure that they would be experienced as a single visual event. Subjects often reported the frame as having appeared around a different picture, usually the immediately pre-

Table 4

<table>
<thead>
<tr>
<th>Confidence</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>Hit</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sure</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>13</td>
<td>41</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Pretty sure</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>21</td>
<td>26</td>
<td>12</td>
<td>19</td>
</tr>
<tr>
<td>Not sure</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>17</td>
<td>19</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td>Guess</td>
<td>7</td>
<td>4</td>
<td>10</td>
<td>16</td>
<td>15</td>
<td>11</td>
<td>37</td>
</tr>
</tbody>
</table>

| Position of reported numeral in the sequence |

<table>
<thead>
<tr>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>Hit</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sure</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>13</td>
<td>41</td>
<td>9</td>
</tr>
<tr>
<td>Pretty sure</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>21</td>
<td>26</td>
<td>12</td>
</tr>
<tr>
<td>Not sure</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>17</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>Guess</td>
<td>7</td>
<td>4</td>
<td>10</td>
<td>16</td>
<td>15</td>
<td>11</td>
</tr>
</tbody>
</table>
coding or immediately following picture in the sequence. The phenomenon, referred to as the visual dissociation effect, was not limited to those situations where the subject war unsure but occurred reliably for all subjects at the highest confidence level.

The results of Experiment 2 provide a replication of the dissociation effect and show that it cannot be attributed to spatial separation and the lack of contiguity between the large frame and the target picture because the same effect was obtained using a small frame photographed in the center of the picture itself. Pilot research has shown that unrelated color backgrounds also dissociate from the target picture and, like the frame, seem to appear behind the immediately preceding or following picture in the sequence. These errors also occurred when subjects were confident of their response.

The pattern of responses obtained in Experiment 3 shows that migration of the frame cannot be attributed to the loss of the target picture via visual masking or conceptual masking (cf. Potter, 1976). Contrary to the masking hypothesis, when subjects were required to report the picture that followed the frame, they often reported the target picture. They did this as frequently as they had erroneously reported the frame as appearing around the preceding picture in the standard condition. This indicates that subjects often perceived the target picture as an "unframed" picture and had integrated the frame with another picture in the sequence.

Finally, in Experiment 4 it was demonstrated that frame migration does not depend solely on presentation rate. The type of stimulus (in this case digits or pictures) had a pronounced effect on whether or not frame migration occurred at a given rate of presentation. Taken together, Experiments 1–4 are consistent with the model of scene processing described previously and suggest that frame migration reflects integration processes that occur once identification has been initiated. Before addressing this position in more detail, it should be noted that these results neither support nor refute Treisman and Gelade's (1980) position that certain classes of features (e.g., color and form) are initially registered independently in the visual system and are conjoined by focal attention. The illusory conjunctions obtained in the present research certainly seem to occur at a later stage of processing than initial sensory registration. They seem to be based on meaning, expectation, and identification time, and they support the view that a very short-term memory which is involved in the type of illusory conjunctions, as well as some showing an effect of word structure on color or letter migration (e.g., Moser, 1983; Prinzmetal & Mills-Wright, 1984), show that a relatively high level of analysis mediates some types of visual dissociation.

Scene Integration and Short-Term Conceptual Memory

The results of Experiments 1–3 are consistent with cognitive models of scene perception in which meaning and expectation based on general knowledge affect the early stages of processing (e.g., Biederman, 1981). When pictures are presented at high speeds, descriptions of pictures obtained during visual search (Intraub, 1981b) and during free recall (Intraub, 1979) indicate that features from one picture do not generally migrate and merge with other pictures. One does not typically perceive odd combinations of features belonging to temporally adjacent pictures. On the other hand, the large frame, small frame, and the red, green, or black backgrounds, which are all unrelated to the pictures, frequently dissociate from the host picture. A major factor contributing to the dissociation effect may be the lack of a meaningful context.

The effect of the meaning or "gist" of a scene on the speed and accuracy of object perception has been demonstrated by Biederman and his colleagues (e.g., Biederman, 1981; Biederman, Glass, & Stacy, 1973; Biederman, Rabinowitz, Glass, & Stacy, 1974). Biederman et al. (1974) required subjects to identify an object presented in either a coherent or a jumbled scene. A visual cue, presented at the picture's offset, indicated the location of the object. The pictures, which were irregular drawings of common scenes, were presented briefly, in some conditions for as little as 50 or 100 ms. They found that accuracy of object identification was enhanced when the object appeared in a coherent scene. Facilitation was also obtained in conditions where the object appeared in a coherent scene that was meaningfully related to it as
computational and cognitive processes involved in perceiving and understanding visual stimuli.
be obtained for different types of stimuli by comparing the presentation rates necessary to obtain a given frame migration frequency.

More important, the dissociation effect should provide an interesting new tool with which to study the role of meaning and expectation in scene perception. Research is currently under way in which instead of searching for the conjunction of a meaningless frame and a picture, subjects will search for the conjunction of a meaningful object and a scene. The conceptual and physical relatedness of the object to the host scene and to the immediately preceding or following scenes will be manipulated to determine the effect of each on the magnitude and direction of the dissociation effect.

Clearly, the relatedness of an object to a scene affects the time necessary to identify that object. Different processing time requirements in conjunction with the rapid succession of pictures on film or in conjunction with a sudden rapid succession of eye movements during visual scanning provide an explanation of the dissociation effect observed in the laboratory, as well as the double takes and erroneous perceptions that occasionally occur during normal visual scanning.

References


Received June 27, 1984
Revision received February 6, 1985

442 HELENE INTRAUB