Illusory Conjunctions in the Perception of Objects

Anne Treisman and Hilary Schmidt
University of British Columbia, Vancouver

In perceiving objects we may synthesize conjunctions of separable features by directing attention serially to each item in turn (A. Treisman and G. Gelade, *Cognitive Psychology*, 1980, 12, 97–136). This feature-integration theory predicts that when attention is diverted or overloaded, features may be wrongly recombined, giving rise to "illusory conjunctions." The present paper confirms that illusory conjunctions are frequently experienced among unattended stimuli varying in color and shape, and that they occur also with size and solidity (outlined versus filled-in shapes). They are shown both in verbal recall and in simultaneous and successive matching tasks, making it unlikely that they depend on verbal labeling or on memory failure. They occur as often between stimuli differing on many features as between more similar stimuli, and spatial separation has little effect on their frequency. Each feature seems to be coded as an independent entity and to migrate, when attention is diverted, with few constraints from the other features of its source or destination.

When we perceive and identify any complex object, we normally register not only its features (its particular shape, size, color, etc.), but also the fact that they are conjoined in a particular configuration. We see the rose as red, its leaves as green, and the vase as gray. The allocation of colors and sizes to shapes, movements to locations, voices to speakers, seems immediate and automatic. We are seldom aware of making errors in the form of illusory conjunctions, for example seeing a green rose, a red vase, and gray leaves. Yet there is both psychological and physiological evidence which suggests that at least some different dimensions are initially analyzed separately, by functionally independent systems (Shepard, 1964; Garner, 1974; Zeki, 1976).

We have recently proposed a feature-integration theory of attention (Treisman, Sykes, & Gelade, 1977; Treisman & Gelade, 1980), which suggests that focused attention is necessary to conjoin features correctly whenever two or more objects are present and these objects vary along the same dimensions, so that their features could be wrongly recombined.

We are grateful to the National Scientific and Engineering Research Council of Canada which provided the financial support for this research. We also thank Daniel Kahneman and William Prinzmetal for many helpful suggestions and discussions and Linda Roberts, Janet Souther, and Jacquie Burkell for their help in running the experiments.

Address reprint requests to Dr. Anne Treisman, Department of Psychology, University of British Columbia, 154-2053 Main Mall, Vancouver, British Columbia, V6T 1Y7 Canada.
Note that we use the term "dimension" to refer to the set of possible, mutually exclusive states of a variable (e.g., the set of orientations, or the set of colors), and "feature" to refer to particular values on a dimension (e.g., vertical or red). The theory proposes that attention is directed to one object at a time, allowing those features which co-occur in the same attention "fixation" to be conjoined into the correct, unitary whole. When task conditions, such as brief exposure, overloading, or the demands of a competing primary task preclude the serial focusing of attention on each of the items present, "illusory conjunctions" may be formed.

We have carried out a number of experiments testing predictions from this hypothesis in a variety of different paradigms. An early result was that visual search for a target specified by a conjunction of features (a green T in a background of green X's and brown T's) requires serial scanning, while search for a disjunctive pair of features (blue or S in the same background) is parallel. The same pattern is observed with conjunctions of discriminable or of confusable features, although the scanning rate may vary considerably. The contrasting patterns of search for conjunctions and for elementary features are largely independent of visual angle and of whether the items are arranged in a regular matrix or haphazardly scattered. They can also be demonstrated with elementary components of shapes: search for target R is serial in distracters P and Q (which jointly contain all the features of R), while search for R in the more similar distracters P and B is not. Texture segregation, which presumably depends on preattentive processing, is automatic and rapid when regions differ in simple features, but there is no perceptual boundary between regions that differ only in the conjunctions of features. Search for a conjunction target benefits greatly when the relevant location is precued, but search for single features does not. Finally, features can sometimes be identified without being localized, while their conjunctions cannot; this supports the hypothesis that focusing attention on a location is a necessary prerequisite for correctly conjoining the features contained in that location.

While these different paradigms all provided support for the theory, there was one notable omission from the predictions we tested: we did not venture to test the central claim of the theory—that illusory conjunctions would occur with the stimuli we were using if focused attention was prevented. Perhaps the main reason why we hesitated was that we found the claim implausible and counterintuitive. It is certainly true that in everyday life we seldom experience "illusory conjunctions." Do they actually occur? Two factors could explain their rarity: First, in natural contexts there are many known constraints limiting which features can sensibly combine: grass is normally green, the sky blue or gray; people have noses, while tables do not. Perception maps sensory data into expected "frames," selecting the conjunctions that make sense. On the
other hand, in less predictable situations, a rapid glance or a diffuse and inattentive gaze may in fact generate many hallucinatory couplings, which we do not always take the time to verify. A friend walking in a busy street “saw” a colleague and was about to address him, when he realized that the black beard belonged to one passerby and the bald head and spectacles to another. In a class demonstration, we flashed slides of colored magazine pictures and asked the students to write down what they had seen. A number of apparently illusory conjunctions were reported, including a blue striped chair for a blue flowered chair on a black and white striped rug, a yellow lamp for a silver lamp near a yellow chair and a red-haired woman for a black-haired woman wearing a red blouse. Both in everyday life and in experiments using these complex displays, it is difficult to control guessing biases and to be sure that the reports reflect a genuine perceptual recombination of separately identified features. Specially controlled situations must be set up in the laboratory to demonstrate that illusory conjunctions do in fact occur and to investigate their frequency and nature.

We decided to try a pilot experiment. We prepared displays containing colored letters which we flashed briefly in a tachistoscope, attempting to report all that we saw. We stopped after a few trials, discouraged and convinced that we were seeing correctly the one or two items that we were able to report. Both of us found it hard to believe that we had written down almost as many illusory conjunctions as correct items. Having clearly seen a pink T, it was hard to accept the evidence on the card, which showed a pink X and a green T.

The first experiment describes an attempt to see whether others shared our experience and to explore the effect more systematically. This experiment used a whole-report task, and was concerned primarily with the phenomenology of illusory conjunctions. One question we would like to answer is whether they are true illusions, in the sense that they are “seen” as part of a subject’s conscious experience. Unfortunately, terms such as “see” and “experience” are poorly defined, making it difficult not only to ensure that subjects know what question is being asked, but also to relate their reports to any systematic model for the data. Experiments II, III, and IV therefore used a number of converging operations to confirm that illusory conjunctions are genuine perceptual phenomena and cannot be attributed to memory failures, response biases, or interchanges of verbal labels. Experiments IV and V also explore further the nature of the feature encoding within which the errors arise and link them to attention overload rather than to sensory failures.

A FRAMEWORK FOR PERCEPTION

Since we hope to relate conjunction errors to conscious experience as well as to information-processing mechanisms, we offer a flow chart of
the processing operations we hypothesize and try to specify which of these operations affect conscious awareness (see Fig. 1). In doing so, we will speculate beyond what we would wish to defend as the essentials of the theory. We should also make it clear that we are discussing only one type of attention. We do not believe that all the limits on performance that have been attributed to attention can be subsumed by a single mechanism (Treisman, 1969). In this paper we use the term “attention” to refer to selective or focused scanning of competing external stimuli, the type of attention involved in “filtering” paradigms (Broadbent, 1958). We distinguish this from the priming or expectancy effects studied, for example, by Laberge (1973) or Posner and Snyder (1975), which selectively preactivate particular perceptual decisions. We also distinguish it from the limited span of conscious experience; in the present theory, focused attention is used to construct the representations which may, but need not, then become “available” to subjective awareness.

The basic claims made by Treisman and Gelade (1980) were (1) that there is an early stage of perceptual processing at which separable dimensions are coded, independently of each other, and form regions defined within separate maps by the presence of particular sets of features; (2) that this stage is followed by a feature-integration stage, where focused attention mediates the formation of perceived objects. Feature registration at the first stage occurs automatically, without attention, and in parallel across the spatial display, although it is subject to “data limits” on accuracy (Norman & Bobrow, 1975) and also to mutual interference be-

![Flow diagram of processing operations involved in the perception of objects.](image-url)
tween different features of the same type (e.g., between colors or between orientations) which are simultaneously present. Separate feature maps are formed for each separable feature (color, orientation, size, etc.).

Before any features can be consciously perceived, they must be recombined. Conscious experience, like the real world, is populated not by disembodied orientations, colors, and movements, but by objects and their backgrounds. We may not always be sure what these are, in the sense that we can confidently or correctly identify which real objects are present. Because of data limits at stage 1, there may be only fragmentary or imperfect information available about one or more features; for example a blue X may be coded only as a diagonal line of a greenish-blue hue. We may prefer to report overtly only the features we can confidently classify, rather than to guess, for example, that the diagonal belonged to an X rather than N. However, if we consciously see an object at all, our perceptual experience would necessarily comprise some information (correct or incorrect) about all the dimensions whose values must be instantiated for the physical object to exist in the real world.

The stage at which objects are constructed is the interface where top-down and sensory information combine to specify what will (or may) be consciously seen. Perception, in our view, is the construction of a temporary representation of particular entities, states, and events at particular places and times. It usually involves matching stored, perceptual information from semantic or conceptual memory with the incoming sense data. However, we do not equate perception with the temporary activation of permanent nodes in semantic memory (Shiffrin & Schneider, 1977; Laberge, 1976). Instead, it builds an episodic structure, which is functionally distinct and separate from the semantic store. Thus we extend the distinction drawn for long-term memory by Tulving (1972) to help in modeling immediate perception.

How, then, is feature integration achieved and the perceived object formed? We suggest that there may be three ways in which features are conjoined to form the objects we experience: (1) we may direct attention to the location of one feature and integrate with it the other features present in the same location; (2) we may fit features, detected in parallel, into predicted object frames, without necessarily focusing attention on their external location; (3) when we are neither paying attention, nor expecting any particular object, the world around us need not become quite invisible. In this case, some of the features detected at stage 1 may be randomly conjoined, although others may remain single and thereby doomed to remain unconscious. Illusory conjunctions will be generated both by random couplings and also by route (2) when our expectations mislead us.

The episodic structure constructed in these various ways will usually
(but need not) become consciously accessible, and may, but need not, be stored in episodic memory. It may also, and perhaps independently of consciousness, prime or inhibit other related elements in the semantic store and thus modify the knowledge used in subsequent perceptual processing. Since, on this view, attention limits are separate from and prior to conscious experience, they may affect processing which remains unconscious in the same way as they affect processing which results in conscious experience. Semantic priming may also be controlled directly by features, as well as by their conjunctions. If enough features are present to identify the letters of a word (e.g., RAT) but not their order, they should prime associates like “mouse,” but might also prime associates of other possible conjunctions of the same features (e.g., TAR, ART).

Finally, we should specify more exactly how we envisage the function of focused attention, in the context of object perception. We accept Eriksen and Hoffman's suggestion (1972) that the area or the number of items receiving attention can vary from time to time. The “spotlight” can be spread over a large or small area to encompass one or several items. Our hypothesis is that the focus of attention determines the accuracy with which features are conjoined. If two or more items fall within the spotlight, their features may be interchanged, as may those of unattended items which fall outside the spotlight. Thus illusory conjunctions may arise either in unattended regions or within the focus of attention when it is divided among several items.

**EXPERIMENT I**

In the first experiment we used a primary task to direct attention to selected items in the display and then asked subjects to report what they saw of another set of items labeled secondary. Displays consisted of two black digits placed one on each side of a row of three larger colored letters. Subjects were to attend to and report the black digits and then to describe anything they had seen of the colored letters. Since the colored letters were closer to the fovea than the attended black digits, they should be registered with at least as high acuity. Subjects were advised to spread their attention across the whole display in order to report both digits correctly; the colored letters were therefore presumably included within a broader spread of attention. However, we used brief exposures to reduce the probability that the secondary items could receive the narrowly focused serial scan which would specify their conjunctions of features correctly.

There are three kinds of errors which can be made in reporting displays of colored letters. Subjects may misperceive a single feature, for example mistaking blue for green or N for X. We call these errors “feature errors.” Subjects may incorrectly conjoin correctly perceived features, for exam-
ple mistaking a red O and a green X for a red X and/or a green O. We call these errors "conjunction errors." Finally, subjects may mislocate either a feature or a conjunction of features. We call these "location errors." In addition, subjects may omit either one feature or a complete item. Our main interest was in conjunction errors and the conditions under which they arise.

It is worth pointing out a problem we encountered in designing experiments on illusory conjunctions. The difficulty comes from attempting to combine two conflicting requirements: (1) The theory claims that illusory conjunctions arise when attention is overloaded; we therefore need to present several items and to use brief exposures. (2) However, illusory conjunctions can be formed only from correctly identified features; the more different values are used on any dimension, the more confusable they become, and the briefer the exposure, the poorer the quality of the sensory information. Thus we were forced to trade off the need to load resources against the risk of introducing data limits (Garner, 1974; Norman & Bobrow, 1975). In most of the experiments, we controlled exposure durations separately for each subject in order to produce a feature error rate of about 10%.

A final question studied in Experiment I was what happens when features are repeated within a display? When two green letters are shown, are subjects more likely to interchange their locations than when one is green and one is brown? If features are conjoined solely via their locations, as the theory predicts, this should not be the case. On the other hand, if there is an intrinsic "glue" holding together a shape and the color of its ink, the probability of a feature switching locations might increase when either the colors or the shapes are identical. We included some trials (38%) on which either two colors or two shapes (never both) were the same. The presence of repeated colors and shapes should also increase subjects' freedom to report what they saw on nonrepeat trials. It is an interesting question whether an illusory conjunction can include a feature which is also seen in its correct conjunction, producing the illusion of a repeated color or shape, or whether a feature presented only once must be perceived as part of a single item only, so that if it switches locations, it leaves no "ghost" behind? The introduction of real, repeated features on some trials allowed us to see whether illusory conjunctions also sometimes generate illusory repetitions from nonrepeated features.

Method

Stimuli. The stimulus displays for this experiment consisted of three colored uppercase letters arranged in a row between two black digits. The letters each subtended a visual angle of 0.61° horizontally and 0.76° vertically. The black digits were slightly smaller in size and subtended an angle of 0.38° horizontally and 0.61° vertically. The row of letters subtended 3.27° and the entire row, including the digits, subtended a visual angle of 6.10°. The lumi-
nance of the display and also of the fixation and masking fields was matched at about 16 cd/m².

The letter set from which each combination of three was chosen consisted of the following five alternatives: TSNOX. The color set consisted of five alternatives, also chosen to maximize discriminability; these were pink, yellow, green, blue, and brown. Since the color and shape of any given item could be confused with any of two other presented colors and shapes and two nonpresented colors and shapes, reports based on a random selection of features would produce as many feature errors (i.e., intrusions of nonpresented colors or shapes) as conjunction errors (i.e., interchanges between presented but nonattended colors or shapes).

Colors and letters were balanced across positions, and color–letter combinations were equated for frequency. The 10 different three-letter combinations occurred approximately equally often, as did the 10 different three-color combinations. A set of 40 cards was made in accordance with these specifications and an additional 24 cards were made which contained repetitions of either a color or a letter. Half of these cards contained letter groups in which one of the letters was repeated but three different colors were present; the other half included one color repetition but three different letters. The particular letter or color which was repeated varied across displays, as did the location of the repeated feature.

The 64 display cards were made by hand, using alphanumeric stencils and colored inks on white cards.

A black and white noise mask was made by randomly blackening half of the 2-mm squares of a 126 × 70-mm grid. In the tachistoscope, the mask subtended a visual angle of 5.34° vertically and 9.53° horizontally. This mask was used in all the experiments to be reported. We used a random checkerboard rather than a pattern mask to reduce the possibility that illusory conjunctions would be formed, combining visual features from the mask with those from the display.

Procedure. The format of the displays and the letter and color sets from which the display items were composed were shown to subjects before the experiment. The subjects were instructed to report first the two digits (which they combined into one two-digit number; for example, if the digits were “6” and “4” the response was “64”) and then the positions, colors, and names of any letters they had seen in addition. The positions on the card were referred to as left, middle, and right. We stressed the importance of accuracy on the primary task of naming the digits. We also emphasized that we were interested in what subjects could see when their attention was focused on other items. They were told not to guess, but to report only what they were fairly confident they had seen. They were asked to indicate any reports about which they felt doubtful and these were noted by the experimenter. No feedback was provided.

The stimulus cards were presented in a Gerbrands four-field tachistoscope, and the experimenter recorded the oral responses of the subjects by hand. The experimenter gave a verbal “Ready” signal and then initiated a trial by pressing a button. Each trial was preceded by a black outline rectangle, 98 × 24 mm in size, which marked the area within which the row of display items would appear. Once the trial was initiated, a black fixation dot was shown for 1 sec. The display card was then flashed briefly, followed by the noise mask of 200-msec duration. The return to the rectangular outline signaled the end of a trial.

The exposure duration of the display card was set at 200 msec for the first few trials and then varied according to the error rate of each individual subject. The experimenter tried to elicit an overall feature-error rate (i.e., reports of colors or shapes that were not presented) of about 10% for each subject, by reducing or increasing the exposure duration of the display cards. The feature-error rate was calculated every 10 trials and exposure adjustments were made, if necessary, according to the following rules: If a subject was able to report at least two features, in addition to the primary item, with less than 10% feature errors, the exposure duration was reduced by 20 msec. This was repeated, if necessary, until the exposure
duration reached 100 msec, when decrements of 5 msec were introduced. If an exposure reduction resulted in either an increase in the feature-error rate above 10% or the inability of the subject to report more than one feature in addition to the digits, the exposure duration was increased by the same amount as the previous reduction. Subjects varied considerably in their efficiency; the mean exposure duration across all the trials was 120 msec, with a range across subjects of 168 to 95 msec.

The set of 64 display cards was presented twice to each subject and the order was randomized within and between blocks for every subject. Each subject was given a few trials for practice (a mean of seven).

Subjects. Twelve students, nine women and three men from the University of British Columbia, volunteered to participate in this experiment. Each was paid $3.00 for a 1-hr session.

Results and Discussion

There were very few errors on the primary task, reporting the digits (3%). We seem to have been successful in controlling attention. We cannot, however, exclude the possibility that on some trials subjects were able to attend to one or more letters in addition to the digits. If the capacity of attention fluctuates from trial to trial, there is no way in which this could be reliably prevented.

The main result was that conjunction errors in reporting the colored letters significantly exceeded errors which combined one correct feature with one not present in the display ($t(11) = 5.33, p < .001$). This confirms the prediction that when attention is overloaded, correctly detected features may recombine to form illusory conjunctions.

Table 1 shows the mean number of items reported per trial in each of a number of categories: correct items, illusory conjunctions, feature errors with and without a conjoined correct feature, and single correct features. The data are for nonrepeat trials only. An average of .28 feature error was made per trial (with or without a conjoined correct feature), which repre-

| TABLE 1 |
|------------------|------------------|------------------|
| Mean Number per Trial of Different Types of Responses in Experiment I |
| In correct location | In wrong location | Total |
| Correct items | 0.44 | 0.08 | 0.52 |
| Illusory conjunctions | 0.33$^a$ | 0.06$^b$ | 0.39 |
| Items with one feature correct and the other wrong | 0.11 | 0.04 | 0.15 |
| Single feature correct | 0.49 | 0.24 | 0.73 |
| Single feature wrong | 0.13 | 0.28 |
| Total features correct | 1.81 | 0.89 | 2.70 |
| Total features wrong | 0.28 |

$^a$ In this case, one feature is correctly located and the other moved from another location.
$^b$ Both features wrongly located.
sents 9.4% of all features reported. The presence of feature errors implies that some of the errors labeled illusory conjunctions in Table 1 were also really feature errors; subjects misperceived a color or shape, but the error happened also to be among the other colors or shapes on the card rather than among the colors or shapes not presented. We can use the number of reports which conjoined either one correct feature with one feature not presented or two nonpresented features (the latter averaged only .02 per card) to estimate the true number of illusory conjunctions formed by incorrectly pairing features, both of which were correctly identified. The excess of conjunction errors over conjoined feature errors averaged .24 per trial (.39 minus .15). However, random selection of a shape and a color from the three shapes and three colors presented would result in an appreciable number of correct conjunctions (one pair in three). Thus the reported conjunction errors represent only two-thirds of the randomly conjoined shape–color pairs; the remaining one-third is included in the correct item reports. The estimate of features randomly conjoined should be 50% higher than the observed conjunction errors (i.e., .36 item per trial), and the estimate of features correctly conjoined by focused attention should be correspondingly lower than the number of correctly reported items (i.e., .32 per trial rather than .44).

The large majority of conjunction errors (mean of 71% and median 78%) were made on the first item each subject reported. This is only slightly less than the proportion of the correct items which occurred in the first report position (mean 80%, median 84%). Thus, memory interference from earlier reports of colors and shapes on the same trial is unlikely to be a major factor in producing conjunction errors.

There were considerable individual differences: the mean number of conjunction errors per card ranged from .13 to 1.07, although the median of .35 per trial was close to the mean. The differences may reflect individual variation in the number of items which subjects could process with attention, when exposures were short enough to produce 10% feature errors; in other words, there may be individual differences in the balance between state and process limits (Garner, 1974).

The proportion of correct features which were mislocated (one-third) is the same for conjoined pairs (correct items and illusory conjunctions) as it is overall for all the correct features reported. If no location information had been available, two-thirds would have been mislocated. This suggests that subjects were able correctly to localize some, although not all, of the features they detected. In an earlier experiment, we found that features could be correctly identified without being correctly located (Treisman and Gelade, 1980). We suggested then that single features were initially free of any location information. However, this hypothesis raises problems: for example, it is hard to see how preattentive texture segregation
ILLUSORY CONJUNCTIONS

would be possible without some spatial mapping within a dimension, at least at the boundaries between homogeneous groups. Our claim should perhaps be modified: features within a single dimension may be partly organized within their own spatial map at the preattentive level; however, they are integrated or brought into register with others only serially, by attentive scanning of each item or group in turn.

**Distance and dimension effects on conjunction errors.** Two further questions can be asked of the data: (1) were conjunction errors more likely between adjacent than between distant locations; (2) were colors and shapes equally likely to switch locations? We compared the proportion of reports per trial which wrongly conjoined adjacent features (.23) with the proportion which wrongly conjoined features from the two outer positions (.16). By chance, we would expect twice as many adjacent as distant conjunctions, since two pairs of items in the display were adjacent and only one pair was distant. The results show that subjects were, if anything, more likely to conjoin features from the two distant locations than those from adjacent locations, although the difference was not quite significant \( t(11) = 2.13 \). This was not due to higher accuracy overall at the center location: 60% of the correct items were on the left, 20% in the center, and 20% on the right. An interesting and fairly frequent error was shown when subjects collapsed together one feature from each outer location and saw the resulting illusory conjunction in the center location. It is clear that illusory conjunctions can be formed across an intervening item as well as across unfilled space.

The absence of a distance effect is surprising. We had, moreover, obtained the same result in an earlier pilot experiment, with displays that included more striking differences in item separation. In the earlier experiment four colored letters were presented at the corners of an imaginary rectangle subtending 4.96° horizontally and 2.06° vertically. A colored dot was also presented in the center of each display. The primary task was to report the shape of the colored letter which matched the color of the dot; subjects were then to report all that they had seen of the other three colored letters. The 10 subjects in that experiment reported as many illusory conjunctions recombining features whose centers were horizontally separated by 4.35° and on opposite sides of the fixation point as conjunctions with features whose centers were on the same side of fixation and vertically separated by only 1.3°. Thus quite large differences in spatial separation appear to have little effect on conjunction errors, and their formation does not depend on the features initially reaching the same hemisphere.

The fact that spatial adjacency seems irrelevant to the formation of illusory conjunctions is surprising and important. It conflicts with an account of illusory conjunctions in terms of local, spatial perturbation of
features (Wolford, 1975). However, it also conflicts with his recent finding (Wolford & Shum, 1980) that a tick on a square is more likely to migrate from an adjacent than from a more distant position. An earlier finding by Snyder (1972) also showed an adjacency effect in what might be illusory conjunctions. Snyder's subjects were asked to report one letter from a larger tachistoscopic display when the letter was cued either by its color or by fragmentation or inversion. On some trials, however, subjects reported the letter adjacent to the cued letter. In explaining the effect of distance on feature migrations, our theory predicts that two factors will be relevant: (1) the probability of detecting both of two separate features, which may vary when these are presented at different distances from each other; (2) the spatial distribution of the subject's attention. Whenever subjects are more likely to detect a pair of adjacent than a pair of distant features, conjunction errors must also decrease with distance. Second, the theory predicts that when attention is focused on a subset of presented items, illusory conjunctions should be formed either within the attended subset, or outside it, but not between the attended and the unattended items. In Snyder's (1972) experiment, the adjacency effect could be explained by partial but imperfect narrowing of attention, so that the focus included not only the cued item but its immediate neighbors. In the present experiment, it is likely that neither of these variables was correlated with distance. Further research manipulating detection probabilities and the spread of attention may be needed before any firm conclusion can be reached about the spatial variables affecting illusory exchanges.

Since subjects reported not only the items they saw but also their locations, we were able to distinguish cases in which a color moved to make an illusory conjunction from cases in which a shape moved. We found that shapes were rather more likely to move than colors (.27 compared to .18), although the difference was not significant ($t(11) = 1.37$). The number of shapes mislocated when reported alone (11) did not differ from the number of colors (13). The opposite result was found in the earlier pilot experiment; there colors moved more often than shapes. Thus there is little evidence for any difference in mobility between colors and shapes.

**Wholistic perception of objects?** Are subjects more likely to detect correctly two features of the same object than two features of different objects? If so, this will introduce a bias favoring correct over illusory conjunctions. The question is also of some intrinsic theoretical interest: Feature integration theory claims that objects are initially analyzed into separate features, independently processed. The only reason, then, to expect any correlation between the detection probabilities for the color and shape of an unattended object would be local differences in spatial sensitivity. For example a foveal color and a foveal shape might both have higher probabilities of detection than more peripheral features. Given the
number of pairs of features correctly detected on any trial, we can predict the number of pairs expected by chance to come from the same item, assuming that the features seen were randomly selected from the three pairs presented. The predicted mean was .65 per trial, while the observed mean was .72. The difference is small and did not reach significance ($t(11) = 1.85, p < .10$); it is certainly less than would be expected on a wholistic model of object recognition (e.g., Lockhead, 1972), in which subjects should perceive both features of an object or neither.

Moreover, of the .72 pair of features correctly reported from the same object, .21 was perceptually split and reported in separate locations, often conjoined with a different feature. So, on a substantial proportion of trials, both features of a single item were seen, but not as an integral, perceptual whole. We cannot account for all of the illusory conjunctions which our subjects experienced by assuming that single features of different items were seen and then conjoined by default, in the absence of their original partners.

These findings are evidence against one interpretation of wholistic perception—that in which an external physical object is seen as a whole or not at all. We have, however, proposed another sense of wholistic perception, in which subjects have a strong tendency to conjoin features, regardless of their origin, and consciously experience only percepts in which both features are specified. Do the data support this speculation? Given the number of shapes and colors correctly identified on any particular trial and the three possible locations in which they could be seen, we can predict for that trial the number of correct and illusory conjunctions expected if subjects had no conjunction information at all: the chance predictions averaged .145 and .29, respectively, with observed rates of .52 and .39. In terms of the present theory, the excess of correct items suggests that the digits did not fully preempt attention; in addition, however, there does seem to be a tendency to conjoin features rather than to report them separately, since the observed illusory conjunctions also exceed the number predicted by random allocation of features to positions. We cannot, in this experiment, rule out the possibility of a response bias to conjoin features rather than reporting them separately. However, Experiments II and III make it very unlikely that illusory conjunctions result mainly or only from a response bias of this kind, since they occur with recognition and with simultaneous matching tasks as well as in free recall. If we are correct in our claim that conscious experience of disembodied features is impossible, the tendency to conjoin, a kind of "feature magnetism," may be an inherent property of the perceptual system.

**Feature repetitions.** The main purpose of trials with a repeated feature was to determine if the probability that a particular feature would move from one location to another depended on whether the other feature was
the same or different in the two locations. For example is an X more likely to exchange locations with an O when both are green than when one is green and one is blue? There was little evidence that this was the case. With displays consisting of 2 features in each of three locations, there are 12 possible single feature moves; of these, 2 would go undetected on repeat trials (the interchange of the two identical features), 2 represent the interchange between the different features of the otherwise identical items, and the other 8 are interchanges between items which differ on both features. The predicted ratio of 1 to 4 for exchanges between otherwise identical and between different items is exceeded by the data: .07 and .16 per trial, respectively. However, the departure from the predicted ratio is due to two subjects only and is not significant across the group as a whole. Thus there is little evidence that illusory interchanges of features are more likely to occur between locations which share another feature than between locations in which the other feature differs. The question is explored more fully in Experiment IV.

There was also little evidence of illusory repetitions: when a feature moved to form an illusory conjunction, it was also seen in its correct location on only 7.7% of trials (most of these instances being due to one subject). In general, features in this experiment appeared to move in an integral way, leaving no trace or ghostly replica behind them. However, Prinzmetal (in preparation) has found frequent "ghosts" when illusory conjunctions are formed between local components of shapes (a diameter switching from one circle to another) rather than between values on different dimensions, as here.

Confidence and subjective experience. A final question concerns subjects' confidence in their reports. Did they really "see" the illusory conjunctions, or were they simply guessing? Did they identify the colors and shapes separately and then guess that they were conjoined, or were they describing a genuine phenomenological experience? No conclusive answer can be given, but the evidence seems consistent with our own conviction, from the pilot experiment, that on some trials at least, illusory conjunctions have the character of perceptual experiences. We encouraged subjects not to guess, but to describe what they saw or thought they saw. They seldom used the "doubtful" category, applying it to only .31 item per trial, of which .04 was correct items, .05 was conjunction errors, .16 was feature errors, and .06 was correct single features. There was therefore little difference in confidence between reports of correct items and reports of illusory conjunctions. When questioned after the experiment, subjects were very surprised to hear that they had made so many conjunction errors. On average only 2.98 features were reported out of 6 presented; this makes it unlikely that subjects were attempting to guess the whole display. Further evidence against a response bias interpretation is given in Experiments II and III.
An incidental finding in this experiment has considerable importance for the question whether conjunction errors are genuine perceptual experiences. When giving subjects their instructions, we told them that the digits would always be black and that they simply had to report their identity. About half the subjects, at some point in the experiment, spontaneously broke off and said something like "Oh, you've started giving me colored numbers." They had not been asked to report the color of the digits, which in fact remained black throughout. These subjects were reluctant to believe the experimenter when told no colored digits were presented and some expressed doubts even when shown the actual cards which they had wrongly perceived.

Conclusions and further questions. The first experiment has demonstrated that illusory conjunctions are experienced on a substantial number of occasions when attention is overloaded and when no a priori constraints can guide the choice of pairings. The emphasis was on demonstrating the effect and exploring its phenomenology. There are, however, some alternative explanations which we should rule out. First, it is important to check whether these results could be due to verbal coding. Subjects were asked to describe the items in words; could they have switched verbal labels rather than feature encodings? Experiments II and III test this possibility, using colored letters in a successive and in a simultaneous matching task, in which subjects responded "yes" or "no" and were not required to label any features verbally. Prinzmetal (in preparation) has also shown what appear to be illusory conjunctions in a task which required subjects to detect a feature rather than to describe an object verbally.

Second, the free report experiment depended to some extent on memory. Although subjects reported only three features, on the average, and most of the conjunction errors involved the first two features reported, it is possible that the feature exchanges occurred in the recall process, some perhaps induced by a response bias to conjoin features when in doubt. The phenomenal descriptions subjects gave made it clear that at least some of the errors were experienced as genuine illusions. However, it would be useful to confirm these subjective reports by demonstrating illusory conjunctions in immediate perceptual matching of items, either to a previously presented probe or to an item in the same display. In Experiment II, we gave a probe item in advance and asked subjects to report whether or not it appeared in the display which immediately followed.

EXPERIMENT II

Method

Stimuli. The display cards for this experiment were the same as those used in Experiment I, excluding the 24 that contained feature repetitions. Each display contained a row of three colored letters between two black digits. For each display card, a set of three probe
cards was made, consisting of one conjunction probe, one feature probe, and one identical probe.

A conjunction probe was constructed by recombining a color and a letter from the corresponding display. For example, conjunction probes for a display that consisted of a yellow N, a green O, and a brown X might be a yellow O or a green X. The positions from which the two features were selected for recombination were varied systematically between displays. One-third of the conjunction probes recombined adjacent features on the left side of the display, one-third recombined adjacent features on the right side of the display, and one-third recombined features from the two outside items.

A feature probe combined one feature from the display with another feature not present in that display. So, a feature probe for the display containing a yellow N, a green O, and a brown X might be a yellow S or a pink O. One-half of the feature probes matched a color in the corresponding display while the other half matched a letter. Again the display position from which the matching feature was selected for the probe was counterbalanced across displays.

An identical probe was exactly the same as one of the items in the display; in other words, a replication of a color-letter conjunction that occurred in the display. The item for this type of probe was selected equally often from each of the three display positions.

All the probe items subtended the same visual angle as the display items and were positioned 0.76° above the center on 4 × 6-in. white cards. The same letter stencil and colored inks were used to make these probes as were used to make the displays.

**Procedure.** The probe cards were separated into three sets, with each set containing an equal number of conjunction, feature, and identical probes. There were therefore more negative than positive trials. However, any response bias this may have induced would apply to feature probes as well as conjunction probes. Each set of probe cards was organized to correspond to a randomized order of the display deck. The deck of display cards was presented three times to every subject, each time with a different set of corresponding probe cards. The first five cards were rotated to the end of the display deck between blocks, to reduce the possibility that subjects would learn the cards. The order of the cards was rerandomized between subjects.

The same equipment for presentation of the stimuli was used in this experiment as in Experiment I. A single trial consisted of the following sequence: a verbal "Ready" signal given by the experimenter, a 1-sec presentation of a probe card, the black rectangle outlining the area within which the display would appear for 1 sec, the display for a variable duration, and finally the noise mask for 200 msec.

Subjects were shown the letter and color sets from which display items were drawn before the experiment began and were given several practice trials to familiarize them with the task. They were instructed to report the digits on the display card and then to state whether or not the preceding probe exactly matched any one of the three colored letters in the display. The importance of accuracy on the digit task was emphasized. Four response categories were used for the matching task to reflect the confidence of the subject's judgment. The categories were "Sure yes," "Think yes," "Think no," and "Sure no."

The exposure duration of the display was adjusted according to the method described in Experiment I to approximate a 20% feature error rate.

**Subjects.** Fifteen students, five men and ten women, from the University of British Columbia, volunteered to take part in this experiment and were paid $3 an hour. They had not previously taken part in Experiment I.

**Results and Discussion**

Table 2 gives the mean probabilities of responses at each of the four criteria, Sure yes, Think yes, Think no, and Sure no. Every subject gave
more Yes responses to conjunction than to feature probes; 13 out of 15 gave more Sure yes responses to conjunction than to feature probes and every subject gave more Sure no responses to feature than to conjunction probes. Misperceived features in this experiment could give rise to correct rejections of conjunction probes as well as to incorrect acceptances of them. Given a blue X probe and a blue S and red X in the display, subjects could correctly reject the conjunction probe either by seeing that its features were wrongly conjoined or by misperceiving the blue as green or the X as N; they could incorrectly accept the conjunction probe either by forming an illusory conjunction or by misperceiving the S as X or the red as blue. This means that we should not, as in the earlier experiment, correct the estimate of randomly formed conjunctions by subtracting from it the false positive rate on features.

The difference between the number of conjunction and the number of feature errors, although significant, is smaller than in the free-recall experiment. This could reflect a contribution from memory failures in the recall task.

Since the task here was perceptual detection, the results confirm our previous conclusion that a substantial number of illusory conjunctions are formed during perception rather than through memory failures. They also rule out the response bias interpretation of the earlier results, which attributed them to a preference for reporting features conjoined rather than singly. Finally, the results are of some importance in that they did not demand verbal coding. Subjects were asked to match one perceptual experience to another and not to label separate dimensions. Thus, they were at no stage forced by the task to decompose their experience into distinct features. The fact that illusory conjunctions were still generated reduces the likelihood that verbal report was responsible for their occurrence in the earlier experiment.

In an earlier study (Treisman, 1977), a similar matching task was used with reaction time rather than accuracy as the dependent measure. Subjects were asked to decide as quickly as possible whether either of two target stimuli (colored letters) exactly matched either of two test stimuli. The stimuli could differ from the targets either in one or more features or only in the conjunction of features. It is interesting, in the present con-
The text, to note that many false positive errors were made, particularly in the condition in which the two test items recombined the colors and shapes of the two targets. In these trials, subjects made a false positive match about 30% of the time. On trials when either two colors and one shape or two shapes and one color matched the targets, subjects made about half as many false positives; here, only one feature move would result in a match, while with all four features matching, either of the two possible moves would have produced an illusory match. Errors were much lower (less than 5%) on trials in which two colors matched and two shapes mismatched, or the reverse. Unlike the present experiment, the displays were not masked and remained available until the subject responded. However, if subjects generated an illusory conjunction before they had time to focus their attention, they could have "seen" the illusory target and triggered a positive response.

**EXPERIMENT III**

Experiment III used another stimulus-matching task. The first aim of the experiment was to reduce further the possibility that illusory conjunctions depend on verbal coding. In Experiment II, although no verbal description was required, the probe was presented before the display, perhaps allowing time for its color and shape to be named and matched to verbal labels encoding the display. In Experiment III, on the other hand, we used a simultaneous matching task with displays of five items presented so briefly as to make verbal encoding of all their features impossible. Subjects were to respond "Same" if any two of the five items were identical in both color and shape, and "Different" otherwise.

The first question of interest was to see whether subjects made a substantial number of false matches in displays in which an exchange of features would generate an illusory conjunction. The second aim was to test more stringently the role of the conjoining process, when the difficulty of feature detection is kept constant. In Experiment II, the feature probes matched items in the display only in one feature, while the conjunction probes matched in two. This could have contributed to the relative difficulty of rejecting conjunction probes. Subjects may sometimes have ignored, or been unable to follow, the instruction to check that the matching shape and color belonged to the same item in the display, and they may have responded partly on the basis of the number of matching features. In Experiment III, we included three conditions in which the number of matching features was the same.

Figure 2 shows examples of the five different types of displays we used. Types a–c all have the same number of matching features (three colors
and two shapes). They differ in the way the matching features are combined with each other or with nonmatching features. In condition a (correct positive displays) there were two identical items; in condition b, a pair of matching items could be generated if a single illusory conjunction was formed (in this example, from the shape of the blue O and the color of the red X or H), while in condition c, two conjunction errors were necessary to generate the illusion of two identical items.

The comparison of displays b and c provides a more stringent test of the conjunction hypothesis than that made in Experiment II. Displays b and c differ only in the way in which an equal number of matching features are conjoined. Any excess of Same responses to type b rather than c can be explained only by a difference in the number of illusory conjunctions needed to generate a false positive match between two color–shape combinations. Moreover, the comparison is loaded, in another sense, in favor of producing Same responses to display type c; here, every item matches at least one other item in one of its features. Display type b, on the other hand, contains a unique item. If subjects took in a random subset of the five items in the display, they would be more likely to sample a set of two or three items which contained no feature matches with type b than with type c. We should emphasize that both these displays can generate illusory conjunctions which would make a pair of identical items. Type c is not a feature-control card like those used in Experiment II. The only

FOUR PAIRS OF MATCHED FEATURES

SIX PAIRS OF MATCHED FEATURES

**Fig. 2.** Examples of different display types used in Experiment III. Mean percentage Yes responses (with standard deviations in parentheses) are shown beside each display.
reason to expect fewer false positive matches with type c than type b is
the fact that type c requires two feature moves to produce a positive
response while type b requires only one. Note, in addition, that the sam-
pling problem discussed in connection with Experiment II will also reduce
the number of conjunction errors reflected in the results, relative to those
reported in free recall. There are 20 possible feature moves for each
display of which only 4 in displays of type b would produce a pair of
identical items. If subjects see one or less illusory conjunctions per trial,
only one-fifth will show up in the data.

Displays of type e were included in an attempt to generate a larger
number of illusory conjunctions. In these displays, any of 8 feature ex-
changes (out of the total of 20 possible) would produce an identical pair of
items, while in display type b, only 4 of the 20 would do so. We would
expect, therefore, a substantial increase in the number of illusory matches
made, if we are correct in attributing these to feature exchanges. If, on the
other hand, they are simply errors in perceiving individual features, they
should be less likely to produce an increased number of false positive
matches. Displays of type (d) contained two pairs of identical items; with
these displays, illusory conjunctions could destroy the objective identity
matches. Thus we predict a certain number of “Different” responses in
type d, more Same responses with type d than with type e, but a substan-
tial number of Same responses with type e as well. On the other hand, if
the conjunction errors in the earlier free recall experiments were due to
verbal coding or to memory loss, we should expect very few false positive
Same responses in this nonverbal, perceptual matching task.

Methods

Stimuli. The displays all contained five colored letters, chosen from the sets X, H, O, S,
and red, yellow, blue, green. The letters, each subtending 1.2° × 1.0°, were arranged at the
four corners and the center of a square subtending 3.3° × 3.3°. There was also a pair of more
peripheral black digits, subtending 1.1° × 0.7° and placed 3.7° to the right and left of the
center.

The six types of display are illustrated in Fig. 2 and were described above. In Fig. 2, the
comparisons of interest all involve the four outer positions; the fifth item (here the central
one) was added because pilot experiments had shown that with four items only, the task was
too easy. Subjects adopted the strategy of locating the pair of matching colors and checking
their shapes. Attention was not sufficiently loaded to produce errors until we added a fifth
item with a matching feature. This extra item was not always in the center; the assignment of
particular features to particular locations was counterbalanced as far as possible across
displays. Figure 2 illustrates the comparisons using displays with two items which match in
shape and three items which match in color; an equal number of displays used three items
which matched in shape and two which matched in color. The particular colors and shapes
which matched were also counterbalanced across the displays. There were 20 examples each
of types a—c and 10 each of types d and e. Since the letters could all be inverted without
changing shape, the number of different displays of each type was effectively doubled.

Procedure. The cards were randomly mixed and the set of 80 was shown four times to each subject, in alternating orientations. The task was to report the two digits (as in the earlier experiments) and also to indicate whether each display contained at least one identical pair of items, which matched in both shape and color. The cards were shown in a Cambridge two-field tachistoscope and were preceded and followed by the random checkerboard noise field. The exposure duration was set initially at 150 msec, and then either increased or decreased for each subject according to his or her accuracy. The aim was to approximate 70% correct across all the displays. The exposure was increased by 10 msec every time an error was made and decreased by 10 msec after two consecutive correct responses. No feedback was given. The mean exposure duration across subjects was 165 msec and the range was 115 to 213 msec.

Subjects. We tested 19 subjects, all students at the University of British Columbia. They included 8 women and 11 men, aged 18 to 30, who volunteered and were paid $4 an hour.

Results and Discussion

One subject was discarded because she was unable to do the task, even at 250 msec; she showed no discrimination between the displays containing an identical pair and those requiring two illusory conjunctions to make a match. She said “Yes” to 43% of type a and 51% of type c. She responded almost entirely on the basis of color matches, and in fact said Yes to considerable fewer type b displays (38%) relative to type c, confirming that a feature-matching strategy was biased against Yes responses to type b.

The mean percentage Yes responses for the other 18 subjects in each condition are shown in Fig. 2. The digits were correctly reported on 94% trials. The critical test for the effect of illusory conjunctions (with feature matches equated) is the comparison of types b and c: the mean difference of 4% was significant \( t(17) = 2.8, p < .02 \). Four subjects had a marked bias to say “No” when in doubt, giving a positive response on fewer than 30% of trials overall. The mean difference between types b and c for the remaining 14 subjects was 5% \( t(13) = 3.0, p < .01 \). All 20 type b cards generated some false positive responses; the range was from 8 to 33 per card, with a mean of 18.2 and a SD of 6.7. We looked for any variable that might correlate systematically with the number of false positives (e.g., the spatial arrangement of the matching features, the colors of the identical shapes, and the shapes of the identical colors) and could find no factor which was reliably associated with a high or low probability of false positive matches. The present experiment is a replication (with minor changes and a new set of cards) of an earlier experiment, in which 15 subjects averaged a difference between display types b and c of 7% \( t(14) = 2.87, p < .02 \). Thus we are confident that the effect, although small, is reliable. Since the only difference between displays of type b and displays of type c
was the fact that a single feature exchange could generate an illusory match in type b, whereas two were required in type c, the difference gives strong support to our claim that illusory conjunctions occur in the absence of verbal coding and of memory delay.

As predicted, a greater number of false positive matches were generated with type e, averaging one in every two or three trials. This is the same frequency as we found for illusory conjunctions with verbal report and memory load in Experiment I. The increase in illusory matches with type e relative to type b is consistent with the idea that they result from feature exchanges. Subjects also missed a large number of identical pairs, particularly when only one was present in the display (50.3%). They may sometimes have missed a match because it was destroyed by an illusory conjunction, rather than simply failed to detect the identity of two items.

In summary, the conclusions from this simultaneous matching paradigm are consistent with those of the free-recall experiment, despite the important differences in the memory and response components of the tasks. They strengthen the hypothesis that illusory conjunctions are frequently formed when attention is overloaded.

**EXPERIMENT IV**

So far the results suggest that colors and letter shapes are separately registered and that, in the absence of focused attention, they may be wrongly recombined to form illusory conjunctions. Are there any constraints on the freedom of movement enjoyed by separable features? For example, could we take the red from a small circle and wrongly re-pair it with the shape of a large and complicated polygon, to fill an area and contour which are strikingly different from those of its physical source? There is an extreme interpretation of separability in which the representation of one separable feature could not be constrained by the value of the others. Redness should be coded in the same way whether its embodiment is large or small, outline or filled in, triangular or circular. If color and size are separable in this sense, colors should be interchanged as freely between large and small shapes as between shapes of the same size. This prediction, however, is highly counterintuitive; for example, it seems implausible that we could at a preconscious level register the presence of green with no reference to its extent or shape. Physiologically there are likely to be more neurons signaling green the larger the green area. However, we are interested here in the functional codes that are synthesized to form perceptual objects. These, we suggest, could be more like discrete labels or propositions ("Red present") than like analog compounds.

The next experiment tests the degree of miscegenation in the illusory recombinations of four different properties—shape, color, size, and out-
line versus filled-in area, which we will refer to as "solidity." We presented four colored shapes again placed between two digits which required primary attention. Immediately after the display, we cued the location of one of the four shapes and asked subjects to report all that they had seen in the cued location. We tried to maximize the discriminability of the features. We therefore used only two sizes, two shapes (triangle and circle) and two ways of coloring (outline versus filled in). However, we also wanted some unambiguous information about the origin of wrongly conjoined features, so we used a different color for each of the four items in the display. This allowed us to see whether color exchanges varied with the identity of the items on other dimensions.

**Method**

**Stimuli.** White cards with four colored shapes and two black digits served as stimuli. The stimulus displays were made by hand, using shape and digit stencils and colored inks. In each display, various combinations of four triangles and/or circles were centered on the corners of an imaginary square, 1.22° from the center of the card. The large shapes subtended 1.3° and the small ones 0.65°. The complete display of four shapes subtended 3.03° x 3.03° if all shapes were large and 2.38° x 2.38° if they were small. Two black digits, subtending 0.45° x 0.72°, were located on either side of the four shapes, each 2.74° from the center of the card. Figure 3 gives an example.

The four shapes in each display were always colored in four different inks: blue, brown, pink, and yellow. They could be outlined or filled in, and large or small; the triangles were equilateral. Nine types of display were designed, and eight examples of each type were made, giving 72 different stimulus displays. The nine types all consisted of four items, each in a different color; they differed also (1) in size only (large or small); (2) in shape only (triangle or circle); (3) in solidity only (outlined or filled in); (4) in both size and solidity, with shape held constant; (5) in both shape and solidity, with size constant; (6) in both size and shape, with solidity constant. The remaining three types of display varied in all three dimensions, with size and shape covarying in (7), size and solidity covarying in (8), and shape and solidity covarying in (9). Whenever a dimension varied in a given display, there were always two items with each value (e.g., two large and two small). Overall, each value of each dimension was presented equally often. The four colors appeared approximately
equally often in each of the four positions, and, for all nine conditions, the contrasting features were distributed equally often across the diagonal, vertical, and horizontal positions.

In order to probe recall of the information contained in one of the four positions, additional cards were made. Single black lines, 0.76' in length, were placed in the periphery of white cards pointing at the position of one of the four colored shapes. Each of the lines was 0.84' from the center of the nearest colored shape, not overlapping any part of the shape, yet clearly indicating which of the four was to be recalled. Seventy-two cue cards were made, with each of the four positions probed 18 times.

Procedure. The stimulus cards were presented in a Gerbrands four-field tachistoscope. At the beginning of each trial, subjects saw a small black fixation dot on a white background. Following a verbal "Ready" signal, the target display was shown for a variable time, adjusted for each subject. It was immediately followed by a mask for 200 msec, then by a cue card which remained in view for 2½ sec. Subjects were informed that their primary task was to report the two digits which appeared on either side of the colored shapes and that only the trials in which the digits were reported accurately would be counted. They were advised to spread their attention across the entire display in order to be successful in this task. Subjects were told that the secondary task was to report, in addition, as much as they had seen of the colored shape in the cued location. They were asked, if they could, to report its color, size, and shape and to indicate whether it was filled in or outlined. It was emphasized that they should not guess when they had failed to see any or all of the information in the cued position and that they should indicate whenever they were doubtful about their report. After each trial, the experimenter recorded all the information reported. No feedback was given.

Subjects were tested on each of the 72 stimulus displays twice, for a total of 144 trials. Three random orderings of the 72 stimuli were made and used on three different groups of subjects. In each of these arrangements, each stimulus display was probed in two different locations. This was done by presenting the cue cards once and then turning them upside down and presenting them once more. Across the three random orderings, each item was probed at least once and half were probed twice.

In all three arrangements, the stimulus displays were presented in blocks of nine. Each block contained one card from each of the nine types of displays, in different random orders. As in the previous experiments, the exposure time was adjusted for each subject individually. The experimenter decided after each block whether to change the exposure duration: If subjects had made no feature errors (reporting a feature that was not presented), the exposure was reduced, by 20 msec decrements down to 100 msec and 10 msec down to 30 msec. If subjects made more than one feature error, the exposure was increased by the same amount. The mean exposure duration was 199 msec with a range across subjects of 151 to 349 msec. The 144 trials were run in one session lasting about an hour.

Subjects. Ten women and two men, aged between 17 and 30, volunteered to be subjects in this experiment and were paid $3 an hour. All were students at the University of British Columbia and none of them had taken part in the earlier experiments.

Results

The digits were correctly reported on 99% of the trials. In the secondary task, subjects normally reported four features for the cued item, omitting on average only 0.13 feature per trial. They rated very few of their responses as "doubtful" (only 0.17 feature per trial of the 3.87 reported; of these, 0.01 was feature errors, 0.08 was correct, and 0.08 was conjunction error). In discussing the results, we will use the display shown in Fig. 3 as an example. The cued location in this example is the top right and the
correct answer is therefore "small pink filled triangle." This is an example of a display in which all four dimensions were varied, with shape and solidity correlated. Like all the other displays with four dimensions varied, it contains one noncued item which differs in two features from the cued item, one which differs in three and one which differs in all four. A number of questions can be asked of the data.

**Illusory conjunctions of size, solidity, and shape.** The first question was whether illusory conjunctions occur with the dimensions of size and solidity as well as with color and shape? On trials when all four dimensions varied, only conjunction errors were possible. On trials when two or three dimensions varied, the subjects could make either a conjunction error or a feature error but not both on each of the dimensions of shape, size, and solidity. Table 3 shows the rates of feature errors and of conjunction errors for the four dimensions. The percentages (throughout this section) are the means for the trials on which each type of error was possible. Significantly more conjunction errors than feature errors were made on all three dimensions ($t(11) = 5.15, 7.73, \text{and } 2.55$, for size, shape, and solidity, respectively). Thus it seems that size and solidity can be interchanged between trials in perceptual processing in the same way as color and shape. An estimate of true conjunction switches is given by the difference between conjunction errors and feature errors. There were significantly more of these on the dimension of shape than on size or solidity ($t(11) = 2.42$ and $2.97$, respectively, $p < .05$) although all occurred on a substantial number of trials.

There were a small number of trials on which subjects reported a complete noncued item, which differed in either three or four features from the cued item. This suggests the possibility that other errors which we classified as illusory conjunctions could really be wholistic reports of noncued items which happened to share more features with the cued item. For example, in Fig. 3, a report of "large brown filled triangle" was classed as a conjunction error of size and of color, but could have been a complete report of a noncued item. We calculated for each type of display the number of possible combinations of one, two, three, and four noncued features which could reflect wholistic movement of a noncued item, and the number which would separate features of one or more noncued items.

**TABLE 3**

<table>
<thead>
<tr>
<th></th>
<th>Size</th>
<th>Shape</th>
<th>Solidity</th>
<th>Mean</th>
<th>Color</th>
<th>Omissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature errors</td>
<td>4.5</td>
<td>5.6</td>
<td>7.8</td>
<td>6.0</td>
<td>26.5</td>
<td>3.3</td>
</tr>
<tr>
<td>Conjunction errors</td>
<td>14.8</td>
<td>24.3</td>
<td>14.8</td>
<td>18.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>10.3</td>
<td>18.7</td>
<td>7.0</td>
<td>12.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Thus, in Fig. 3, wholistic errors would give one report with two features differing from the cued item (large brown filled triangle), one with three and one with four; in addition there are 6 possible nonwholistic combinations which would differ in one feature from the cued item, 11 which would differ in two features, 9 which would differ in three features, and 2 (large brown outline circle and large blue outline circle) with four features different from the cued item. The observed proportions of errors which could be classed as wholistic were close to those expected by chance for responses differing in one or in two features from the cued item (5.6 to 6.6% predicted for one feature error, 2.8 to 2.3% predicted for two feature errors), but slightly exceeded the prediction for responses differing in three or four features (2.2 to 0.7% for three feature errors and 1.0 to 0.3% for four feature errors). The extra 2.2% may in fact reflect errors in perceiving the location of the cue rather than wholistic transpositions of noncued items. The fact that wholistic errors were so rare is of some interest, since it suggests that features were much more likely to be transposed than whole items.

**Heterogeneity, feature errors, and illusory conjunctions.** The second question concerns the effect of the heterogeneity of displays on feature and on conjunction errors. Table 4 gives a further breakdown of the results, now classified by the number of dimensions which varied within each display. (The reason the means in Table 3 are not the same as the means of the data in Table 4 is that a different number of cards contribute to each category in Table 4. For example, only one set of cards allows conjunction errors on size when two dimensions are varying, whereas three sets allow them when four dimensions vary.) Attention load, as defined in relation to feature integration, depends on the number of different items in the display and not on the number of different dimensions on

<table>
<thead>
<tr>
<th>Feature errors</th>
<th>Two dimensions</th>
<th>Three dimensions</th>
<th>Four dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>3.6</td>
<td>6.3</td>
<td>—</td>
</tr>
<tr>
<td>Shape</td>
<td>6.0</td>
<td>4.7</td>
<td>—</td>
</tr>
<tr>
<td>Solidity</td>
<td>7.8</td>
<td>7.8</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Conjunction errors</th>
<th>Two dimensions</th>
<th>Three dimensions</th>
<th>Four dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>12.5</td>
<td>15.9</td>
<td>14.6</td>
</tr>
<tr>
<td>Shape</td>
<td>27.1</td>
<td>22.7</td>
<td>22.4</td>
</tr>
<tr>
<td>Solidity</td>
<td>12.5</td>
<td>15.9</td>
<td>16.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Color errors</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25.5</td>
<td>27.4</td>
<td>24.0</td>
</tr>
</tbody>
</table>
which the items vary. Thus we expect heterogeneity to affect performance, if at all, only by increasing confusability between features. Heterogeneity should have no direct effect on the formation of illusory conjunctions with the present task (although it could, with other displays, affect perceptual grouping, and thus indirectly change the pattern of attention scanning (Treisman, in preparation)). In the present experiment, we found no difference in feature errors between trials in which two and trials in which three dimensions varied \((t(11) = .71)\). We compared conjunction errors on trials on which two, three, or four dimensions varied and again found no effect of display heterogeneity \((F(2,22) = 0.49, p = 0.62)\).

**Similarity constraints on illusory conjunctions.** The main question of interest in the present experiment was whether there are constraints on the formation of illusory conjunctions of features. Within any given display, are subjects more likely to exchange a pair of features between two items which are otherwise identical than between two items which differ on other dimensions? The clearest tests of this question are based on color conjunctions, since each color was present once only in each display. This allows us unambiguously to locate the physical location of each color reported. We analyzed first the displays which varied only in one dimension besides color. The frequency of illusory conjunctions involving color was compared for colors originating from two types of noncued items, those which differed from the cued item only in color and those which also differed on another dimension (size, solidity, or shape). Each display contained one of the first type and two of the second. If subjects were equally likely to select the color of any of the three noncued items, we would expect a one to two ratio of conjunctions involving the color of the otherwise identical item and those involving the color of a different item. The results, if anything, favored color exchanges between items differing also on another feature: means of 7.3 and 18.1\%, respectively.

We carried out the same analysis for the most heterogeneous displays, this time comparing the frequency of color migrations from items differing from the cued item in one, two, or three other dimensions. Again there was no significant difference between the means: 8.7\% for colors from noncued items which differed in one other feature from the cued item, 7.8\% for items which differed in two other features, and 7.3\% for items which differed in all three other features. It seems, therefore, that illusory conjunctions of color are formed as often when the appearance of the color must be altered in extent, solidity, or spatial distribution as when it can retain its original configuration.

**Independence between features within a single item.** Another question we can ask is whether the features of a given item are switched independently of each other, or whether they tend to cohere and to move together.
Do noncued features come from the same noncued item, when more than one is moved to the cued location, or are they equally likely to come from different noncued items? In this analysis, we excluded the small number of transposition errors of whole items differing in three or four features from the cued item, since these may reflect errors in cue location. We look first at illusory conjunctions which include color as one of the two noncued features. If errors on both color and another feature (e.g., shape) were selected at random from the three noncued items, they could make any of three possible pairs, two of which are combined in noncued items and one of which is not: for color and shape in Fig. 3, these would be blue circle and yellow circle (both presented) and brown circle (not presented). We would therefore predict two cases where the incorrect features were conjoined in a single noncued item to each one case in which they came from different noncued items. The observed rates for all pairs including color were 13.5% for conjoined pairs and 4.7% for separated pairs. For color and shape, the ratio was exactly that predicted by random selection of features (4.8 to 2.5%), but for color with solidity and color with size, there was some dependence (8.7 to 2.2%).

With dimensions other than color, we must consider triples of features rather than pairs, since all possible conjunctions of pairs were present on every trial. With the displays which varied in all four dimensions, we tested whether illusory conjunctions more often involved triples of shape, solidity, and size which were conjoined in noncued items or triples which recombined features from more than one presented item. Three noncued conjunctions of size, solidity, and shape were presented in each display (large filled triangle, small outline circle, and large outline circle in Fig. 3) and four other possible recombinations of the same features were not presented as such (small outline triangle, small filled circle, small outline circle, and large filled circle). One of the presented conjunctions differs in all three features from the cued item (large outline circle) and none of the recombinations does; to avoid confounding accuracy with the test of feature independence, we omitted this item from the analysis. This leaves one presented conjunction for every two possible recombinations. We can then test whether the ratio of responses which match noncued conjunctions to those which recombine noncued features conforms to the 1-to-2 ratio expected if the choice were made randomly. The means were 13.7% presented combinations to 22.3% illusory recombinations. This is quite close to the predicted 1-to-2 ratio.

There is another sense of dependence which could be confounded with dependencies in feature migrations: this is dependence in the probability that different features of the same item will be detected and identified. In Experiment I there was a slight tendency for reported features to come from the same item rather than from different items. Since illusory con-
ILLUSORY CONJUNCTIONS

junctions can be generated only from features whose presence in the display has been detected, dependence in detection probability could mask independence in illusory interchanges. Dependence in detection need not reflect wholistic perception; it could arise from local variation in visibility from trial to trial. With size and with solidity, there could also be objective interactions between features, which make one easier to see when combined with another which is also more salient. Subjects were in fact more likely to report illusory conjunctions involving both color and size when the noncued item from which they presumably originated was large than when it was small (6.0% vs 4.1%), and illusory conjunctions involving both color and solidity when the noncued item was filled in than when it was outlined (6.2% vs 3.1%). In both cases there was physically more color to see in the more frequent type of conjunction error. Dependencies of this kind in detection probabilities could explain the minor departures from the predicted independence of feature movements in illusory exchanges.

Spatial constraints on illusory conjunctions. In this experiment, for the first time, we found a tendency for some directions of movement to be preferred over others when features migrated to form illusory conjunctions. In color errors (the only ones whose origin was unambiguous), 10.7% moved horizontally, 8.5% vertically, and 6.4% diagonally. The differences were significant ($F(2,22) = 4.67$, $p = .02$). In the present displays (unlike those of the free-recall pilot experiment) the spatial distances were the same for vertical and horizontal movement. In the earlier experiment, a bias to move horizontally rather than vertically may have been counteracted by the greater distance involved. Alternatively, the directional bias could reflect a task-dependent bias in the shape of the attention field. In Experiment IV, the primary task digits were horizontally aligned and could have induced or enhanced a bias for that direction in the spread of attention, while in the pilot experiment the attended item was in the center of the display. We predict that conjunction errors will occur either between items within the same fixation of attention or between items outside it, but not across the boundary. Thus, if attention is spread horizontally, conjunction errors should occur most frequently between items which are also horizontally aligned.

EXPERIMENT V

The last experiment attempts to strengthen the link between attention and the correct integration of features. In this study we drastically reduced the load on attention, both by eliminating the primary task of digit identification and by precueing instead of postcuing the relevant location. We kept the overall difficulty (measured by the total number of errors) the same as in Experiment IV, but instead of loading attention, we
reduced the exposure duration. This should result in a change from process or resource limits to state or data limits (Garner, 1974; Norman & Bobrow, 1975). The prediction was that illusory conjunctions would no longer be the main source of error, since attention could be accurately focused on the relevant items. Errors should result only from inadequate feature information, so that intrusions of nondisplayed items would be as likely as exchanges with displayed items. Moreover for colors, similarity should become more important in determining errors than it was in the previous experiment and the spatial relations between correct items and intrusions (e.g., horizontal versus vertical versus diagonal) less important.

Method

Stimuli. The same set of cards and the same apparatus were used as in Experiment IV.

Procedure. The procedure was the same as in Experiment IV except for the following changes. The cue was given 150 msec in advance of the display instead of after it. The display was presented immediately after the cue, initially for 100 msec; its duration was then adjusted for each subject following the same rules as in the previous experiment, decreasing in 10-msec steps down to 60 msec and thereafter in 5-msec steps. When the display duration was less than 100 msec, a blank interval of darkness was introduced between the cue and the display, with the correct duration necessary to maintain the total interval between cue offset and display offset at 100 msec. The display duration was never increased above 100 msec. The mean exposure duration in this experiment was 89 msec with a range across subject means of 65 to 100 msec. The display was immediately followed by a 150-msec masking field.

Subjects. Twelve new subjects, six men and six women, aged 18 to 34, were run in this experiment and paid $3 an hour. Nine were students, two were in business, and one was a teacher.

Results and Discussion

The mean feature error and conjunction error rates for the cards on which both were possible are shown in Table 5, which also gives the mean omissions. The total error rate was very closely matched to that in Experiment IV, averaging just over one feature wrong or omitted on every card.

The main interest is in the distribution of those errors across the categories of feature errors, omissions, and illusory conjunctions. The

<table>
<thead>
<tr>
<th>TABLE 5</th>
<th>Mean Rates of Feature and Conjunction Errors in Experiment V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature errors</td>
<td>Size</td>
</tr>
<tr>
<td>12.5</td>
<td>8.2</td>
</tr>
<tr>
<td>Conjunction errors</td>
<td>10.8</td>
</tr>
<tr>
<td>Difference</td>
<td>-1.7</td>
</tr>
</tbody>
</table>
change from divided attention to brief exposure induced a clear shift from conjunction errors to feature errors and omissions. The overall difference between conjunction and feature errors is no longer significant in the precued experiment ($t(11) = 2.06$), which makes it likely that almost all the errors of commission were actually misperceived features, half of which happened also to be on the card. An analysis of variance showed no significant difference in the overall number of errors between the precued and postcued experiments ($F(1,22) = .35$), but a highly significant interaction between cue condition and type of error, feature versus conjunction ($F(1,22) = 51.16, p < .0001$). There was also a significant interaction between feature and conjunction errors and the dimension (size, shape, or solidity) on which the error was made ($F(2,44) = 13.09, p < .0001$). The shape dimension had fewest feature errors and most conjunction errors. The interaction confirms that these two kinds of errors are different in origin.

With the color responses, feature errors cannot be distinguished from conjunction errors. The total error rate is similar in the two experiments, but the theory predicts that the source of the errors should shift from conjunction exchanges in Experiment IV to feature errors in Experiment V. One way of testing this prediction is to look at confusability effects in the two experiments: If subjects misperceive the cued color without being affected by the colors in the other locations, the distribution of errors should be determined by the similarity of the correct and incorrect colors and not by their spatial relations in the display. On the other hand, if the errors were due solely to conjunction exchanges with colors in other locations, we would expect the same spatial biases as in Experiment IV, but a random distribution across the particular colors (pink, blue, brown, and yellow) which were exchanged. We look first at color confusability effects: we drew up confusion matrices for the color errors in the two experiments and calculated the numbers expected by chance in each cell as follows: given the total number of errors on a particular color, we assume that every other color will be substituted for it in proportion to its share of the total erroneous reports. We can then calculate $\chi^2$ for deviations from this zero correlation model for the precued and for the postcued conditions. $\chi^2$ was considerably higher in the precued than in the postcued condition (30.8 compared to 9.3) and was significant only for the precued condition. This is what we would expect if the precued errors (those made with focused attention but poor stimulus information) were simply misperceptions of the cued item, while a substantial proportion of the postcued errors were illusory conjunctions formed by spatial exchanges with other items, regardless of the particular colors involved.

We expect spatial constraints, on the other hand, only when the color errors are due to illusory conjunctions. In the postcued condition (Ex-
periment IV), we found significant differences in the pairs of spatial locations between which color exchanges occurred. The same analysis for the precued condition (Experiment V) showed no significant effect of the display location of the wrongly reported color: it was horizontally removed in 7.3% of trials, vertically removed in 7.9%, and diagonally removed in 7.5%. The systematic biases we found in the postcued or divided attention condition (10.7% horizontal, 8.5% vertical, and 6.4% diagonal) disappeared when attention was precued. Thus we find an effect of spatial relations in the divided attention condition and an effect of color confusability in the focused attention condition. This dissociation supports our belief that most of the color errors in the postcued, divided attention condition were illusory conjunctions, involving incorrect re-combinations of separately perceived features, and that attention load plays a significant part in determining their occurrences.

GENERAL DISCUSSION

The main conclusions which emerge from these experiments are: (1) when attention is loaded, subjects make many conjunction errors involving each of the different dimensions we tested. While color and shape seem to be slightly more prone to these illusory exchanges than size and solidity, all four dimensions appear, by this criterion, to be separably coded and can be wrongly recombined. (2) At least some conjunction errors are consciously and confidently experienced as perceived physical objects rather than reflecting simply guesses in the absence of information. (3) The number of illusory conjunctions can be increased while feature errors decrease. Illusory conjunctions occur when attention is loaded (resource limits), while feature errors primarily reflect poor visibility (data limits). This differing vulnerability to experimental manipulations supports the hypothesis that the two types of error reflect different functional levels in the perceptual coding of objects, and that attention is involved primarily in the resynthesis of separately identified features. (4) The illusory conjunctions are not dependent on verbal mediation. They occurred both in a successive matching task, in which the only responses were Yes and No, and no decomposition of the items into features was required, and also in a simultaneous matching task, in which the exposure was probably too brief to allow naming of more than one or two features. (5) Nor do they depend on failures in recall, since they occurred both when the subject had advance knowledge of a target item and with simultaneous presentation of all the relevant items. (6) There seems to be little effect of distance, as such, on the frequency of illusory conjunctions. Such spatial constraints as we found may be mediated by attentional grouping rather than by physical distance (Prinzmetal, in preparation). (7) To a first approximation, each feature appears to be coded as an independent entity
and to migrate without constraints from the other features of its source and destination. An object can be as confidently seen when its conscious representation is generated from the color and size of one object and the shape and solidity of another as when it veridically matches the features of a physically presented stimulus. Features are exchanged as freely between objects which differ maximally as between otherwise identical objects, even though this usually requires some change in the conscious representation of the migrating feature. For example, moving a color between objects of different sizes or between one outline and one filled object must also change the spatial distribution of the color. The results were consistent with the conclusion that the only dependencies between the different features of a single object were those which affected their probability of being detected. Once a feature was detected, it appeared, in most cases, to be free to move, with or without its partners, and to recombine with one or more features from other sources. Minor departures from complete independence occurred mainly in combinations involving solidity with color or size.

The implications of these conclusions, if we accept them, are quite far reaching. They suggest that the internal representation on which conscious experience depends contains discrete labels of values on each dimension separately. The whole object must be resynthesized from a set of these feature labels. If in a brief glance only the labels "blue," "small," and "triangle" are registered, we supply our conscious image with the correct quantity of blue coloring to fill the specified area, regardless of how much was initially presented: moreover, we use it to color the area within the specified shape, again regardless of whether this matches the shape that was originally blue. This hypothesis involves an extreme interpretation of the notion of separability and places conscious seeing at a greater remove from the physical stimulus than we might intuitively assume.

What, then becomes of Gestalt properties, of the ways in which the whole is more than the sum of its parts? When features are conjoined in a real external object, new "emergent" features are sometimes formed (Pomerantz, Sager, & Stoever, 1977; Pomerantz, 1981). For example conjoining parentheses with convex edges outward, (), produces a very different configuration from conjoining them with convex edges inward, ). Pomerantz and Garner (1973) have shown that classification performance with a single parenthesis relevant is strongly affected by the irrelevant parenthesis whenever the two are close together. Similarly, discrimination of \ from / is facilitated by the addition of the same two lines, L, to form Δ on the one hand and L on the other. Pomerantz attributes these effects to new emergent features, such as closure with the triangle.

It would be interesting to know whether such new features also emerge
from illusory conjunctions. The model outlined in Fig. 1 suggests that they should not, since conjunctions are formed at a level beyond that at which features are detected. If this is the case, there should be constraints on which illusory conjunctions are likely to be perceived. Suppose, for example, that a vertical, a horizontal, and a diagonal line are detected at the feature level and conjoined, by chance, as a triangle at the object level. The decision that a triangle is actually present, as a candidate for conscious perception, would then be based on conflicting evidence: the appropriate lines would be conjoined, but emergent features, such as angles and closure, would be missing. If these speculations are valid, they predict that illusory conjunctions which, in a physical stimulus, produce clear or salient emergent features, should be less likely to be consciously experienced than illusory conjunctions which do not.

The particular dimensions which we studied may produce few emergent characteristics. "Circular yellowness" and "outline triangularity" are not compelling as wholistic or integral properties. The features for which we found least independence (solidity and size with color) both produce a combined variation in the total amount of color. Color–shape conjunctions are unlikely to produce emergent features, and this may explain the greater frequency with which they gave rise to illusory objects in our experiments. Emergent features are more likely to result from combining different elementary components of shapes (e.g., lines, curves, angles), and this may make it more difficult to generate illusory conjunctions within the dimension of shape than across different nonshape dimensions (e.g. size, brightness, movement, color). On the other hand, the separation and rigid temporal ordering of the two stages shown in Fig. 1 could be misleading. There could, for example, be recursive loops, allowing feature processing of internally generated conjunctions as well as of sensory data. This issue awaits further research.

REFERENCES

ILLUSORY CONJUNCTIONS


(Accepted July 1, 1981)