# Illusory Conjunctions of Objects and Forms: Integration Errors in a Very Short-term Store

Helene Intraub

Department of Psychology
University of Delaware

ABSTRACT

Two experiments tested the predictions of an integrative buffer model of visual processing, regarding the illusory conjunction of components of rapidly presented displays. Color pictures of objects were presented at a rate of 9/s, in the same spatial location. Experiment 1 used a modified report procedure to test the hypothesis that Stroop-like response competition, during naming, not a perceptual error, resulted in the high confidence "illusory conjunctions" reported in previous research. Subjects were provided with the name of a picture in advance and reported "yes" or "no" to indicate if that picture was the one in the frame. Contrary to the response competition hypothesis, high confidence errors occurred frequently under these conditions. Experiment 2 tested the hypothesis that the direction of migration (preceding or following picture) is the result of a difference in the sequential allocation of attention to the frame first or to its "host" picture first on different trials. As predicted by the integrative buffer model, subjects were faster in detecting the frame when they confidently reported it around the preceding picture than around the following picture in the sequence, and reaction times associated with correct reports fell between the two.

## INTRODUCTION

The purpose of these experiments was to test the predictions of a model of the early stages of scene processing that Intraub (1985a) proposed might account for temporal migration and illusory conjunctions of components of visual displays. Temporal migration is a type of illusory conjunction of visual components that occurs when stimuli are presented in rapid succession (e.g., 9-20 items per second) in the same spatial location. Although components of a single display are simultaneously presented they are sometimes reported, with high confidence, as having occurred separately. Subjects will report a component as having been an integral part of the preceding or following display in the sequence. There have been reports of color migration among letters (McLean, Broadbent & Broadbent, 1983), letter case among words (Lawrence, 1971), form migration among letters, words, numbers and pictures (Gathercole & Broadbent, 1984; Intraub, 1985a) and object migration among scenes (Intraub 1985b).

The model proposed to account for these phenomena, is based upon a model of the early stages of visual processing, in which a very short-term conceptual/visual buffer memory plays a central role (Potter, 1976; Avons & Phillips, 1980). Several experiments using search tasks or memory tasks have provided evidence for a very short-term post-categorical store that is distinct from the icon (Avons & Phillips, 1980; Intraub, 1981, 1984; Loftus & Ginn, 1984; Potter,1976). Intraub (1985a) proposed that this conceptual short-term store may play a role in the integration of visually presented information.

According to this view, migration occurs when identification time is slow relative to presentation rate because one display is still being analyzed in the buffer at the same time that processes are initiated on the next new display. Thus a black outline frame migrates among pictures, but parts of the pictures themselves do not (Intraub, 1985a) because the former display requires more identification time. Similarly, to obtain the same level of frame migration among different types of stimuli, a faster presentation rate must be used for stimuli that are relatively easy to identify, (e.g., numbers and letters) as compared with those that are more difficult (e.g., colored objects and words; Gathercole and Broadbent, 1984; Intraub 1985a).

The reason that components sometimes migrate to the preceding display and sometimes to the following one, can be explained in terms of the allocation of attention. Consider Intraub's (1985a) task in which subjects must report which object was presented with a black frame around it. The preceding picture in the sequence is in the short-term buffer when the picture with the frame is presented. Although presented simultaneously, the frame and picture are not rapidly integrated because they are not meaningfully related. As a result, if the subject pays attention to the frame first, it may become integrated with processing of the previous picture, which is still in the visual/conceptual buffer. If the subject pays attention to the target picture first, then the frame, which is now in the buffer, may become integrated with the next picture in the sequence.

Two experiments tested different aspects of the integrative buffer model. Experiment 1 used a modified report procedure and showed that high confidence migration reports cannot be attributed to difficulties in naming rapidly presented displays. Experiment 2 tested the validity of the attention allocation hypothesis as an explanation of differences in the direction of migration, using a reaction time task in conjunction with a standard temporal migration task.

## EXPERIMENT 1

To determine if high confidence error reports obtained in temporal migration experiments are an artifact of the naming requirement during high speed presentation, the report procedure was changed from the unconstrained naming procedure typically used in temporal migration experiments (which requires subjects to access and articulate the proper name for the target) to a "yes-no" detection task. A picture was named by the experimenter and the subject's primary task was to answer "yes" or "no" to indicate whether or not the frame was around the specified picture. If high confidence error reports to -1 and +1 pictures are the result of response competition during naming, then no high confidence error reports to specified -1 and +1 pictures should be obtained using this procedure. Specification of -2 and +2 pictures by the experimenter in some sequences, served as "catch trials" to allow a measure of "yes" response bias.

## Method

Subjects and Apparatus. Subjects were 18 individually-tested undergraduates. They 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, data analyzer 16mm projector at silent speed (18 frames per second). The size of the field was 14 x 20 cm, which subtended a visual angle of approximately $5^{\circ} \times 8^{\circ}$.

Stimuli. The stimuli were twelve objects that were cut out from magazines and photographed on a gray background. These are the same stimuli that were used in Intraub (1985b: Experiments 2 and 3). The objects were: 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.

Design. Each sequence contained all twelve pictures, with one of the twelve in the black frame. Each picture was photographed with the black frame around it six times, yielding 72 sequences. On the six occasions that a picture was the target it appeared with three different pairs of flanking pictures (-1 and +1 pictures), such that on one occasion the order was ABC and on the other it was CBA (B is the picture with the frame). Subjects were divided into three equal groups. What differed among the three groups, was the position of the specified picture (-1, target, or +1) in a given sequence. The position of the specified picture in each sequence was counterbalanced across subject group. Twelve additional sequences were filmed to serve as "catch trials" (-2 and +2 pictures specified).

Procedure. Before the experiment, subjects were familiarized with the pictures and received practice naming them. Following this, they were told that prior to each sequence, the experimenter would name a picture while they looked at the fixation point. Their task was to respond "yes" or "no", followed by a confidence rating (sure, pretty sure, not sure, guess), to indicate whether or not the frame was around the specified picture. On trials eliciting a "no" response they were asked to report the picture that they saw the frame around, along with a second confidence rating.

## Results

The results support the claim that the migration effect obtained with pictures and forms (Intraub,1985b) is not due to difficulty in naming rapidly presented pictures. All subjects persisted in reporting the frame around temporally adjacent pictures in the sequence, with high confidence. As in Intraub (1985b) the yes/no procedure yielded the same general pattern of results as the unconstrained naming procedure (Intraub, 1985a, 1985b).

Table 1 shows the percent of trials in each specification condition to which the subject reported "yes" with high confidence. Collapsing over minus and plus positions, a repeated measures ANOVA showed a significant main effect of position, ($F_{(1,34)} = 61.48$, $p < .001$). High confidence "yes" responses decreased as the specified picture was further removed from the target. A planned comparison of the percent of "yes" responses in the -1/+1 condition and the -2/+2 position shows that the large number of erroneous "yes" responses to -1/+1 pictures cannot be accounted for by a guessing bias because there were significantly fewer "yes" responses to -2/+2 pictures ($p<.001$). The means for the -1/+1 and -2/+2 conditions, were 44.9 (SD = 21.1) and 13.8 (SD = 18.0), respectively.

Whether the relatively small number of "yeses" to -2/+2 pictures, reflects a guessing bias or contained actual integration errors is being addressed in other research. The major point being made here is that a strong migration effect was obtained with these sequences using the yes/no task.

Table 1

Mean percent of total possible responses in each specification condition that were high confidence "yes" responses (Exp 1)

| | Specification Condition | | | | |
|---|---|---|---|---|---|
| | -2 | -1 | 0 | 1 | 2 |
| Mean | 14.9 | 43.6 | 63.3 | 46.1 | 13.8 |
| SD | 17.0 | 23.1 | 18.7 | 24.9 | 21.3 |

## EXPERIMENT 2

The purpose of Experiment 2 was to test the attention allocation assumption of the model using a reaction time task in conjunction with the standard temporal migration task. According to the attention allocation assumption, -1 errors occur when subjects attend to the frame before the "host" picture and +1 errors occur when the subjects attend first to the "host" picture and then to the frame. If the assumption is true, then reaction time to frame detection (measured with a key press) should be faster when subjects report the frame around the -1 picture, than around the +1 picture. The times obtained on trials in which the frame is reported around the host picture should neither be faster than the -1 time nor slower than the +1 time.

### Method

Subjects and stimuli. The subjects were 27 undergraduate volunteers who were paid $3.00 for their participation. The filmed sequences were the same as in Experiment 1.

Apparatus. The apparatus was the same except for the addition of a hand-held key press which the subject depressed with his or her thumb. Reaction times were measured using an Apple II Plus computer that was interfaced with the digital frame counter of the variable speed data-analyzer projector. Reaction times from the onset of the frame to the key press were accurate to 1 msec.

Procedure. Subjects were familiarized with the pictures, and were given practice naming them. Following this, they were told that their task was to press the key in the shortest time possible in response to the frame, and then to immediately report which picture it had occurred with, along with a confidence rating. Each subject was provided with 24 practice sequences followed by 168 experimental sequences. After each group of approximately 28 sequences, subjects received a brief break and viewed a relatively slow presentation of the 12 pictures which they were asked to name.

### Results

Table 2 shows the percentage of pictures reported as occurring with the frame as a function of the picture's position in the sequence and confidence rating. The table shows that addition of the reaction time task did not alter the pattern of results typically obtained with these and other pictorial stimuli using the unconstrained naming procedure (Intraub, 1985a, 1985b).

Table 2

**Percent of responses reporting the frame around the correct scene (0), preceding scenes (-), or following scenes (+) in the sequence as a function of confidence level (Experiment 2)**

| | Position of the reported scene in the sequence | | | | | | |
|---|---|---|---|---|---|---|---|
| Confidence | -3 | -2 | -1 | 0 | 1 | 2 | 3 |
| Sure | 0 | 1 | 8 | 81 | 9 | 0 | 1 |
| Pretty Sure | 2 | 2 | 19 | 58 | 15 | 1 | 3 |
| Not Sure | 7 | 6 | 23 | 36 | 18 | 3 | 6 |
| Guess | 12 | 7 | 19 | 28 | 15 | 6 | 13 |

The mean reaction time to frame detection as a function of the position of the reported picture in shown in Table 3. As may be seen in the table, the pattern of results followed the prediction of the model. Reported position of the frame significant affected reaction time ($F(2,52) = 19.35$, $p < .0001$).

Table 3

**Mean reaction time to detecting the frame as a function of the position of the picture the subject reported seeing it with when the subject reported high confidence in the response (Exp 2)**

| | Position of the reported picture | | |
|---|---|---|---|
| | -1 | 0 | +1 |
| Mean | 327 | 332 | 353 |
| SD | 42 | 40 | 46 |

Frame detection was faster on trials in which subjects reported the frame around the -1 picture than around the +1 picture ($t_{(26)} = 5.33$, $p < .005$). This comparison was also significant by sign-test, with 22 of the 27 subjects showing faster RTs to -1 pictures.

## Conclusions

The results of these experiments are consistent with the integrative buffer model proposed by Intraub (1985a). According to this view of illusory conjunctions, they occur because information is still being processed in a conceptual/visual buffer, at the time that the new display is presented. Parts of the pictures themselves do not migrate at these speeds because components of the pictures can be quickly integrated (e.g., eyes in a face). This integrative buffer would not necessarily play a role in integrating the information from successive fixations (see Pollatsek, Rayner & Collins, 1984), because it seems to be working during the a time interval comparable to a single fixation. Indeed if film speed is reduced so that rate of picture presentation approximates the average eye fixation frequency of about 3 or 4 per second, no frame migration occurs. These phenomena seem more related to temporal integration times in perception, and provide a new means with which to study the perceptual organization of the components of simultaneously presented visual information.

## 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.

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

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

Intraub, H. (1985b). Illusory Conjunctions: Integration errors in a very short-term store. Psychonomic Society Convention, Boston, MA, November, 1985.

Intraub, H. (1985c). Conceptual masking of briefly glimpsed photographs In Groner, R., McConkie, G.W. & Menz, C. (Eds.), Eye movements and psychological processes. North Holland Publishing Co.

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,

Pollatsek, A, Rayner, K. & Collins, W.E. (1984). Integrating pictorial information across eye movements. Journal of Experimental Psychology: General, 113, 426-442.

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.