

Close Binding of Identity and Location in Visual Feature Perception

James C. Johnston
NASA-Ames Research Center

Harold Pashler
University of California, San Diego

The binding of identity and location information in disjunctive feature search was studied. Ss searched a heterogeneous display for a color or a form target, and reported both target identity and location. To avoid better than chance guessing of target identity (by choosing the target less likely to have been seen), the difficulty of the two targets was equalized adaptively; a mathematical model was used to quantify residual effects. A spatial layout was used that minimized postperceptual errors in reporting location. Results showed strong binding of identity and location perception. After correction for guessing, no perception of identity without location was found. A weak trend was found for accurate perception of target location without identity. We propose that activated features generate attention-calling "interrupt" signals, specifying only location; attention then retrieves the properties at that location.

Many important tasks require people to determine object identities and their locations, what Sagi and Julesz (1985) have called the "what" and "where" of visual perception. What we do with information about the identity of an object depends upon where it is located. Not only do overt responses to objects need to be directed to the appropriate location, but also further perceptual analysis of objects typically requires controlling the direction of eye movements or the locus of spatial attention. Achieving proper binding of identity perception and location perception is therefore a critical function in human visual perception.

From a commonsense perspective, it would not appear that the perceptual binding of object identities and their locations is problematic. Under ordinary viewing conditions, people typically succeed in perceiving objects in at least approximately their correct locations. Furthermore, that success has a clear basis in the physiology of the visual system. Localization in vision begins with a set of receptors sensitive to information from different directions: thus, the initial registration of sensory information already achieves a form of location coding. As one continues further into the visual system, coding of location in retinotopic coordinates persists through area V1 and into subsequent areas (cf. Cowey, 1985).

So if people are generally successful in binding perception of identity and location, and visual-system physiology provides a firm basis for that success, what problem remains? The problem is that the physiological evidence for retinotopic feature maps may only be telling us about processing at a very early level, corresponding to what has been called preattentive processing (Neisser, 1967; Treisman, 1985; but see Moran and Desimone, 1985). There remains, however, the problem

of transferring information to more central, attentional processes that can respond to task goals, control overt responses, store information in memory, and so on. How this transfer of information to central levels is accomplished remains largely a mystery. In particular, we do not know whether information about identity and location are subject to decoupling during transfer. The fact that identity-location binding appears to be quite reliable under ordinary conditions is not revealing. When viewing time is unconstrained, failure in the transfer of identity or location information would be unlikely, even if two entirely separate transfer processes were involved. For this reason, we restrict our review of the literature to studies using brief masked displays, which, by limiting viewing time, provide an opportunity to dissociate the transfer of identity and location information.

Baron (1973) presented subjects with brief masked displays of a pair of letters. Subjects pressed one of four keys to jointly register two binary forced-choice judgments. One judgment was of the side of the display on which the letter pair appeared; the other judgment (which varied across experiments) depended on the identity of the letters. With easy identity judgments (exact identity matches, or discriminating Os from Xs), performance was significantly above chance when the location judgment was wrong. With a more difficult identity judgment (whether the pair of letters rhymed), performance was only at chance when the location judgment was wrong. Baron concluded that a perceptual dependence model (identity can be judged correctly only if location can be judged correctly) could be rejected for the easy identity judgments, but was supported for the difficult identity judgment. Baron proposed a sequential processing version of perceptual dependence in which stimuli must be located before their properties can be processed. Baron noted that his results could also be explained by assuming that less stimulus information is required for the location judgment than the identity judgment.

Logan (1975b) used a tachistoscopic masking paradigm in which one of four characters was displayed in one of four locations. Instead of reporting location and identity on the same trials, subjects ran in four blocked conditions judging target location (L), target identity (I), target location given its

This research was supported by funds from the NASA-Ames Research Center, Moffett Field, California, under Interchange No. NCA2-225.

We are grateful to D. Eggleston and S. Kelly for help in data collection and to C. Sanguinetti and R. Edmiston for help in preparing the manuscript.

Correspondence concerning this article should be addressed to James C. Johnston, Mailstop 262-2, NASA Ames Research Center, Moffett Field, California 94035.

identity (L|I), and target identity given its location (I|L). Logan tested two hypotheses: (a) that identity and location are perceived independently, and (b) that identity perception is contingent on location perception. Logan argued that the independence hypothesis predicts that condition I and condition I|L should produce the same performance. Because the data showed that performance was higher for I|L than for I, he rejected independence. Logan argued that the contingency hypothesis predicts that the probability of correctly judging identity should equal the probability of locating the target multiplied by the probability of identifying it contingent on locating it: that is, $P(I) = P(L) * P(I|L)$. Logan found that this prediction was confirmed when targets (the letters a, o, e, or c) differed substantially from nontarget elements (either blanks or dots) but disconfirmed—obtained $P(I)$ was higher than $P(L) * P(I|L)$ —when targets and nontargets were more similar (letter Xs).

Logan's (1975b) method has the appealing virtue of requiring subjects to report on any trial only identity or location, but not both, thus sidestepping the problem of how to jointly indicate two responses. Unfortunately, the validity of the method is open to serious question. Particularly questionable is the hidden assumption that performance on I|L measured directly is equivalent to performance on the I|L process presumed to follow a preliminary locating process in the I task. One problem is that, with processing time restricted by a masking field, it hardly seems likely that the I|L process could fail to be harmed by the time lost in the preliminary locating process. There also may well have been a selection artifact because performance in the I|L condition includes all trials, whereas the presumed I|L process in the I condition operates only on trials in which locating was successful. Conditions for this subset of trials are likely to have been especially favorable. A third problem (noted by Logan) is that subjects could fixate the target location in condition I|L but not in condition I. That the contingency hypothesis fared as well as it did may be due to these problems offsetting each other. The first and third problems result in the I|L condition overestimating implicit I|L performance, whereas the second problem results in an underestimate.

Logan (1975a) argued that if identity perception were contingent on location perception, any factor that harmed the latter should also harm the former. Logan performed an experiment showing that differences in target-distractor discriminability affected locating but not identifying (naming). He therefore rejected the contingency hypothesis, concluding that "locating and naming processes function independently." Unfortunately, Logan's (1975a) basic premise is flawed. It is not a general truth that if B is contingent on A then manipulations that affect A must also affect B. The manipulation may instead increase or decrease subsets of A that do not contain any of B. That a mammal can read is contingent on it surviving infancy. Nevertheless, one can easily imagine manipulations (e.g., altering infant mortality in cats) that would decrease the number of mammals that survive infancy without decreasing the number that could read. In Logan's experiment, suppose that whenever there is sufficient visual information for identifying a stimulus, there is also enough for locating it. It remains possible for a manipulation—such

as degree of target-distractor similarity—to alter the likelihood that information *insufficient* for identifying would still permit locating.

Krumhansl and Thomas (1976) required subjects to locate and identify a target letter (from among A, E, I, O, and U) in a row of background characters (dollar signs). Krumhansl and Thomas's data were well fit by a model treating locating and identifying as independent processes. One problem with Krumhansl and Thomas's paradigm is that a location report required selecting one of 11 locations in a linear array, which would appear to invite nonperceptual location-reporting errors at a very late processing locus. In fact, location errors were virtually absent for the middle positions adjacent to a fixation X available as a landmark. This study is also open to the objection that the information needed to identify the target may have differed from the information needed to locate the target among dollar signs.

The Different Properties Problem

Each of the studies reviewed so far has (from our perspective) a common difficulty, which we will call the *different properties problem*. We wish to investigate failures in binding perception of the identity of a stimulus object and perception of its location. In doing so, it is critical to distinguish binding failures from other reasons why accuracy and location judgments might differ in accuracy. The different properties problem arises when subjects can determine which stimulus item is the target (the discrimination critical for a location judgment) using different properties than those required to identify it. Consider first an extreme case. The task is to identify the letter C or G embedded in a background array of filled black squares, with a very brief stimulus exposure. Suppose we find that subjects are near chance in identifying the target as a C or a G, but are very good at discriminating which item is the target (mediated by the gross luminance difference), and hence do very well on the location judgment. Thus, location performance is drastically dissociated from identity performance. However, what would such results tell us? We argue that nothing has been revealed about the binding between the perception of an object's identity and perception of its location. What we have found are conditions under which it is very easy to discriminate one property (luminance, which is sufficient to specify which item is the target, and hence support a location judgment) and very difficult to discriminate another property (a slight form difference required for the identity judgment). Put more succinctly, we have found out that it is easier to locate "blobs" than to identify letters. This example is quite crude, but the same problem will persist in more subtle guises unless a very demanding condition is met: *that the target item cannot be distinguished from background items except by using a stimulus property that would be sufficient to identify it*. This condition is quite difficult to meet.

The different properties problem is especially severe for Baron's (1973) study in which locating the target required only the minimal stimulus information to discriminate it from a blank field—clearly not the same information needed for identification. Some of Logan's (1975a, 1975b) conditions

also had empty nontarget locations. Others used dummy symbols, some of which were not grossly dissimilar to the targets: neither were Krumhansl and Thomas's dollar-sign dummy symbols. However, in none of the studies can a strong case be made that selecting which item was the target required exactly the same information as identifying it. This is only possible if we know that target letters do not share a common property differentiating them from nontarget characters. Avoiding the different properties problem is therefore difficult, because we lack an exhaustive list of usable features. The list of features might even be open-ended, if features can be formed by experience (cf. Dumais, 1979).

Sagi and Julesz (1985) devised a paradigm that might appear to succeed in evading the different properties problem. Subjects viewed brief masked displays of a large homogeneous array of parallel lines with a diagonal orientation. Several targets—either vertical or horizontal lines—were present. Sagi and Julesz found that in a task of counting targets, intended to require locating but not identifying, the number of targets had little effect, indicating that parallel search was possible. However, in a task of determining whether several simultaneous targets were identical, intended to require identifying, number of targets had a substantial effect, indicating that serial search was required. On the plausible supposition that serial search requires focal attention while parallel search is performed preattentively, Sagi and Julesz argued that preattentive processes can locate targets (based on a low-level process detecting local differences from background elements), but that focal attention (typically to a location specified by preattentive processes) is required to identify them. Sagi and Julesz therefore asserted that locating and identifying need not be tightly bound: preattentive processes can locate targets without identifying them.

Sagi and Julesz's (1985) design appears on the surface to provide a viable solution to the different properties problem. Vertical and horizontal targets tilt in opposite directions relative to the nontarget diagonals. The two targets thus appear to be more similar to nontargets than to each other, and share no obvious property differentiating them from nontargets. So it might appear that a target cannot be selected except by using a feature that identifies it. There is another possibility, however. An array of parallel lines may produce a perceptual "flow field"; both vertical and horizontal targets share the common property of interrupting the flow. If "break in flow field" is a visual feature, its detection would be sufficient to specify the target without being sufficient to identify it. Whether correct or not, this hypothesis underlines how difficult it is to avoid the different properties problem.

Treisman and Gelade (1980) offered a different perspective on the binding of location and identity perception. Treisman and Gelade proposed that the binding problem has a different answer depending on whether or not a target is specified by a conjunction of features. In visual-search experiments with conjunction targets, Treisman and Gelade found large effects of stimulus set size and approximately a 2:1 ratio of slopes for negative and positive trials: they concluded that conjunction targets can be located only by serial application of focal attention (but see Pashler, 1987). Treisman and Gelade proposed that the act of focal attention necessarily fixes at least

the approximate location of stimuli, so that for conjunction targets locating and identifying should be tightly bound. Experiments VIII and IX, which required reporting both the location and identity of conjunction targets, found extremely tight binding: When location reports were not even approximately correct, identity reports were at chance accuracy and vice versa.

Treisman and Gelade (1980) also used a disjunctive feature-search paradigm in which each of the two targets differed from the background along a different single dimension (e.g., color and form). Feature search produced only small effects of stimulus set size and approximately 1:1 ratios of slopes for negative and positive trials. From these results, Treisman and Gelade concluded that feature search can be accomplished in parallel at a preattentive level.

Of particular interest are the disjunctive feature-search results from Experiments VIII and IX in which subjects jointly reported both target location and target identity. Background items were pink Os and blue Xs, and targets were either orange (in the shape of O or X) or the letter H (in either pink or blue color). Thus, one target differed from the background only in color and the other differed only in form. This stimulus set appears to provide an excellent solution to the different properties problem, because it hardly seems likely that the color orange and horizontal line orientation (or any form feature in H) share a common property not in the background that would specify the target without permitting its identification. This conclusion is especially compelling because color and form have been shown by various lines of evidence to be separable perceptual dimensions (Shepard, 1964; Garner, 1974). With this paradigm, Treisman and Gelade found a striking decoupling of identity and location performance. Even when location was entirely wrong, identity was correct far more often than chance (.678 in Experiment VIII and .748 in Experiment IX using longer exposures; chance level of .5). Although the results of the converse analysis were not clear-cut, Treisman and Gelade summarized that "location responses were generally at chance when the target was wrongly identified" (p. 131). The combination of these results led Treisman and Gelade to conclude that "we can detect and identify separable features in parallel across a display... locating any individual feature requires an additional operation." This conclusion is just the opposite of that reached by Sagi and Julesz (1985) in their study using homogeneous backgrounds.

The mechanism for identification without location in feature search has been subject to revision by Treisman and her associates. Treisman and Gelade (1980) proposed that features can be extracted preattentively without any precise marking of location: they argued for "registration of unlocalized features in separate maps" (p. 131). The notion of a map that does not store location is puzzling, and appears inconsistent with physiological evidence about early visual feature extraction. Treisman and Souther (1985) and Treisman and Gormican (1988) rectified this problem by arguing that location is stored in feature maps after all, but that this information can be lost during transfer to central attentional processes. More specifically, they proposed that it is possible to read from a feature map a pooled response, a composite measure

of the total activity level for one feature type within the attended region. A pooled-response signal provides no information about the location of the feature within the attended region. This model explains not only why simple features can be identified without being located, but also why detection of feature presence is easier than feature absence (Treisman & Gormican, 1988; Treisman & Souther, 1985).

Although the paradigm of disjunctive search for features on separable dimensions provides a good solution to the different properties problem. Treisman and Gelade's (1980) experiments are subject to two other critical problems. Most important is what we will call the *negative-information problem*: If subjects can tell which target was more likely to have been seen *if it had been present*, then even when they cannot find the target they may be able to guess its identity better than chance. Suppose a subject suspects that the two types of targets occur equally often, but notices that he or she is seeing more orange targets than Hs. The subject can now infer, on trials in which no target is detected, that the actual target was more likely to have been H than orange. This negative-inference strategy will support better-than-chance guessing for identity even though location responses will still be at chance (except as noted shortly). Later we describe a formal model that embodies this guessing strategy.

Note that the negative-information problem arises regardless of how the subject knows which target was less likely to be detected. In the previous example, this is accomplished by observing which feature target is perceived less frequently on average. Such knowledge might also be gotten from direct introspection trial by trial. For instance, the subject might know on a particular trial that more attention was paid to orange targets, so that a detection failure was more likely for an H. The problem is difficult to avoid completely. Note that analogous problems can in principle contaminate location reporting, if subjects can tell in which location a target would be more likely to be detected. (Fortunately, there appears to be no evidence for such behavior in either Treisman and Gelade's data or our own.)

Treisman and Gelade's (1980) results also are subject to what we will call the *location-reporting problem*. We wish to make inferences about perceived location, but what we measure is necessarily only *reported* location. Location perception is accomplished in internal perception coordinates, but measured only after it is translated into experimenter-defined coordinates. If we mistake errors in translation for errors in location perception, our analyses will be misleading. We noted earlier that Krumhansl and Thomas's (1976) use of long linear arrays of characters would appear to increase this risk. Treisman and Gelade (1980) used displays with two rows of six letters each (with targets in only the middle four locations). This is an improvement, but the middle positions of each row are still not anchored to any usable landmarks. Treisman and Gelade also used a poststimulus masking field that apparently did not mark the possible positions, since it "consisted of colored segments of the target and distractor letters scattered at random over a rectangular area slightly larger than that of the letter array." Consider the plight of the subject who, after a brief exposure, has correctly coded both the identity of the target and its location (in internal coordinates). The subject

must translate those coordinates into a stimulus position, but the positions cannot be determined from the masking field. Thus, the subject may be required to encode, while the stimulus array is still available, not only the target location itself, but the relation of the location to the structure of the stimulus array. This will lead to errors in reported location, inflating the experimenter's estimate of identity-location decoupling.

The Present Approach

The present study attempts to determine the degree of binding of location perception and identity perception while minimizing the three problems with previous research. To avoid the different properties problem, we adopt the disjunctive feature-search paradigm of Treisman and Gelade (1980), with a color target (orange) and a form target (H). To minimize the location-reporting problem, the geometric arrangement of stimulus locations provides strong anchoring of all target locations, and the same display geometry is used in the masking field. To deal with the negative-information problem, we attempt to equalize the difficulty of the two feature targets, and (in Experiment 1) allow subjects to make a "no-target" response rather than forcing an identity guess. We also include trials on which no target is presented to provide a direct measure of biases in guessing identity.

Experiment 1

The goal of Experiment 1 was to measure the degree of binding of feature-identity perception and location perception with limited viewing time. We used a variant of the disjunctive feature-search paradigm of Treisman and Gelade (1980). Subjects searched brief, masked, eight-character displays looking for either the letter H (in pink or blue) or an orange character (X or O) in a background of pink and blue Xs and Os. Thus, the different properties problem was dealt with by using a disjunctive search for either of two single-feature targets (H or orange) with values on separable dimensions (color and form).

To ameliorate the location-reporting problem—the possibility of error in translating perceived location in subjective coordinates to experimenter-defined coordinates—we took the eight locations that Treisman and Gelade displayed in two rows and rearranged them into a hollow square with three locations per side. In this configuration every possible target location has either a unique corner code (e.g., upper left) or a unique side code (e.g., left side). As a further aid to location reporting, a prestimulus and poststimulus masking field was used that also consisted of eight characters arranged in the same hollow square configuration as the stimulus array and in the same position. Thus, if subjects found themselves after a trial knowing that they had seen a target in a certain location where a mask character now appeared, it would be clear which location to report.

Two different measures were taken to ameliorate the negative-information problem. Because negative-information guessing is more advantageous if the two features differ sub-

stantially in visibility, considerable efforts were made to equalize them. By taking advantage of the fact that perceptibility of the color target is differentially impaired by adding more of its hue to the mask, we were able, using pilot studies, to design a mask producing roughly equal overall difficulty. We also adaptively varied relative difficulty during sessions by varying the size of the form target (H). The second measure taken to discourage negative-information guessing was including trials without either of the target features and allowing subjects to make a "no-feature" response.

Method

Subjects. Fifteen people with normal or corrected-to-normal vision recruited in the Mountain View, California, area were paid for their participation in the experiment.

Apparatus and stimuli. The experiment was run on IBM PC microcomputers, equipped with Princeton Graphics SR-12 monitors, and Sigma Design Color-400 graphics cards, providing 640 x 400 pixel display resolution at a full 60-Hz refresh rate. The four bit-planes of picture memory were configured so that two predrawn screens (each with three colors plus black) could be alternated between successive frames (with the changeover accomplished during vertical screen retrace). Background characters consisted of X and O in either pink or blue color. The characters were drawn on a relatively large grid, so that X was 19 pixels across by 18 pixels high (approximately 7 x 8 mm, 0.9 x 1.0 degrees of visual angle at a viewing distance of 45 cm) and the O was 17 x 18 pixels. Use of this large grid allowed the H target character (pink or blue) to be drawn in 15 different sizes ranging in small steps all the way from 4 x 5 pixels (2 x 2 mm) to 21 x 24 pixels (9 x 11 mm). Color targets were the X and O characters in the color orange. Mask characters consisted of segments of blue, pink, and orange within an area slightly larger than a stimulus character (25 x 30 pixels, 9 x 1.2 mm). The entire masking field consisted of a matrix of eight mask characters arranged in three rows and three columns, with the ninth center element a small multicolored fixation element. The entire matrix was approximately 5.2 x 6.3 cm (6.6 x 7.9 degrees of visual angle), surrounded by a blue box about 6.8 x 6.3 cm (using a blue box around the mask helped to offset a slight advantage for pink targets over blue targets). As a slight further contribution to equalizing performance, pink Hs were always presented one size smaller than blue Hs.

Design. The experimental data were collected in 8 blocks of 48 trials each. Each block contained 16 trials with no target, 16 with a form target (one blue H or one pink H in each of the 8 possible target positions), and 16 with a color target (one orange X or one orange O in each of the 8 possible target positions). Individual displays were generated by first permuting arrays with 2 each of the 4 types of background items (blue and pink Xs and Os), and then substituting the proper target (if any) in the target position. The order of trial presentation was randomly permuted.¹

Procedure. The subjects were given both written and oral instructions explaining what types of trials could occur and how to classify both the identity and location of targets. Accuracy rather than speed of responding was emphasized.

Each trial began with display of the masking field: subjects were instructed to use the fixation mark to position their eyes. After 1 s, the stimulus field was displayed for a (variable) number of 16.67-ms frames followed immediately by the masking field. Subjects used the number pad keyboard to make their responses. Target identity was selected by pressing one of the three keys on the bottom row: the "O" key (form-feature target), the "." key (color-feature target), or the "Enter" key (no-feature target). Subjects were encouraged to make a

positive feature response if they had a hunch about which one had been present. After making an identity response, subjects used the digit keypad keys (arranged in a square, excluding the middle "5" key) to indicate in which of the eight matrix locations the target had occurred. A location response was required even if the subject had made the no-feature identity response. Immediately after the location response, the masking field was turned off and remained off for an intertrial interval of 1.2 s before the next trial began.

Before collection of the experimental data, two practice blocks were run to familiarize subjects with the procedure and select appropriate exposure duration settings. In the initial block of 62 trials, the first 12 trials were run with gradually descending exposure durations. In the remainder of the block staircase procedures were used with two algorithms: one to find the exposure duration producing 50% average performance, the other to find the H size producing equal form and color performance. The second practice block of 50 trials was run with a constant exposure duration and H size. After the second block, exposure duration was adjusted further, and H size was adjusted if errors on the two types of targets were substantially unequal. Each of the 8 experimental blocks consisted of 48 trials, preceded by 2 throwaway transition trials. Exposure duration and H size were further adjusted between experimental blocks.

Results

The mean exposure duration across the 15 subjects in Experiment 1 was 71.2 ms (range: 53.4-114 ms). Of the total of 5,760 experimental trials, 76 (.013) were excluded from analysis because of missing location or identity responses. On control trials where no target was present, subjects correctly made the no-feature response on .649 of trials. The remaining responses on these control trials were divided unevenly between the two positive responses: .128 for color false alarms and .228 for form false alarms. Thus, there was a substantial bias toward the form response.

Before examining the data from target-present trials, we need to discuss how to score location judgments. Treisman and Gelade (1980) scored a reported location as approximately correct if it were adjacent to the correct location, either horizontally or vertically, but as entirely wrong if it was adjacent diagonally. In pilot experiments, we found that identity responses were more often correct when the reported location was adjacent to the correct location, not only vertically or horizontally, but also diagonally, so we decided to count them all as approximately correct (the chance level of correct location guesses is then exactly .5, an average of .375 for the four corner targets and .625 for the four side targets).² This criterion was fixed before Experiment 1 to avoid post

¹Because blocks were run at different exposure durations, we felt it desirable to equalize target types and positions within a block. This raises the slight possibility that subjects could do better than chance guessing toward the end of blocks without any stimulus information. As we see, the pattern of data we obtained does not leave room for any nonnegligible effects of this kind.

²In principle, it would be possible for subjects to achieve higher than .5 correct on location guesses (using approximate location scoring) by systematically choosing side rather than corner locations. There are no signs that the data are distorted by this strategy, because location performance above .5 with identity incorrect is due entirely to an excess in exactly correct location reports.

hoc inference problems. Insofar as there is any doubt about the correct criterion, ours is appropriately conservative. That is, it is less likely to treat a trial with a small location error as if it embodied no information about location when it did. We wish to find out whether identity information can be obtained without location information, so we ought to base our tests on those trials for which it is least likely that location information might actually have been present.

The bottom matrix of Table 1 shows the overall data from the target-present trials of Experiment 1. It is clear, first of all, that our adjustment procedure produced average performance that was nearly perfectly centered between ceiling and floor levels (1.0 and .5) on both types of responses. Location responses were correct on .765 of trials. The identity results (assuming "no-feature" reports should be divided equally between identity correct and incorrect) are equivalent to two-alternative forced-choice performance of .763. Thus, overall accuracy was virtually identical for location and identity. The proportion of "no-feature" responses on target-present trials was a substantial .263. Subjects were virtually at chance on these trials in reporting location (.135 location right, .128 location wrong).

What is critical for assessing identity-location binding is the degree of contingency in the accuracy of the two responses. The two left data columns of Table 1 cover performance when the subject made a positive feature-identity response. It is helpful to first consider how responses would be distributed among these four cells if we assume perfect binding of identity and location perception and no losses of information in later stages. Only two kinds of trials should occur. When a target identity is not perceived, subjects' responses must represent random guesses on both identity and location, producing equal proportions in all four cells: identity right-location right (Iright-Lright), identity right-location wrong (Iright-Lwrong), identity wrong-location right (Iwrong-Lright), and identity wrong-location wrong (Iwrong-Lwrong). When a target identity is perceived, location should also be perceived, so such trials should end up in the upper left cell, Iright-Lright. Thus, the perfect binding model predicts equal response proportions in the three other cells with errors: Iright-Lwrong, Iwrong-Lright, and Iwrong-Lwrong.

The obtained data show an excess of .023 for both of the cells with one type of error only (Iright-Lwrong and Iwrong-Lright) compared with the cell with both types of errors

(Iwrong-Lwrong). The excess trials are a small proportion of the total, but, looked at differently, are a substantial proportional increment over the .042 Iwrong-Lwrong responses. The optimal statistical model for testing the significance of the excess responses is unclear it is certainly fair, however, to use a sign test across subjects for whether individuals do better than chance on location responses when identity responses are wrong and vice versa. Because it is reasonable for true performance to be above chance but not to be below it, the use of one-tailed rather than two-tailed tests would appear to be appropriate: significance levels for both will be given, however. When location is reported wrongly, the .023 difference between the Iright-Lwrong cell and the Iwrong-Lwrong cell is significant (two-tailed, $p < .05$; one-tailed, $p < .025$). When identity is reported wrongly, the .023 difference between the Iwrong-Lright cell and the Iwrong-Lwrong cell is also significant (two-tailed, $p < .02$; one-tailed, $p < .01$). It is also useful to consider, for trials on which identity was reported wrongly, the breakdown of location reports into the following categories: entirely wrong, .393 (chance, .5); approximately right (off one location vertically, horizontally, or diagonally), .397 (chance, .375); and exactly right, .210 (chance, .125). Note that there is a large excess in trials with location exactly right, but only a small excess with location approximately right.

The top two panels of Table 1 show the data broken down by whether the target presented was the form target (H) or the color target (orange). For the Iwrong-Lright cell, the excess in proportion of responses relative to the Iwrong-Lwrong cell was consistent for both targets. For the Iright-Lwrong cell, however, there was clearly a marked asymmetry. Form-target trials showed a striking excess of above-chance identity responding when location was wrong: .089 for the Iright-Lwrong cell versus only .033 for the Iwrong-Lwrong cell ($p < .002$ by two-tailed sign test). Earlier in this article, we raised the possibility that such an excess might occur because of guessing biases, and control target-absent trials did in fact show a pronounced bias toward the form response. Consistent with this point, color-target trials actually showed slightly worse than chance performance on identity responding when location was wrong: .041 for the Iright-Lwrong cell versus .051 for the Iwrong-Lwrong cell ($p > .1$ by sign test).

Discussion

The results of the first experiment, using a paradigm that minimizes the problems plaguing previous studies, indicate that feature perception and location perception are rather tightly bound. Performance above the chance level on identity responses came almost entirely from trials with correct location responses (including approximately correct) and vice versa.

Although performance on feature identity given an incorrect location response was significantly above chance, the excess was quite small. Treisman and Gelade (1980) found in two experiments .678 and .748 correct feature performance when location was reported wrongly. After correcting for guessing, they argued that "the identity of the target was correctly perceived on perhaps 40% of trials on which the

Table 1
Data by Target Type for Target-Present Trials in Experiment 1

	Identity judgment		"No Feature"
	Right	Wrong	
Form-feature (H) trials			
Location right	.603	.052	.113
Location wrong	.089	.033	.110
Color-feature (orange) trials			
Location right	.530	.077	.157
Location wrong	.041	.051	.145
Total target-present trials			
Location right	.566	.065	.135
Location wrong	.065	.042	.128

location was completely misjudged." To generate a comparable estimate from our data, we note that the excess trials in the Iright-Lwrong cell (.023) are about 10% of the total trials with the location report entirely wrong (.235). Assuming no location information on these trials and identity information only on the excess trials, one can estimate that feature identity was known on about 10% of trials when location was unknown. The estimated proportion of all trials on which identity but not location was perceived (.023) can be compared with an estimate of .524 for the proportion of all trials where identity and location were both perceived (observed proportion of Iright-Lright equals .566 minus the estimate of .042 guesses per cell). This comparison indicates that subjects knew target identity but not location less than one twentieth as often as they knew both location and identity. Even these estimates of perceiving identity without location may be inflated. Excess performance in the Iright-Lwrong cell occurred only with form-target trials on which a correct response could be obtained by the observed bias to guess the form target (see the model presented for Experiment 2).

The other possible direction of perceptual decoupling—reporting location better than chance when feature identity was incorrect—also occurred on a small but statistically significant proportion of trials. We noted earlier that in principle, location performance above chance can be produced by a negative-information guessing strategy analogous to that for features. Although this is plausible a priori, the data show that virtually the entire excess of performance above chance is due to trials where the reported location was exactly correct. Because there is always only one exactly correct location and several approximately correct locations, it is hard to conceive of any guessing strategy that would not lead to most of the excess above chance occurring with only the approximate location correct. After attempting a further replication of this finding in the next experiment, we consider its importance in the General Discussion section.

Although the overall results show strong binding of location and identity perception, a small residual decoupling is present, and we should consider whether it might be the tip of an iceberg. Perhaps perceptual decoupling occurred frequently, but the accompanying subjective experience was tentative and untrustworthy, leading subjects to make a "no-feature" response rather than select a feature. The force of this argument is somewhat lessened by our use of instructions encouraging subjects to report tentative perceptions and by the high false-alarm rates observed on target-absent trials. Nevertheless, this argument deserves to be tested. This can be done by simply forcing a feature choice. If identity-location decoupling occurred in the previous experiment, but was hidden in "feature-absent" responses, it will now be forced to appear in positive feature responses: the decoupling should then be apparent in the contingency analysis. Because no-target responses on target-present trials were numerous in Experiment 1, the potential for a large change is present.

Experiment 2

To observe whether perceptual decoupling of identity and location might occur but with low subjective confidence,

Experiment 2 modified the procedure in Experiment 1 by requiring subjects to choose a feature identity on every trial (while continuing to include control trials on which no target actually appeared). Forcing an identity choice might have the unwanted side effect of increasing use of the negative-information guessing strategy. Fortunately, increasing the number of trials with positive identity responses in Experiment 2 should also provide a much larger data sample to which a mathematical model of guessing behavior can be fit. It is advantageous to be able to estimate guessing biases independently of the sample of trials to be fit, which can be accomplished using the control target-absent trials. To maintain motivation for subjects to make the best choice possible, subjects were not told about target-absent trials. In a tachistoscopic masking paradigm with adaptive control of exposure duration, subjects frequently miss actual targets anyway, so there is no easy way for them to know the difference.

Method

The apparatus, stimuli, design, and procedure were the same as in Experiment 1, except as noted here.

Subjects. A new group of 15 individuals with normal or corrected-to-normal vision from the Mountain View, California, area were recruited as paid subjects.

Procedure. The procedure was the same as that used in Experiment 1, except that subjects were no longer allowed to make a "no-target" response. Subjects were instructed to make the best choice they could between the color targets and form targets on every trial. The instructions described both color targets and form targets without mentioning the possibility of trials with no target. To disguise the presence of such trials, they were excluded from the first two practice blocks (where exposure durations might not have reached threshold levels).

It was considered desirable to hold subjects to about the same average level of stimulus detectability as in Experiment 1. A small change was made in the algorithm for adjusting exposure duration between blocks to take into account the redistribution of some "no-target" responses in the first experiment into correct guesses of target identity.

Results

The mean exposure duration across the 15 subjects in Experiment 2 was 68.6 ms (range: 41.1-155 ms). Of the total of 5,760 experimental trials, 102 (.018) were excluded from analysis because of missing responses for either location or identity. On the control target-absent trials, responses divided .561 for form and .439 for color, so there was again an overall bias favoring the form response.

The bottom matrix of Table 2 shows the overall data from the target-present trials of Experiment 2. Our revised procedure for controlling exposure duration worked well, producing an average of .755 correct on identity judgments and .748 correct on location judgments. Not only are these numbers almost perfectly centered between chance guessing (.5 for each judgment) and perfect performance (1.0), but, as in Experiment 1, they are almost exactly equivalent. Thus, there was no overall difference in performance on identity and location information.

Table 2
Data by Target Type for Target-Present Trials in
Experiment 2

Location judgment	Identity judgment	
	Right	Wrong
Form-feature (H) trials		
Location right	.609	.137
Location wrong	.162	.092
Color-feature (orange) trials		
Location right	.619	.131
Location wrong	.119	.131
Total target-present trials		
Location right	.614	.134
Location wrong	.141	.111

The prediction from the hypothesis of perfect coupling—that performance would be equal in the three cells Iright-Lwrong, Iwrong-Lright, and Iwrong-Lwrong—was approximately confirmed, but once again there were minor but statistically significant deviations. There was an excess of .030 in the Iright-Lwrong cell over the Iwrong-Lwrong cell ($p < .005$ by one-tailed sign test). There was an excess of .023 in the Iwrong-Lright cell over the Iwrong-Lwrong cell ($p < .05$ by one-tailed sign test). Although the differences are small, they are very close to those observed in Experiment 1. Pooling the results of both experiments, these two deviations from independence are highly significant even by conservative tests ($p < .002$ by two-tailed sign test for Iright-Lwrong vs. Iwrong-Lwrong, and $p < .001$ by two-tailed sign test for Iwrong-Lright vs. Iwrong-Lwrong). In Experiment 2, when identity was wrong, approximately right location reports did not exceed chance at all, whereas exactly right location reports were well above chance. The observed proportions for location reports on these trials were .188 exactly right (chance, .125), .357 approximately right (chance, .375) and .455 entirely wrong (chance, .5). A similar pattern was found with the smaller sample in Experiment 1. Pooling over both experiments, the observed proportion was .194 exactly correct (chance, .125), .369 approximately correct (chance, .375), and .436 entirely wrong (chance, .5). Thus, both in Experiment 2 and in the overall data, the excess in accuracy above chance came entirely from exactly correct location responses. For the two experiments together, the deviation of these proportions from chance is highly significant, $p < .001$, $\chi^2(2) = 62.3$.

The two top panels of Table 2 show the data broken down by target type. As in Experiment 1, all of the excess in responses in the Iright-Lwrong cell compared with the Iwrong-Lwrong cell occurred on form target trials. On those trials the excess was a substantial .070, whereas on color trials there was actually a .012 difference in the other direction. Combined with the response bias toward the form target on target-absent trials, this again raises the possibility that some subjects were using a negative-information guessing strategy. Surprisingly, the excess trials in the Iwrong-Lright cell also occurred entirely on form trials. Because the corresponding trend in Experiment 1 was in the opposite direction and only a small absolute number of trials are involved, no interpretation of this anomaly is offered.

Discussion

The basic results show clearly that identity-location binding is quite close. The results do also suggest, on the surface at least, that such binding is not perfect. If identity-location binding was so complete that when subjects were wrong about one response they had no information about the other, the three cells with one or more wrong responses would have occurred in equal proportions. In fact, there was a small excess of trials in both the Iright-Lwrong cell and the Iwrong-Lright cell over the Iwrong-Lwrong cell. The excess in these cells suggests, using a simple correction for pure guessing, that the subjects knew one property about 10% of the time when they were wrong on the other. This is very close to the results of Experiment 1, even though the proportion of responses in the relevant cells was much larger due to the absence of "no-feature" response. In fact, the results of Experiment 2 are remarkably close to that obtained by allocating "no-feature" responses in Experiment 1 equally among all four remaining cells, a sensible prediction if subjects had picked up no information. Thus, Experiment 2 confirms the main results of Experiment 1 and reassures us that allowing a "no-feature" response did not mask any sizable number of trials on which identity and location information were decoupled.

Knowledge of identity without location. On the face of it, our results indicate that when location was not even approximately right subjects knew the feature identity just over 10% of the time. This value indicates much less knowledge of identity without location than the 40% estimate from Treisman and Gelade's (1980) results (both estimates rely on a straightforward correction for pure guessing). There are good reasons to question whether even the 10% figure from our data may be too high. The entire excess in feature-identity performance with location incorrect occurred on form-target trials, whereas control target-absent trials show directly a strong bias to guess form rather than color.

Let us consider in more detail whether the data from Experiment 2 might be due to the negative-information strategy of guessing, on trials where no feature is detected, the feature less likely to be detected. This strategy could produce the largest advantage if it flexibly tracked any changes in the relative perceptibility of the features. It turns out, however, that the observed level of performance is nicely fit with just a crude division of the subjects into two groups, one consisting of subjects biased to guess the form feature more frequently and the other consisting of subjects biased to guess the color feature more frequently. Subjects were assigned to the groups depending on whether they made more form responses or more color responses on control target-absent trials. Thus, classification of subjects into bias groups (and quantitative estimation of bias parameters) was accomplished using a different set of data than the data to be fit.

The model applied to the target-present trials is quite simple. On each trial the feature target is either perceived or not; if it is perceived, both feature identity and location are reported correctly. If it is not perceived, the identity response is made according to the guessing response-bias, and the location is reported with chance accuracy of .5. Each group of subjects is characterized by three parameters. There is a single param-

eter, β_F , for the proportion of form guesses when no feature is seen (regardless of which feature was present). Because only two responses are allowed, the proportion of color guesses is just $1.0 - \beta_F$. There is a parameter α_C for the proportion of color features perceived on color-feature trials, and a parameter α_F , for the proportion of form features perceived on form-feature trials.

With this simple structure, the parameters of the model can be estimated without use of free parameters. First, β_F is set to the level measured on control target-absent trials. According to the model, the Iright-Lright cell proportion depends only on β_F and either α_F (for form-feature trials) or α_C (for color-feature trials). We set the model proportion for the Iright-Lright cell equal to the proportion actually obtained for that cell, both for form-feature trials and for color-feature trials. We can then solve for α_F , the only unknown in the equation for Iright-Lright on form-feature trials, and for α_C the only unknown in the equation for Iright-Lright on color-feature trials. The Appendix contains the equations and parameters used for the model fit.

Fitting the model parameters from the obtained data uses only performance on control target-absent trials, and the Iright-Lright performance on target-present trials. Thus, the model can now be used to predict performance for target-present trials on the remaining three cells, where at least one response was in error. The critical prediction is the excess in the cell with identity right but location wrong (Iright-Lwrong) relative to the cell with both wrong. (We argue later that excess responses in the Iwrong-Lright cell have a different cause; the model presented here by its nature predicts equal proportions for that cell and the Iwrong-Lwrong cell.)

Table 3 shows the target-present data broken down by subject group (form bias and color bias) and the model fits for each data set. Notice that for Group 1 (form bias), the Iright-Lwrong cell is elevated for form trials and depressed for color trials, while the reverse is true for Group 2 (color bias). The model shows the same modulation. As expected, the parameter fits indicate that form-bias subjects saw color targets substantially more often ($\alpha_F = .305$, $\alpha_C = .580$), whereas color-bias subjects saw form targets more often ($\alpha_F = .636$, $\alpha_C = .385$). Thus, the model parameters indicate that subjects did indeed tend to guess whatever target they saw less often.³ Thus, when the results for different groups and different targets are pooled together, the model shows an overall elevation of the Iright-Lwrong cell relative to the Iwrong-Lwrong cell of .026, very close to the actual difference of .030. That the fit is this close may be fortuitous, because the model does not track differences in bias for each individual subject or fluctuations in bias within a session. On the other hand, it is unlikely that the model is correct in assuming that all subjects decided to follow the strategy all of the time.

What can we conclude from the performance of the model? Without free parameters, it successfully predicts: (a) the overall degree of excess performance in the Iright-Lwrong cell, (b) its modulation by target type displayed, and (c) its modulation by the direction of target guessing bias for different groups of subjects. If we take the quantitative perfection of the fits seriously, we can conclude that virtually all of the excess in performance in the Iright-Lwrong cell is due to the negative-

Table 3
Model Fits for Experiment 2

Location judgment	Data		Model	
	Identity right	Identity wrong	Identity right	Identity wrong
Form-feature				
Group 1 (<i>n</i> = 9)				
Location right	.552	.143	.552	.101
Location wrong	.215	.092	.247	.101
Group 2 (<i>n</i> = 6)				
Location right	.698	.128	.698	.120
Location wrong	.082	.092	.062	.120
Total				
Location right	.609	.137	.609	.109
Location wrong	.162	.092	.173	.109
Color-feature trials				
Group 1 (<i>n</i> = 9)				
Location right	.641	.134	.641	.149
Location wrong	.088	.139	.061	.149
Group 2 (<i>n</i> = 6)				
Location right	.588	.126	.588	.105
Location wrong	.167	.119	.203	.105
Total				
Location right	.619	.131	.619	.131
Location wrong	.119	.131	.118	.131
All trials				
Total				
Location right	.614	.134	.614	.120
Location wrong	.141	.111	.146	.120

information guessing strategy. If so, then subjects virtually never have trials on which they perceive a feature but not its location. One could, more cautiously, suppose only that some substantial portion of the Iright-Lwrong excess is due to negative-information guessing. However, because the data before adjustment already showed only an excess of about 3% in the overall proportion of Iright-Lwrong trials, the residual cannot be very impressive.

Knowledge of location without identity. As in Experiment 1, the present experiment found a small excess of trials in the Iwrong-Lright cell relative to the Iwrong-Lwrong cell. Pooling both sets of results together, the combined statistical test shows that the excess is very highly significant. In absolute magnitude, the present result and that from Experiment 1 indicate that on about 10% of trials when the feature target was wrongly identified, it was located correctly. It is in principle possible for this trend to occur because of negative-information guessing strategies (e.g., on trials where no feature is seen, guessing the locations seen less well), but there is no evidence for this possibility. We examined the fine grain of the data broken down by particular positions for such trends.

³Note that if bias is higher toward the identity target seen less often, equal identity performance on the two target types does not mean equal perceptibility. Table 2 shows that the total identity correct performance was very close to .75 for both groups of subjects on both types of targets, whereas the model fits show substantial differences between α_F (probability of seeing the form feature) and α_C (probability of seeing the color feature).

but found none. Furthermore, the excess occurred entirely on trials where the location was reported exactly correct, and no guessing strategy we have considered has such precision.

General Discussion

The experiments reported here, using a modified version of Treisman and Gelade's (1980) disjunctive feature-search task, indicate that perception of feature location and feature identity are closely, but not quite perfectly, bound together. Our discussion deals first with the accuracy of identity judgments when location reports are wrong, and then with the accuracy of location judgments when identity reports are wrong. We then outline an explanation for our findings and other related findings.

Feature Identity When Location Is Unknown

The overall data, adjusted for simple guessing, indicate that when location is unknown feature identity is known only about 10% of the time. The 10% estimate is far less than Treisman and Gelade's estimate of 40%, and even so is almost certainly inflated by the negative-information guessing strategy (guessing whichever feature was less likely to have been seen). Strong evidence for such a strategy is provided by the combination of an asymmetry in overall response bias (most subjects guessed form more often than color on feature-absent trials) and a corresponding asymmetry in excess correct identity responses when location was wrong. Furthermore, when subjects are grouped by the direction of their individual response bias (to report form or color), each group shows excess identity performance only for the half of the data favored by their bias.

We showed above that a simple mathematical model based on the negative-information guessing strategy can, without free parameters (estimating response biases from control feature-absent trials), provide a good fit to both the overall level of excess identity performance when location is wrong and the distribution of the excess across different subject groups and target types. Although the model can account for essentially all of the observed excess, it is not critical that the residual be absolutely zero; a small number of decoupling trials might result from keypressing errors or forgetting location while reporting identity. Clearly, however, our data do not leave room for much perception of identity without location.

Relation to Treisman and Gelade's Results

Why should the accuracy of identity judgments when location responses were incorrect be so much lower in our experiments than in those of Treisman and Gelade (1980)? There are two important possibilities. First, some of the dissociation in Treisman and Gelade's data may have been due to the location-reporting problem. In our paradigm, all locations can be easily coded as unique corner elements or side elements. Furthermore, our masking field, unlike Treisman and Gelade's, repeats the geometric array structure of

the stimulus field in the same location; this may be critical for accurately translating internal subjective-location coordinates to experimenter-defined coordinates. Second, in Treisman and Gelade's study, the negative-information problem was probably worse. We performed extensive pilot studies to equalize feature difficulty. We also used an active staircase procedure during each session to adjust the relative difficulty of the two target features. Even with these efforts, some subjects were poorly balanced in the proportion of errors on the two types of targets. Treisman and Gelade did not report data broken down by target type, so it is impossible to tell how well the difficulty of their targets was matched. However, because they did not adaptively adjust the relative perceptibility of targets, it seems likely that their matching was less effective than ours.

Implications for Treisman's Theory of Attention

Let us consider the implications of our findings for Treisman's feature integration theory of attention (Treisman & Gelade, 1980; Treisman, Sykes, & Gelade, 1977). This theory asserts that, outside of attention, perceptual processing is restricted to the extraction of component features. Focal attention is the "glue" that permits individuals to form coherent objects from component features. Focal attention is directed to a region on a "master location map"; activated features within the attended region are assembled into an object. Outside of focal attention, objects with a conjunction of properties are either not perceived at all or perceived with haphazard spatial match-up of features.

Feature integration theory claims that there are qualitative differences between the processing of conjunction targets and single-feature targets. One important difference is that location and identity are alleged to be tightly bound for conjunction search but not for single-feature search. Our results show that under properly controlled conditions tight identity-location binding does after all occur for single-feature search. Although our experiments did not investigate conjunction search, any difference from feature search would have to be in the direction of weaker identity-location binding, which is unlikely, and would not be in the right direction to support Treisman's theory anyway. Thus, our results make it very unlikely that single-feature search and conjunction search differ in identity-location binding in the way predicted by feature integration theory.

Several other empirical claims supporting Treisman's theory have been seriously challenged by recent investigations. Treisman et al. (1977) and Treisman and Gelade (1980) argued that differences in stimulus set-size effects and positive/negative slope ratios supported the conclusion that single-feature search can be accomplished in parallel, whereas conjunction search requires the serial self-terminating application of focal attention. Pashler (1987) showed that the visual set-size effect for conjunction search, examined on a fine scale, is not linear. For set sizes up to about eight, the effect of additional stimulus elements was small and the negative-to-positive slope ratio was closer to 1:1 than 2:1. Pashler argued that search is serial only across groups of stimuli, and that parallel search for conjunction targets is possible within

groups of about eight elements. If it is true that multiple conjunctions can be processed in parallel, object integration cannot be accomplished by conjoining all activated features inside focal attention into one object. A reasonable conjecture is that the separate properties of objects are indeed represented on maps in their correct locations (now conceded by Treisman & Gormican, 1988), but that the properties of objects are conjoined together on the basis of common location *per se*, provided that the objects are attended.

Other evidence for feature integration theory comes from Treisman and Schmidt's (1982) finding that when multidimensional objects are perceived outside focal attention, subjects mistakenly perceive objects combining features from different objects (illusory conjunctions). Virzi and Egeth (1984) made a strong case that Treisman and Schmidt's data can be attributed to postperceptual reporting difficulties, without postulating illusory perceptual combinations. There is also direct evidence that illusory conjunctions are not common in the search for conjunction targets: Errors are overwhelmingly "misses" rather than the "false alarms" that would be expected if illusory conjunctions were common (cf. Treisman & Gelade, 1980, p. 117; also van Santen & Johnston, 1982). Actually, aside from the disputed evidence for illusory conjunctions, it has never been clear why Treisman's overall theory should require features to conjoin falsely outside of attention rather than simply remain unconjoined.

Implications for the Pooled-Response Model

Although the main goal of Treisman's feature integration theory is to explain how attention permits perception of multidimensional objects, the goal of the pooled-response model of Treisman and Gormican (1988) is to explain search for individual features. This model postulates that the overall activity level in a feature map can be determined without accessing the location of the activity. According to the model, this is accomplished by reading a signal that is pooled (e.g., summed or averaged) across the attended region of the map. Because this pooled signal has a single dimension of strength, it lacks any location information.

The pooled-response model is supported by two lines of empirical evidence. The most direct support comes from Treisman and Gelade's (1980) evidence for decoupling of identity perception and location perception in search for individual features. Our data, collected under conditions designed to ameliorate the location-reporting problem and negative-information problem, indicate that people virtually never perceived the identity of a feature without perceiving its location. Thus, our data show no signs that feature-identification judgments can be based on pooled responses from feature maps.

Treisman and Gormican (1988) buttressed the pooled-response model by using it to explain why parallel detection ("popout") occurs in detecting feature presence but not feature absence (Treisman & Gormican, 1988; Treisman & Souther, 1985). Treisman and Gormican (1988) pointed out that the pooled-response model, combined with Weber's law, can account for this result: One feature in an "empty" feature map (1 vs. 0 signals) produces a much higher signal-to-noise

ratio than a "hole" in a "full" feature map ($N - 1$ vs. N signals). The elegance of this account does not, of course, prove that it is correct. If, as our data indicate, the pooled-response model is wrong, then search asymmetry will require a different explanation. If other plausible alternatives were available, our confidence in abandoning the pooled-response model would be strengthened.

We believe that search asymmetry is best explained not as a byproduct of some other dynamics (e.g., Weber's law), but as a direct consequence of how spatial attention is controlled. Our proposal is simply that positive activation on a feature map can generate an attention-calling signal, but absence of activation cannot. Such a system is clearly feasible, and there are several reasons why it might be desirable. First, because the time for features to register on maps varies, calling attention to missing features entails either waiting until the slowest features were registered or tolerating numerous attentional false alarms; the positive activation proposal avoids the problem because valid attentional signals can be generated as soon as the target element itself is processed. Second, many studies have found that feature errors tend to be misses much more than false alarms (cf. Garner & Haun, 1978; Townsend & Ashby, 1982), so attentional signals for feature absence would frequently be erroneous. Third, in the three-dimensional world, elements in a texture may be missing because of occlusion, without any break in the distal world worthy of attention. Whether or not these functional considerations are decisive, abandoning the pooled-response model clearly leaves open many promising explanations for search asymmetry.

Location Perception When Identity Is Unknown

The data from each of our experiments, adjusted for simple guessing, suggest that location is known on about 10% of the trials when feature identity is not known. This performance cannot be explained by any location-guessing strategy considered, because essentially the entire 10% consists of trials where location was reported exactly correct. This aspect of our data also differentiates our results from those of Atkinson and Braddick (1989). They reported that targets in homogeneous arrays can be located more accurately than they can be identified only when the location judgment required is very coarse (substantially less precise than even our approximately correct standard).

Sagi and Julesz (1985) proposed a model in which parallel preattentive processes compute only differences between elements and their neighbors; the identity of elements, or even the nature of the differences, is not computed. Preattentive processes make available to central processes only a map of the location of "odd" elements differing strongly from the background. If a task can be performed with this information, parallel processing is possible. If a task requires knowledge of feature identity, elements must be searched serially using the "odd-element" map to guide the search.

The evidence in this article that subjects do sometimes know the exact location of a feature target, but not its identity, might appear to support Sagi and Julesz's (1985) proposal. This apparent confirmation would be especially useful because, as noted earlier, our paradigm is less likely to founder

on the different properties problem. Explaining our results with the Sagi and Julesz model faces serious obstacles, however. Their model is well suited to displays with homogeneous backgrounds, where elements not near a target would produce very low difference scores (zero, except for processing noise). With our heterogeneous backgrounds—a mixture of pink and blue Xs and Os—elements not near a target would frequently produce large deviation scores (adjacent elements frequently differ in both color and form), whereas targets themselves, which share a feature with many background elements, would sometimes produce only moderately high difference scores. If one also considers the inevitable presence of some processing noise, it becomes difficult to see how overall difference scores would permit targets to be discriminated reliably from the most deviant background items. Pashler (1988) recently concluded from related considerations that "popout" detection of discrepant items in multidimensional arrays is governed by computations on individual maps rather than a composite difference computation.

Another reason why Sagi and Julesz's model is difficult to reconcile with our data is that it predicts a complete dissociation of location and identity (complete enough to support the argument that identity is not processed preattentively at all). Although Sagi and Julesz's own data are consistent with that prediction, our data show a much less dramatic dissociation (targets that cannot be identified can only rarely be located correctly). We argue shortly that our data support a model in which preattentive processes do more than Sagi and Julesz's proposal; actually identifying visual features rather than only computing local differences.

It is possible, of course, that our paradigm and Sagi and Julesz's paradigm tap different processes, and it might therefore appear that our conclusions and theirs could both be right. Certainly there are major differences in the paradigms. The background texture in their displays was homogeneous, whereas ours was heterogeneous, and their tasks required subjects to deal with multiple targets on a single trial, whereas ours did not. In spite of the difference in paradigms, though, both theoretical conclusions cannot be right. Sagi and Julesz (1985) made general claims that preattentive processes *cannot* determine feature identity. If we are right that in the paradigm we studied subjects make use of feature identities computed by preattentive processes, then Sagi and Julesz erred in underestimating the overall capability of preattentive processes.

Toward a Model of Ideality and Location Binding

We now consider how to explain feature search in large arrays consistent with our new data and other findings. Our proposal borrows certain ingredients from the theories of both Treisman and Julesz, but in total functioning differs from both. Our theory is motivated both by the direction of the asymmetric decoupling we found—identity is only perceived if location is perceived, but location can sometimes be perceived without identity—and by the very small absolute size of the decoupling.

Our proposal begins with the standard assumption that preattentive processes produce a number of separate maps of

feature activation. Positive activation on feature maps is assumed to have the potential to generate attention-calling signals analogous to computer interrupt signals: such signals have a one-dimensional strength parameter but do not specify feature identity. The strength of attention-calling signals can be modulated by priming and by the strategic concerns of the subject. Subjects can exert control over what features should produce strong interrupt signals (e.g., red items) and what size regions (e.g., isolated feature singletons: cf. Pashler, 1988). The location specified by an interrupt signal (in subjective coordinates) is used to direct spatial attention.⁴ When a location is attended, the features at that location are perceived centrally, and can be used for overt reports, memory storage, and so on.

Our theory, unlike Treisman and Gormican's (1988), does not allow feature maps to be accessed through pooled signals. Consequently, our theory does not permit identity perception without location perception; any observed decoupling in that direction must be due to guessing artifacts or reporting artifacts. Our theory does permit location perception without identity perception, if central processes receive the interrupt signal—which contains a specification of location—but are unable to fetch the feature from that location before it is disrupted by masking. Our theory differs from Sagi and Julesz's (1985) theory in that calls to attention are generated from preattentively computed feature identities rather than only local difference signals. What is deferred until servicing of the interrupt signal is only the transfer of the feature identity to central processes, not its computation. In addition to fitting better with physiological evidence for massive parallelism in feature detection, our theory explains why the identity-location decoupling in our experiments was so small. If the feature-identity computation were itself deferred, severely limiting viewing time should have produced a much more severe dissociation.

A similar objection could be made to our theory. If the location of a feature becomes available centrally before fetching its identity, why does the subject not get caught more frequently in "feature-fetchus interruptus," with feature location but not identity known centrally? This objection is especially relevant because our paradigm uses poststimulus masking. Our rejoinder is that the fetching process must be quite rapid (perhaps several tens of milliseconds), and that a feature representation somewhat reduced in quality will still be sufficient to read a queried feature. To avoid false alarms in calling attention, interrupt signals should be generated only by very high-quality feature representations. A lower quality representation should be sufficient to retrieve the feature; in our experiments success requires only determining which of two feature maps has more activity in one location.

We conclude by discussing how our findings relate to three other lines of work. First, we need to reconcile our results

⁴For present purposes, it is not critical to specify the extent to which attention-calling interrupt signals either compel handling by central processes (stimulus-driven mode) or await a strategic decision that they should be processed (cognitively driven mode).

with the evidence from Mewhort, Campbell, Marchetti, and Campbell (1981) that location and identity information are severely dissociated when masked strings of letters are probed with a spatial bar marker; delaying the marker disrupts location information more than identity information. Mewhort et al.'s own discussion (p. 62) suggests two different possibilities. It may be that although our search task taps processing at the feature level, the bar-probe task taps processing at the level of identified letters. The other possibility is that delaying the probe makes it difficult to judge the relative alignment of the probe and the stimulus array, resulting in late-stage location-reporting errors. If either of these accounts is correct, then Mewhort et al.'s results do not cast doubt on our conclusion that identity and location are tightly bound in accessing feature maps.

Second, our conclusions fit comfortably with Nissen's (1985) data on the relation between form, color, and location perception. She presented subjects with brief displays of four colored shapes, followed by a poststimulus masking field with a reporting cue.⁵ In Experiment 1 she found equal accuracy reporting shape from a location cue and location from a shape cue, consistent with our finding that identity and location reports are equally accurate and tightly bound. In Experiment 2, subjects reported color and shape given a location cue and reported location and shape given a color cue. The location-cue data support the conclusion that color and form were retrieved independently from separate maps. The color-cue data fit a double-retrieval model in which color is used to retrieve location, and then location is used to retrieve shape. Our model is consistent with these conclusions. A color feature and a form feature are represented on two different maps; although one act of focal attention should be sufficient to retrieve all properties of an object (cf. Kahneman & Henik, 1981), under data-limited conditions (Norman & Bobrow, 1975) what is critical is how many representations need to be intact. Finding the location of the cued color requires only the use of a color-feature map, while finding its shape requires that both a color map (to establish the location) and a form map be intact.

A final matter is the relation of our findings to the neuro-psychological and neurophysiological evidence for two major pathways in the processing of visual stimuli: a ventral (temporal lobe) path dealing primarily with object recognition and a dorsal (parietal lobe) path dealing primarily with object location (cf. Desimone, Schein, Moran, & Ungerleider, 1985; Ungerleider & Mishkin, 1982). Given this evidence, it may appear to be paradoxical that identity and location perception should be as tightly bound as we have found. One resolution of the paradox, of course, is to argue that anatomical separation and functional dissociation by lesions do not necessarily imply that two processes will be dissociated by limiting processing time. We prefer, however, a resolution along different lines, either (a) that the feature-search task can be accomplished at an early level before the split in pathways, or (b) that our task only requires a form of location coding that can be extracted from within the dorsal (temporal lobe) path, in which many features remain coded on topological maps. If

either of these two answers is correct, dorsal lesion patients might be able to perform our task reliably.

⁵Because Nissen (1985) used displays of only four items and the retrieval cue was not shown until the stimuli disappeared (and further time must elapse while the cue is processed), it is possible that the main constraint on performance was how well subjects could encode the objects in a postattentive "whole-report buffer." It would be valuable to repeat her experiments with precues (presented just before the display, or if substantially earlier, with monitoring of eye position).

References

- Atkinson, J., & Braddick, O. J. (1989). 'Where' and 'what' in visual search. *Perception*, 18, 181-189.
- Baron, J. (1973). Perceptual dependence: Evidence for an internal threshold. *Perception & Psychophysics*, 13, 527-533.
- Cowey, A. (1985). Aspects of cortical organization related to selective attention and selective impairments of visual perception: A tutorial review. In M. Posner & O. Marin (Eds.), *Attention and performance XI* (pp. 41-62). Hillsdale, NJ: Erlbaum.
- Desimone, R., Schein, S., Moran, J., & Ungerleider, L. G. (1985). Contour, color and shape analysis beyond the striate cortex. *Vision Research*, 25, 441-452.
- Dumais, S. T. (1979). *Perceptual learning in automatic detection: Process and mechanism*. Unpublished doctoral dissertation, Indiana University, Bloomington, IN.
- Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.
- Garner, W. R., & Haun, F. (1978). Letter identification as a function of type of perceptual limitation and type of attribute. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 199-209.
- Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 181-211). Hillsdale, NJ: Erlbaum.
- Krumhansl, C. A., & Thomas, E. A. C. (1976). Extracting identity and location information from briefly presented letter arrays. *Perception & Psychophysics*, 20, 243-258.
- Logan, G. D. (1975a). On the independence of naming and locating masked targets in visual search. *Canadian Journal of Psychology*, 29, 51-58.
- Logan, G. D. (1975b). On the relation between identifying and locating masked targets in visual search. *Quarterly Journal of Experimental Psychology*, 27, 451-457.
- Mewhort, D. J. K., Campbell, A. J., Marchetti, F. M., & Campbell, J. I. D. (1981). Identification, localization, and "iconic memory": An evaluation of the bar-probe task. *Memory and Cognition*, 9, 50-67.
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229, 782-784.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Nissen, M. J. (1985). Accessing features and objects: Is location special? In M. Posner & O. Marin (Eds.), *Attention and performance XI* (pp. 41-62). Hillsdale, NJ: Erlbaum.
- Norman, D. A., & Bobrow, D. (1975). On data limited and resource limited processing. *Cognitive Psychology*, 7, 44-64.
- Pashler, H. (1987). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception & Psychophysics*, 41, 191-201.

- Pashler, H. (1988). Cross-dimensional interaction and texture segregation. *Perception & Psychophysics*, *43*, 307-318.
- Sagi, D., & Julesz, B. (1985). Detection versus discrimination of visual orientation. *Perception*, *14*, 619-628.
- Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, *1*, 54-87.
- Townsend, J. T., & Ashby, F. G. (1982). Experimental test of contemporary mathematical models of visual letter recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 834-864.
- Treisman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing*, *31*, 156-177.
- Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97-136.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*, 15-48.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*, 107-141.
- Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, *114*, 285-310.
- Treisman, A., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. In S. Domic (Ed.), *Attention and performance VI* (pp. 333-361). Hillsdale, NJ: Erlbaum.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behaviour* (pp. 549-586). Cambridge, MA: MIT Press.
- Van Santen, J. P. S., & Johnston, J. C. (1982). [Illusory conjunctions in letter search]. Unpublished raw data.
- Virzi, R. A., & Egeth, H. E. (1984). Is meaning implicated in illusory conjunctions? *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 573-580.

Appendix

The quantitative model fit to the data from Experiment 2 is based on a simple event tree. It is assumed that on each trial the subject has a certain probability of perceiving the target (α_F for form trials, α_C for color trials). If the target is perceived correctly, both the identity and location of the target are reported correctly. If the target is not seen, it is assumed that the subject guesses both the target identity and location randomly. Feature identity is guessed to be "form" with a probability of β_F , and "color" with a probability of $(1 - \beta_F)$. On trials where the target is not seen, the location is guessed randomly (with a probability equal to .5 of being correct, using the scoring system that counted adjacent location reports as correct). The following equations give the model predictions for the four data cells for form-target trials:

$$P(\text{Iright-Lright}) = \alpha_F + (1 - \alpha_F)(\beta_F/2),$$

$$P(\text{Iright-Lwrong}) = (1 - \alpha_F)(\beta_F/2),$$

$$P(\text{Iwrong-Lright}) = (1 - \alpha_F)((1 - \beta_F)/2),$$

and

$$P(\text{Iwrong-Lwrong}) = (1 - \alpha_F)((1 - \beta_F)/2).$$

For color-target trials, the equations for the four data cells are as follows:

$$P(\text{Iright-Lright}) = \alpha_C + (1 - \alpha_C)((1 - \beta_F)/2),$$

$$P(\text{Iright-Lwrong}) = (1 - \alpha_C)((1 - \beta_F)/2),$$

$$P(\text{Iwrong-Lright}) = (1 - \alpha_C)(\beta_F/2),$$

$$P(\text{Iwrong-Lwrong}) = (1 - \alpha_C)(\beta_F/2).$$

The model was fit separately to the data from Experiment 2 for two groups of subjects. Group 1 contained the 9 subjects whose data on target-absent trials showed a response bias toward the form re-

sponse, and Group 2 contained the 6 subjects who were biased to the color response. The purpose of the model was to test whether the negative-information strategy can account for the excess of trials in the Iright-Lwrong cell compared with the two other error cells, Iwrong-Lright, and Iwrong-Lwrong. Because the model was not intended to account for the overall level of correct performance, we ensured a perfect fit of the Iright-Lright cell, and used it for parameter estimation. The mean response-bias values for the two groups on target-absent trials ($\beta_F = .709$ for Group 1; $\beta_F = .339$ for Group 2) were used as fixed parameters for fitting the target-present trials (because the target-present trials and target-absent trials were mixed within blocks, the same guessing biases should have held for both). With β_F fixed in this way, the equations for $P(\text{Iright-Lright})$ each contain only one remaining parameter, the probability of seeing the target (α_F for form trials, and α_C for color trials). With the value of α_F and α_C now set from the equation for $P(\text{Iright-Lright})$, there are no remaining parameters in the model. Thus, the distribution of trials in the three error cells follows from the equations with no need to estimate "free" parameters. The model fits in Table 3 were produced with the following parameter values: for Group 1, $\beta_F = .709$, $\alpha_F = .305$, $\alpha_C = .580$; for Group 2, $\beta_F = .339$, $\alpha_F = .636$, $\alpha_C = .385$.

Note that in keeping with the spirit of the negative-information strategy, the parameters for Group 1 reflect seeing color targets substantially more often than form targets, but guessing form more than half the time when no target was seen. Conversely the parameters for Group 2 reflect seeing form targets substantially more often than color targets, but guessing form much less than half the time when no target was seen.

Received November 4, 1988

Revision received November 20, 1989

Accepted December 11, 1989