Feature and Conjunction Information from Brief Visual Displays

By

Marcia F. Grabowecky

B.A., The University of Calgary, 1981

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(Department of Psychology)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

July 1987

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Psychology

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date August 9, 1987

Abstract

The feature integration theory of object perception (Treisman & Gelade, 1980) suggests that the perception of multidimensional stimuli requires that attention be serially directed to the items in a visual display in order to correctly conjoin features into objects, while the perception of features does not require serial attention. Under conditions in which the serial focusing of attention is disrupted by reducing display duration, available information about conjunctions of two features should not exceed the independent information available about the constituent features. Three experiments using a partial report paradigm employing a location cue were conducted in order to test this prediction. Subjects viewed colored letter displays that varied in cue-display stimulus onset asynchrony. The dependent measure was accuracy of response. Results suggest that a small amount of information from a separate representation of conjunctions of features may be accessible.

Table of Contents

List of Figures

Acknowledgement

I would like to thank Anne Treisman, in whose laboratory and under whose supervision this work was conducted, for her generous support, and to thank Daniel Kahneman for his support and useful comments. I also appreciate the assistance and encouragement of my committee, Lawrence Ward, Raymond Corteen and Janet Werker. I thank my colleagues in the Attention Labs at both the University of British Columbia and at the University of California, Berkeley for their invaluable comments and support. Lastly, I thank my parents, my husband Jim MacDonald, and my son Devin for their patience and encouragement.

The research reported here was supported by a Postgraduate Scholarship from the Natural Sciences and Engineering Research Council of Canada.

List of Tables

Feature and Conjunction Information from Brief Visual Displays

There are at least two distinctly different ways of looking at the perception of objects. If we can rely on our intuitions, perception is holistic. We are aware of complete and richly detailed objects and scenes in the visible environment. Objects have locations in space, they have parts and features and as much detail as we care to see. In order to know this all we need do is open our eyes. Seeing is effortless.

An alternative to this conception is that what we do is the antithesis of this process. Rather than instantly and automatically representing the whole and then narrowing and focusing to extract its parts, we first register the many components of the whole and only later recreate objects and scenes. This view of seeing implies a multitude of local processes rapidly encoding and recoding the visible scene many times over.

Both of these approaches are currently taken by researchers, and both have long traditions in psychology. The first view found support among Gestalt psychologists and now is mainly found among direct perception theorists (after Gibson, 1959). Proponents of direct perception argue that

the very richness of information in the visible environment is enough to account for the ease with which we see, that structure is inherent in the stimulus and all we need do is use it. The other view of perception as synthesis has roots in the associationist tradition. From this perspective the complexity of the information present in a visual scene poses problems with which the perceptual system must successfully deal in order for the organism to survive. Because of these differences in approach the two views have also emphasized different aspects of seeing, the first dwelling on stimulus structure and the second dwelling on structure in processing.

Statement of Program

The research reported in this thesis is composed of several facets. The hypothesis tested originates from Treisman's feature integration theory (Treisman & Gelade, 1980). This theory suggests that object perception is partially dependent upon the focusing of attention for combinations of features to be constructed correctly. The present hypothesis is that no representation of conjunctions of features exists until attention is directed to their location in space. In order to establish the context for the research, I will discuss the problem of feature integration, feature integration theory and some of its

alternatives, and finally, the nature of the paradigm used to test the prediction.

Feature Integration

The assumption that feature integration is an important process is fundamental to many approaches within the information processing paradigm. For example, in Anderson's (1981) information integration approach, information processing consists of a sequence of operations that condense multiple stimuli or dimensions into a single result which then becomes the input to the next level of processing.

The importance of feature integration may be assessed by sampling the literature in both physiology and perception. The literature provides strong (though not conclusive) evidence that the decomposition of input into its component features is the function of some brain mechanisms. Early evidence from Hubel and Wiesel (1959, 1968) indicated that single cells in the striate cortex responded best to particular configurations in stimuli, for example, to lines of a particular orientation. At least in some cells these responses are invariant with changes in other properties of the stimulus, such as its color. More recently the idea of individual cells single-handedly representing features has been replaced with the notion that

groups of cells acting in concert are needed to code different values within the same dimension.

Another possibility is that the visual system decomposes input, but not into such natural real world features as Hubel's and Wiesel's lines and edges. The visual system may instead perform something like a rough Fourier analysis of the scene into its component spatial frequencies (De Valois & De Valois, 1980). In addition to spatial frequency the system needs to determine the phase of stimuli in the field. For simple stimuli, phase corresponds to or encompasses positional information. Information about relative phase, or the relationships of components of patterns to one another, must also be coded. A system which can capture this type of information could then recognize a pattern in spite of changes in its orientation in the visual field. If this type of analysis is the decomposition which takes place at the earliest levels in the visual system, then features which are more closely related to real world features are later constructs in visual processing.

Evidence for separate analysis of features is also available from performance in such perceptual tasks as texture segregation and visual search. Treisman and Gelade (1980) found that texture segregation was easy when two areas differed along a single dimension, but difficult if they differed only in the conjunction of values from more

than one dimension. As well, Treisman (1982) found that an element in a display could be hidden if it was located at the border of two groups each of which shared one of its features. These data suggest that separate values and not their combinations are what is coded at the level of texture segregation.

A second type of evidence comes from asymmetries in search tasks (Treisman, 1985). In some search tasks when the roles of target and distractor are reversed, search times change. A task which is rapid and parallel in one condition becomes serial and slow in the other. This asymmetry can be used to determine which of two potential features the visual system is using in making discriminations. For example, in discriminating circles from arcs, closure or boundedness in the case of the circle could allow it to "pop-out" of a field of arcs. Alternatively, the line ends or terminators of the arcs could pop-out of a field of complete circles. This later condition seems to be the one which holds, at least where the distinction between circle and arc is not grossly apparent.

Some of the strongest evidence for the independent representation of features is the occurrence of "illusory conjunctions" (Treisman & Schmidt, 1982). Illusory conjunctions are miscombinations of features into objects

that did not appear in a display. Treisman and Schmidt showed subjects brief displays of colored letters flanked by black digits. To ensure that attention was loaded and spread across the display, the subject's primary task was to report the identities of the two digits, and only then to report the letters and their colors. Trials were included in the analysis only if the digits were correctly reported. The types of errors made by subjects were analysed to compare conjunction errors to intrusion errors. Conjunction errors were defined as the reporting of features that had appeared in the display, but that did not appear in that combination. Intrusion errors were reports of features that were not present in the display. Conjunction errors occurred significantly more often than did intrusion errors. Since they occurred more often than expected, it was unlikely that all the conjunction errors were simply guesses. The nature of conjunction errors made by subjects is also illuminating. They were as likely to conjoin the color from a small item with the shape of a large one, or to fill an outline object as they were to recombine features from more similar items. It appeared that subjects were detecting the presence of a feature, and then providing enough of it to create the perceived object. The distance separating features that contribute to illusory conjunctions also had no effect on conjunction errors. Subjects were as

likely to pull features from widely separated stimuli to conjoin as they were ones that were closer together.

Other evidence for feature analysis comes from speeded classification tasks. In these tasks subjects are asked to selectively attend to a particular dimension or dimensions in a set of stimuli and quickly sort the set into classes. The usual paradigm has one or more irrelevant dimensions which vary either redundantly or orthogonally with the dimension(s) of interest. If variation on the irrelevant dimension does not facilitate or interfere with sorting or reaction time for the relevant dimension it is assumed to be separable from that dimension. Garner (1974; Garner & Felfoldy, 1970) found some dimensions that fit the definition of separability and some that did not. Color and form were separable even within the same object, while dimensions such as saturation and brightness were not. Dimensions that were not separable were termed "integral" by Garner (who adopted Lockhead's (1966) term). Yet other dimensions appear to combine in ways that are neither separable nor integral. A redundant combination of the horizontal and vertical positions of dots causes facilitation of response to a single dimension (either horizontal or vertical position reported separately), which supports a conclusion of integrality, but when these are varied orthogonally they cause much less interference than

do orthogonal values of brightness and saturation (Garner & Felfoldy, 1970). Dimensions of this sort have been called configural, and the original dichotomy, separable versus integral dimensions, has been replaced by a continuum on which these mark endpoints.

Blob Perception

A theory in the tradition of holistic perception is Lockhead's blob perception. Lockhead (1966, 1972) suggests that a stimulus is perceived holistically in the first stage of perception, and he calls the resulting representation a blob. If the task set for the subject is designed so that a response can be made based upon this information alone then processing stops here. An example of a task Lockhead thinks could be accomplished directly from a blob representation is identifying or positioning an item in a multidimensional feature space. Only if the task requires decomposition of the stimulus into its components for response selection will later serial stages of processing occur. Lockhead proposed his model to deal with data indicating that a saw-tooth pairing of levels of two or more features was easier to discriminate than were linearly arranged pairings of the same features (Lockhead,1966). Parallel models (e.g. Biederman & Checkosky, 1970) assume that the analysis of the features of an item is concurrent and that the first feature

to be identified will provide information for a response to be made if responses can be based upon either feature. Lockhead argued that if these models are correct then linear and saw-tooth pairings of dimensional values should not differ in speed or accuracy in sorting tasks.

Lockhead's blobs are integral stimuli; he does not have parallel processes in his model because his unitary stimuli do not need to be processed for independent decisions since subjects are responding to the stimulus as a whole. He proposes that any kind of filtering task, a task where some component of a composite stimulus is relevant and another is irrelevant, will require more time and effort than an identification task. It takes less processing time to respond to the item as a whole than to respond to a part of it. Presumably, responding to the total stimulus would be more accurate as well.

Lockhead is not clear about which types of stimuli will produce single blobs and which might produce more than one "integral" representation. He suggests, in an apparent concession to feature analysis, (Lockhead, 1972) that clearly separable features may produce more than one blob even from a single stimulus. Lockhead contends that his blobs are different from other theorist's features in that blobs may consist of combinations of several physically specifiable dimensions. This, however, is also true of

features such as color, which vary in hue, saturation, and brightness. What is not obvious is how Lockhead's multiple/unitary blobs better account for object perception than feature integration theories do.

Early and Late Selection

An enduring question in attention research is, where in processing does selective attention intervene? Theorists generally divide into two camps, those who propose that attention is used to select the appropriate response after perceptual processing has specified the stimuli that require decisions, and those who propose that attention is required to select the stimuli that will benefit from perceptual processing. As with most other controversies, this one is rooted in differences in emphasis. Early selection theorists have largely been concerned with perceptual issues, and the fact that potential inputs are so numerous that perceptual overload would certainly occur if some sort of stimulus selection did not occur. Late selection theorists, on the other hand, have been mainly concerned with output, and thus response selection. The major difficulty is insuring that the system eventually chooses only one out of a number of incompatible responses.

The role of a limited capacity mechanism is also a concern and marks another difference in emphasis. Some

models assume a central capacity limit (Kahneman, 1973) and others limits only within relatively independent subsystems (Allport, Antonis, & Reynolds, 1972; Treisman, 1969). This distinction is orthogonal to the early-late selection issue, since proponents of both views can be found in each camp. Of course, the two types of limits are not exclusive. It is probable that both central and local capacity limits exist.

The early-late dichotomy really marks the ends of a continuum. The real question for us is, where in either perceptual or response processing does attention act? Does attention select stimuli for perceptual processing or select only the result of that processing? If attention is applied post-categorically, then a number of operations such as figure-ground segregation, feature integration and stimulus identification must occur preattentively. Post-categorical models (e.g. Coltheart's, (1983) model of iconic memory) imply that attentional manipulations should have no effect on any of these processes.

A series of experiments that directly tested the effect of attention on feature integration suggested that attention improved both the quality of information obtained from displays, and the integration of those features (Prinzmetal, Presti, & Posner, 1986). Spatial attention was manipulated by employing a partially valid spatial cue. This cue correctly indicated the position of stimuli in the display

for 80 percent of the trials, but misdirected attention for the remaining 20 percent. Since subjects were required to determine whether or not a target stimulus was present at the location of the stimulus array, selective attention was also necessary once their spatial location had been determined. Their observation that attention improved the quality of stimulus encoding, as evinced by a small though significant reduction in the number of feature errors on valid cue trials, appears to contradict the claim that features are coded automatically. However, Prinzmetal et al. did find a large effect of cue validity on the number of conjunction errors that subjects made as well.

Pashler (1984), tested a prediction of late selection theory that encoding of familiar stimuli to the level of identification occurs whether or not attention is applied to them and whether or not the subject wishes it (e.g. in Stroop interference). If this process is automatic, then identification of items in a display should be computed in parallel, and a probe or location cue would simply retrieve the identity of the item at the cued location. Pashler manipulated factors such as stimulus discriminability and contrast, which he assumed would affect stimulus encoding but have no effect on the retrieval of already encoded items. He assumed that even with degraded stimuli any encoding process would be complete by 300 ms, since choice

reaction time for stimulus classification may be less than 300 ms. Pashler found that even with 300 ms of previewing time in which to encode stimuli, discriminability and contrast had an effect on reaction time to respond to a cued item in the display. Previewing the entire display did not produce a reaction time advantage over pre-cueing a location in the display. Stimulus quality affected reaction time in both cases. Pashler concluded that evidence supported early pre-categorical selection rather than post-categorical selection. He left open the possibility that some processes, such as feature extraction, could proceed in parallel across the display.

Feature Integration Theory

Feature integration theory (Treisman & Gelade, 1980; Treisman, Sykes & Gelade, 1977) assumes as a starting point that feature integration *is* a problem for the perceptual system. Based partly upon the results described above, feature integration theory assumes that features are coded early, automatically and spatially in parallel across the visual field. Different dimensions are coded in specialized modules with different values on the dimension represented in separate feature maps within a module. The feature maps may or may not be spatially contiguous although highly discriminable values on a dimension would be effectively

separable even if maps are continuous. This allows rapid and parallel detection of distinctive features in a display. These automatically coded differences mediate figure-ground segregation and are thus the first step in object perception, parsing the field into potential objects.

When the detection of specific objects defined by a conjunction of features is necessary, attention is required to ensure the correct combination of features. A map of locations that shows where features are, but not which features are present in those locations, is the medium from that attention selects. When attention is focused on a particular location, links to the modular feature maps retrieve all features which are currently active in that location. These links are the only way in which locations of features in the feature maps may be specified. The spatial extent of attention is variable, and a trade-off between accuracy of location of features and conjunctions and speed of processing is a result. The map of locations may be either earlier or later than the feature maps in processing order (or possibly, concurrent with feature mapping). If it is before or concurrent with feature mapping, then an early representation of conjoined dimensions is implied, and if after, no such representation is necessary. Feature analysis occurring after location mapping is consistent with the data that at least some

neural coding is of combinations of features on two or more dimensions.

It is reasonable that the features conjoined to form object representations are features that encode real world properties and more complex structures than those found in early visual encoding. Treisman and Paterson (1984) obtained results suggesting that for some subjects the emergent feature closure could migrate from one stimulus to another. In displays where subjects searched for the presence of a triangle, more false positive responses occurred when a circle, providing the emergent feature closure, was present (in addition to lines and angles) than when it was not.

The proposed final stage in feature integration is one in which the different properties extracted from the environment are recombined to create objects and scenes. Conscious perception relies on the creation of temporary object representations which contain collected information about the features present in a particular location in space and any information about changes that have occurred to those properties (Kahneman & Treisman, 1984). These representations are object centered, and may contain semantic or lexical information if the object is familiar; however, the absence of such information poses no difficulty for the construction of this type of representation.

The Partial Report Paradigm

The partial report paradigm was developed by Sperling in 1960, prompted by reports from his subjects telling him that they could see more in short duration visual displays than they were capable of reporting. Sperling reasoned that if this were the case, when subjects were cued to report a part of the display their accuracy on this subset could be used to estimate the amount of information they could "see" in the whole display. Further, if some sort of complete, though brief representation existed, by varying the interval between the offset of a display and the onset of a cue to report a part of it the duration of the representation could be measured. Sperling (1960) demonstrated the expected superiority of partial report over total report and also that this superiority disappeared by 300 ms after the offset of the display.

Sperling reasoned from his data that the iconic store contained a holistic and unanalyzed representation of the items in the visual field. He concluded that the icon was a holistic representation since his subjects could report from any portion of the display even with a delayed cue. This suggested that they had available a complete "copy" of the display from which they could read off items if a cue came before the representation decayed. He argued that the icon

was a pre-analytic representation since his subjects were unsuccessful in reporting items from the display based upon categorical information, although they could report based upon distinctions in such features as color and location.

Since Sperling's experiments were conducted, many other researchers have used the partial report paradigm to examine the nature of iconic storage and its relationship to more enduring representations and to decisional and response processes. The clear picture of iconic storage as a complete unanalyzed visual store has been considerably muddied, and there are now well-founded doubts that data referred to under the name "iconic memory" reflect a unitary process.

The major issue in iconic memory research became the nature of the information it contained. Early evidence in the field indicated that it was a pre-categorical store. Sperling found (1960) that subjects could use physical cues to selectively report from displays, but not cues such as letter-digit category membership. Von Wright (1968) found that selection by color, location, brightness, and size was easy, while selection by the orientation of letters was difficult.

Other evidence suggests that some discriminations based upon category membership are possible even when categories are not defined by physical features. Some researchers have

found partial report superiority with letter-digit categories (Merikle, 1980). Variations in methodology make assessing the contradictory findings on this issue difficult.

In an attempt to clarify some of these issues, Coltheart (1980) analysed the iconic memory literature to see whether the concept of an iconic memory as an informational representation distinct from after-images and visible persistence could be supported. He classified the studies he found into those that recorded neurophysiological data, those that relied upon phenomenological reports (called visible persistence), and iconic memory. Only the latter category is primarily concerned with how subjects extract and use information from visual displays.

Coltheart suggested that iconic memory is different from visible persistence or after-images. A partial report advantage occurs in report conditions that do not create after-images, and while visible persistence is inversely related to stimulus duration and luminance, iconic memory is not (but see Eriksen & Schultz, 1978, for a dissenting view). Coltheart concludes that while after images and visible persistence are pre-categorical, iconic memory reflects a post-categorical store.

Haber (1983) published a controversial attack on the idea of the icon in information processing psychology. He

rejected the traditional image of the icon as a static retinotopic store, claiming that such a system would not only be useless, it would be detrimental to vision. An idea of the icon as a static retinal representation is problematic. Successive images should mask one another, and saccades should cause smearing of the image. Since our vision is not occasionally missing chunks, and blurry the rest of the time, Haber is correct in rejecting such a store.

Iconic memory need not be either retinotopic or static. Davidson, Fox and Dick (1973) showed that metacontrast masking acted upon a retinal location but allowed a stimulus presented in the same spatial location but at a different retinal location (by allowing an intervening saccade) to be processed unhindered. This contrasts with an earlier conclusion (Averbach & Coriell, 1961) that static viewing (in the same fixation) of a display and metacontrast mask acts upon retinal location, but in that experiment retinal and spatial locations were confounded.

Icons need not be static either. Treisman, Russell and Green (1975) presented subjects with moving dot stimuli. The subject's task was to report the direction of movement of the dots (clockwise or counter-clockwise). Data showed a partial report superiority in this task, indicating that iconic storage of dynamic stimuli was possible.

A likely source of the confusion over the nature of iconic memory is the metaphor of the icon itself. An icon implies a static, photograph-like image, and this implication was imported into the theory. However, the implications may be rejected without discarding either the research or the concept. I will continue to use "icon" and "iconic" in this paper, while referring to a process that is dynamic and spatiotopic rather than retinotopic.

The Present Investigation

Partial report studies have not directly addressed the question of whether or not iconic representations are unitary. Although partial report has been discussed as if it were from some kind of holistic store, be it pre- or post-categorical, no evidence in the literature requires that it be. Report may just as well be based on activation of independent feature maps as on a single object representation. Certainly evidence that subjects are capable of selecting by features does not contradict this assertion. Unfortunately, evidence suggesting that iconic storage is post-categorical has in general used letters and digits to define categories, and these stimuli seem to behave differently than other shapes (Butler, 1980; Butler,

Browse, & Mewhort, 1986; Ceraso, 1985). It is possible that letters may be treated as features by a literate visual system. For the purposes of the present study I will not discriminate between letters and their component features when discussing feature report, because I chose letters that could be distinguished from each other by a single feature.

Feature integration theory suggests that information about conjunctions of features will not become available unless the observer attends to the location of a conjunction of features. An early holistic representation or map of locations and features may or may not exist, but even if it does it may not be accessible for report. When the selection of a particular location is delayed, as with a delayed cue to report, conjunctions of features may be constructed if the activation of the features in their respective feature maps is continuing and if information about the location of these features is preserved until they can be attended. If attention is prevented from selecting or retrieving the correct features, features from the display may be miscombined and illusory conjunctions may occur. If, on the other hand, location information is unavailable but information about feature combinations is available, subjects may report items that were present elsewhere in the display if these are already conjoined in some representation. Finally, if subjects are working from

a decaying trace that makes features increasingly confusable, they may produce a large number of intrusion errors.

If subjects have only an initial holistic representation of the objects in a display, as Lockhead (1972) suggests, then reports of conjunctions of features should be easier than reports of their component parts. With a delayed cue, Lockhead might predict that mislocations would occur more frequently, but he would not predict that illusory conjunctions would occur.

A third possibility is a compromise between an early holistic representation or only independent feature representations. Perhaps both representations exist, and we have separate feature representations and holistic representations. If this is the case, information about both conjunctions and features could be retrieved using either representation, although features would take longer to extract from a conjunction representation, and conjunctions would require time to be constructed from feature representations. However, accuracy could be improved at the expense of speed because if one representation was degraded the other might still contain enough information to report items from the display, although this would involve additional processing. If stimulus conditions were such that one type of report was

easier to use than the other, subjects might be able to strategically report from one or the other representation. If conjunctions were far apart in a holistic representation, but their component features were not, a conjoined representation would be easier to use. If the features were more discriminable on their own independent feature representations may be easier to use.

The present investigation tested the prediction from feature integration theory that no holistic representation of conjunctions of features exists until attention is focused on their location. When attention is constrained by using brief exposure durations, there will be no information about the conjunctions in a display in excess of what is available about their component features. In other words, there may be no additional holistic information arising from the pairing of individual features -- at least in the case where those features are separate visual elements in the coding used by early vision <u>and</u> where no emergent features (such as closure) are formed.

A partial report paradigm employing short stimulus durations and variable stimulus onset asynchronies (SOA) was used in these experiments. Subjects separately reported the color of displayed letters, the letter identity, or the conjunction of both color and identity. Color and shape generally behave as separable features (Garner, 1974) and

therefore these were chosen as features for the experiments. Both colors and letters were chosen for their discriminability. If features are represented independently, the probability of correctly reporting the conjunction of features should be the product of their independent probabilities of being correctly reported. In this case:

$$P(\text{conjunction}) = P(\text{color}) \times P(\text{letter})$$

I compared the observed values [p(conjunction)] with the predicted values (calculated from observed performance in the feature conditions, [p(color) x p(letter)]), to see whether any additional information specifically about the conjunction or pairing of features was available.

If any additional information was available, it would suggest that some type of holistic representation may be available.

Experiment 1

The first experiment was designed to test the prediction from feature integration theory that conjunction

performance could be accounted for by performance on the features forming the conjunction. This experiment also measured total report performance on color, letter and the conjunction of the two features.

Method

Subjects: Six volunteers from the Attention Lab subject pool were employed for five one-hour sessions for which they were paid four dollars per session. All subjects had normal or corrected-to-normal vision and none suffered from any known deficits in color vision.

Displays: The displays were shown on an Intelligent Systems Corporation Intecolor 8001 terminal (model BS001G) controlled by a Digital Equipment Corporation PDP 11/34 digital computer. They were made up of eight characters in a circular array that subtended a visual angle of 2.7 degrees horizontally and 2.3 degrees vertically. The eight characters appeared on a dark background and were composed of four colors: red, green, blue, and white, and four letters: T, O, S, and V. The letters and colors were chosen for their discriminability. The characters subtended an average visual angle of .6 degrees horizontally and vertically. The generation and presentation of the display was randomized with the following constraints: each color and letter appeared twice and no color-letter combination

was repeated within a trial. At the beginning of all trials, a fixation dot appeared in the center of the screen for 100 ms followed by a 700 ms blank interval, and then the display was presented for 100 ms. The cue to report an item was a purple dot which appeared adjacent to and outside of the position of the relevant item. The cued item was chosen randomly. The cue could appear at the following SOAs: -200 (a pre-cue), 0 (a simultaneous cue), 200, 300, 400, 600, and 1100 ms. These SOAs were mixed randomly within blocks. The cue was presented for 100 ms. After the appropriate interval the display, the cue and the display, or the cue alone was presented.

Procedure: Each subject participated in five fifty-minute sessions of the experiment. The first and last sessions were total report sessions, measuring total report for conjunction, color and letter, with trials in each condition blocked and separated by a rest period. The order of presentation of these report conditions was counter-balanced across subjects in a Latin square design. Subjects' heads were held at a fixed distance from the monitor by insuring that they used a chin rest. In the total report sessions, subjects received 24 practice trials before each condition and then 32 trials of data collection in each of the three conditions. This resulted in the same number of "trials per interval" as the partial report

conditions. In total report sessions, the cue appeared at one of the eight positions 200 ms before the display. Subjects reported all the information they could starting at one position counter-clockwise from the cued location. This instruction was given to induce subjects to spread their attention to either side of the cued location and was based on the assumption that information counter-clockwise from the cued location might be obtained and then forgotten by the time subjects worked around the perimeter of the circle. Subjects responded verbally and their responses were entered by the experimenter on the PDP 11/34 keyboard. Entering a response initiated the start of a new trial. This procedure created a variable inter-trial interval. In addition, at the end of the final session, subjects did a written version of the full report task, in which they wrote down the initial letter of the colors and the letters in their observed positions on a "clock face" diagram supplied by the experimenter. This condition was arranged as was the verbal total report, but subjects were free to report items in any order and no cue was presented. They received 16 practice trials and 24 test trials in this last sub-session. The written total report was hand scored by the experimenter, again for correct conjunction, color correct, letter correct or incorrect.

The three partial report sessions were run on consecutive days and consisted of 48 practice trials and 224 data collection trials. The procedure was as described above, except that subjects verbally reported only the item at the cued location, and SOAs varied as described above. The order of presentation of conjunction, color, and letter conditions was counter-balanced across subjects in a Latin square design.

Responses were scored with a scoring program on the PDP 11/34, and counted as either correct conjunction, color correct, letter correct, or incorrect, depending upon what portion of their responses matched the items displayed.

Results and Discussion

There was a partial report superiority for conjunction reports over total report at all SOAs except 1100 and a partial report superiority for the feature report at all SOAs for color, and at all except 1100 for letter (see Table I).

Performances in first and last total report sessions were not significantly different ($t[10] = 1.74$), and as a result we may conclude that practice with these displays had little effect upon total report performance. Superiority of partial report for conjunctions is much higher than that of

the features with a pre-cue (-200 SOA), but after that SOA, falls to the same level as the features.

Table I

Superiority of partial report over total report for Experiment 1.

SOA	Report	t	df	p	Advantage (# of items)
-200	Conjunction	31.77	16	< 0.01	5.91
	Color	12.39	16	< 0.01	4.76
	Letter	15.87	16	< 0.01	4.98
0	Conjunction	9.65	16	< 0.01	4.19
	Color	6.27	16	< 0.01	3.81
	Letter	9.11	16	< 0.01	3.77
200	Conjunction	2.86	16	< 0.01	1.06
	Color	2.01	16	< 0.05	1.56
	Letter	2.63	16	< 0.01	1.37
300	Conjunction	2.52	16	< 0.05	0.95
	Color	2.92	16	< 0.01	2.11
	Letter	3.08	16	< 0.01	1.61
400	Conjunction	2.04	16	< 0.05	0.78
	Color	1.86	16	< 0.05	1.51
	Letter	2.51	16	< 0.05	1.31
600	Conjunction	1.90	16	< 0.05	0.83
	Color	1.87	16	< 0.05	1.28
	Letter	2.19	16	< 0.05	1.34
1100	Color	2.16	16	< 0.05	1.47

The experimental apparatus in this experiment included a color monitor (the Intecolor 8001 monitor) with phosphors that persisted for varying amounts of time after display offset. In order to estimate the contribution of

persistence of the phosphors to task performance, I ran a version of the task using a shutter which blocked out the display and opened at the same onset times as did the cue. Three new subjects ran in this test. On average, subjects were able to report items at levels above chance up to 120 ms after the display offset without actually seeing the original display. When the reports from the persistence alone were subtracted from the obtained reports and they were again tested for partial report superiority the significance of the partial report superiority over total report remained unchanged.

A somewhat surprising finding was that performance on the verbal total report task was better than performance on the written total report task ($t[16] = 3.106$, $p < .01$). The written total report task was added to the first experiment because I was not certain that forcing subjects to start reporting from a particular location in the display would provide an adequate measure of all they could report from the display. Subjects volunteered a preference for the verbal total report task. It seemed that the presence of the cue helped subjects to direct their attention to a section of the display and gave them a starting point for reporting the items present. When this cue was absent they floundered.

The finding of most importance for this study was the lack of a significantly better performance on the conjunction report than would be expected from the independent report of features. The performance on the conjunction and feature reports and the difference between obtained and predicted conjunction performance are presented below (Figure 1). A repeated measures analysis of variance was performed to establish whether there was a significant difference between the conjunction performance as predicted from feature performance (henceforth called "between session"), and observed performance. The difference between observed and predicted performance did not reach significance ($F[1,5] = 1.04$ $p > 0.3$).

We can also look at the independence prediction within the conjunction report condition. Color and letter performance may be measured by adding correct reports of color only (i.e. trials where letter is reported incorrectly) to correct conjunction reports to obtain a measure of the total number of correct color reports. The same may be done with letter reports. The performance values for the features obtained in this way were multiplied to produce predicted values for conjunction performance for each subject (This will be referred to as "within session prediction"). Here I found a small advantage for observed conjunctions over predicted conjunction reports, averaging

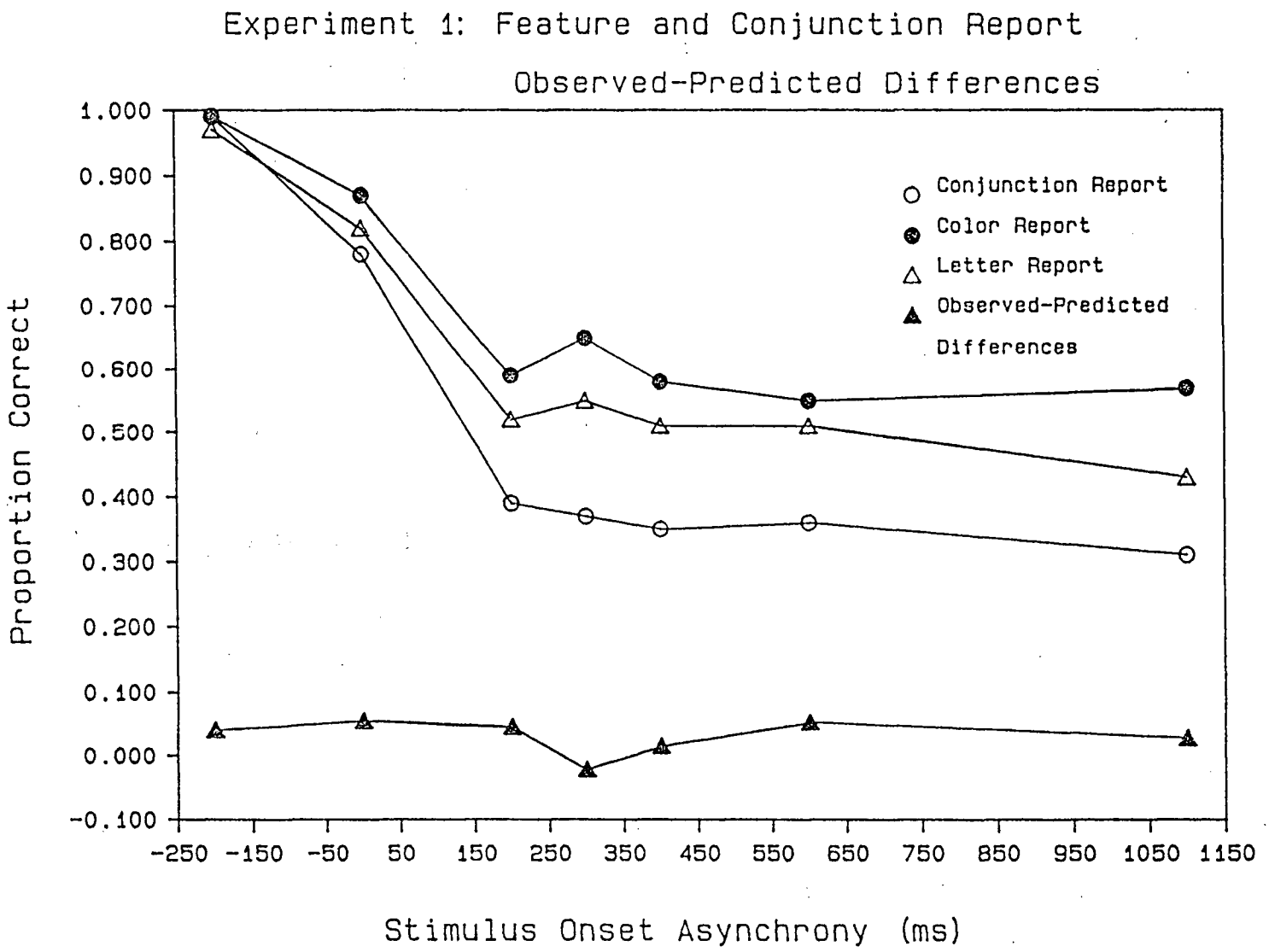
Experiment 1: Feature and Conjunction Report
Observed-Predicted Differences
Proportion Correct
1.000
0.900
0.800
0.700
0.600
0.500
0.400
0.300
0.200
0.100
0.000
-0.100
-250 -150 -50 50 150 250 350 450 550 650 750 850 950 1050 1150
Stimulus Onset Asynchrony (ms)
Conjunction Report
Color Report
Letter Report
Observed-Predicted Differences

Figure 1.

5.6 % and present at all intervals between cue and display ($F[1,5] = 63.6$, $p < .01$).

There was a significant interaction between the observation-within session prediction difference and SOA ($F[6,30] = 2.95$, $p < 0.05$). The interaction is accounted for by the near perfect performance with the advance cue and smaller standard deviations in the within session measure. This within session measure provides an upward bound for the availability of conjunction information. Within a session, state variables such as alertness will produce correlations between feature and conjunction report even if the representations are independent; for example, if subjects looked away from the display they would get neither feature nor conjunction correct. The same session report also controls for memory and response load, factors which could inflate scores when only color or letter are reported.

Experiment 2

Since the assumed duration of the icon may be as short as 150 ms with SOA rather than ISI the important variable (Di Lollo, 1977), the duration of the displays in Experiment 1 and the length of SOAs may have prevented a conjunction

advantage from being observed there. It is possible that the intervals examined passed over the time period where an iconic representation of conjunctions is present and accessible. Without looking at the intervening period it is not possible to conclude that conjunctions do not have an advantage over features. Experiment 2 was run to look at shorter SOAs. The duration of the display was reduced to 20 ms, and the first post-display SOA was shortened to 70 ms.

Method

Subjects: Four volunteers from the Attention Lab subject pool were employed for three one hour sessions for which they were paid four dollars per session. All subjects had normal or corrected to normal vision and none suffered from any known deficits in color vision.

Displays: The displays, cues and apparatus for this experiment were the same as for experiment 1. The cue could appear at the following SOAs: 0, 70, 100, 140, 320, and 1020 ms. The displays were presented for 20 ms, as were the cues.

Procedure: Subjects received 48 practice trials and 288 test trials in each session, resulting in 48 trials per condition. The order of presentation of conjunction, color and letter conditions was counter-balanced across subjects in a Latin square design. Responses were entered and scored

as in Experiment 1. The total report conditions were not run in this experiment.

Results and Discussion

The performance on the conjunction and feature reports and the difference between obtained and predicted conjunction performance are presented below (Figure 2). A repeated measures analysis of variance was performed to establish whether the differences between predicted and observed values were significant for the between session prediction and for the within session prediction. The between session prediction difference was not significant ($F[1,3] = 0.16$, $p > 0.71$) while the within session prediction difference was significant ($F[1,3] = 16.58$, $p < 0.03$). This suggests that conjunctions have only a small advantage, averaging 3.9 percent over all SOAs, over what would be expected from the performance on features.

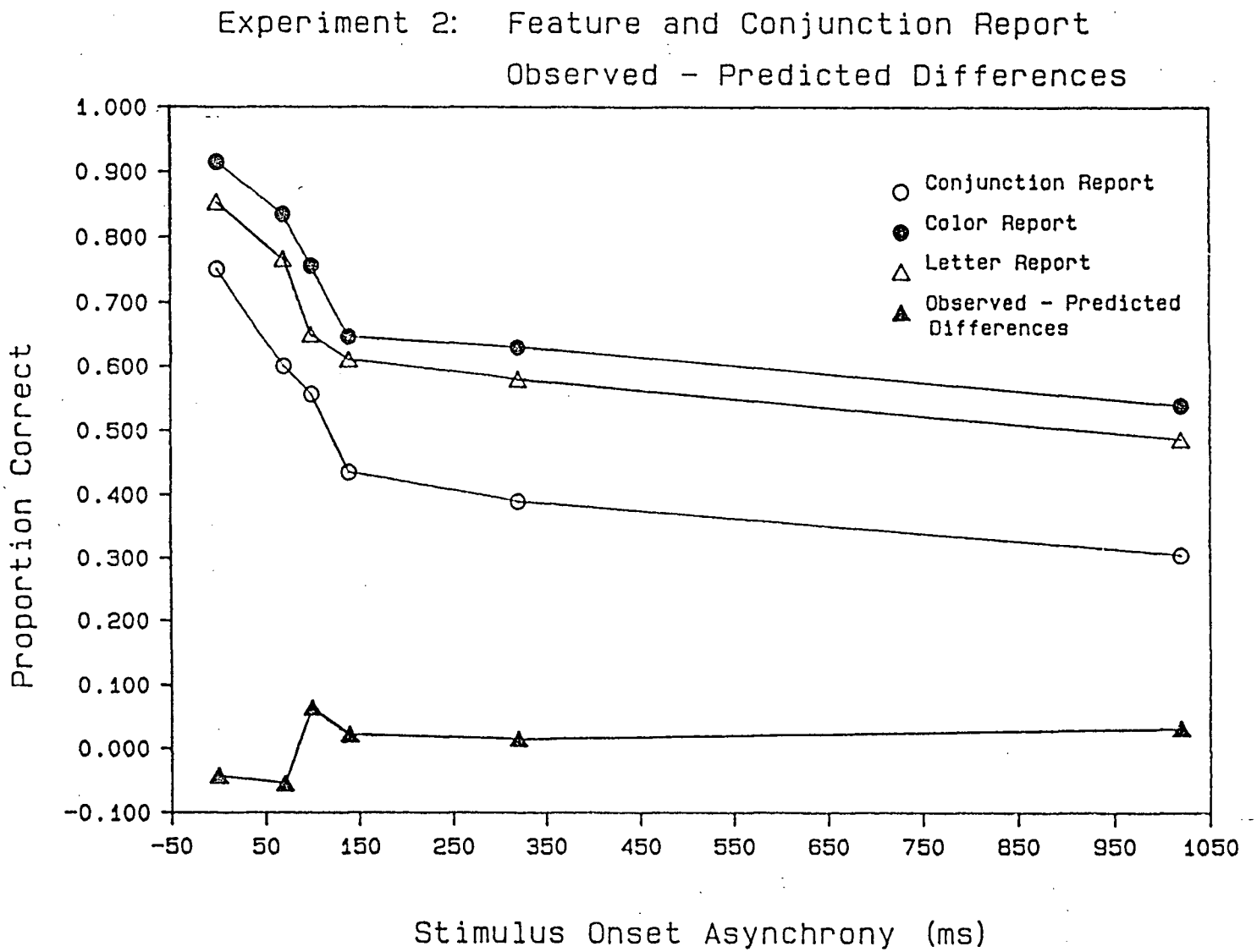
Experiment 2: Feature and Conjunction Report
Observed - Predicted Differences
Proportion Correct
1.000
0.900
0.800
0.700
0.600
0.500
0.400
0.300
0.200
0.100
0.000
-0.100
-50
50
150
250
350
450
550
650
750
850
950
1050
Stimulus Onset Asynchrony (ms)
Conjunction Report
Color Report
Letter Report
Observed - Predicted Differences

Figure 2.

Experiment 3

This experiment was conducted to confirm the results of the first two experiments. An initial pilot study was run to attempt to replicate the earlier results on a different color monitor system. This system had the advantage of possessing short persistence phosphors for all three color guns, allowing me to assess the contribution of the long persistence to performance. This pilot also switched subjects' responses from verbal to typed. The SOAs used were the ones which experiments 1 and 2 had in common, 0, 100, 300, and 1000 ms. Finally, the pilot employed a smaller display size, reducing the number of elements to four from eight. The reduction removed the redundancies in color and letter that were present in the earlier experiments, and thus should reduce within display confusions.

Results of the pilot replicated the first two experiments, and suggested that the long phosphor persistence was not responsible for the observed partial report superiority, and a third experiment was then conducted to extend the other results.

The third experiment was designed to look again at shorter SOAs and to allow a more detailed analysis of errors than was possible in the earlier versions. One of the five

levels of each feature was absent from each display in this version to allow location and intrusion errors to be compared. As well, two groups of isoluminous colors were constructed in order to determine whether subjects were responding to the color or to the luminance of presented items. Two cue durations, 50 and 500 ms were used in this experiment to determine whether or not short cue durations were interfering with subjects' ability to locate the correct display position for report. The long (500 ms) cue duration should allow ample time for the position of the cue in the display to be determined, and if there is no difference between the long and short cue durations, failure to located the cue may be discarded as a potential source of error.

Method

Subjects: Subjects were sixteen paid volunteers from the University of California (Berkeley) Attention Laboratory subject pool who were employed for two one-hour sessions. They were paid either five dollars (eight subjects) or seven dollars (eight subjects) per session for their participation. Three subjects were discarded for failing to achieve an accuracy of 75 percent with 0 cue delay, and one subject was discarded because of a computer malfunction during data collection that resulted in the loss of data.

All had normal or corrected-to-normal vision, and none suffered from any known deficits in color vision.

Displays: The displays were shown on a Mitsubishi High Resolution RGB color monitor controlled by an Artist Plus color graphics board installed in an IBM AT computer. Displays were composed of eight characters arranged in a circular array that subtended a visual angle of 5.84 degrees horizontally and 7.32 degrees vertically. These items averaged 0.69 degrees horizontally and 0.97 degrees vertically. The items in the display were selected from the color set red, green, blue, yellow, and purple, and from the letter set N, O, S, T, and X. Stimuli were shown on a dark background. Items were selected randomly with the following constraints: each color or letter selected appeared twice in the display with no color-letter combination repeated. This resulted in four colors and letters being present in a display and one color and one letter being absent. At the beginning of a trial a fixation dot appeared in the center of the monitor screen for 500 ms. This was followed by a 200 ms blank interval and then the display was presented for 50 ms. The cue to report an item was a white dot which appeared adjacent to and outside of the position of the relevant item in the display. The item cued was chosen randomly. The cue was 50 ms in duration for six of the subjects and 500 ms for the remaining six. Cues appeared at

the following SOAs: 0 ms (a simultaneous cue), 125 ms, 200 ms, 300 ms, and 1050 ms. The experimental room was lit with a 75 watt light bulb in a lamp directed upward to provide low level diffuse light.

Procedure: Each subject participated in two 50 minute sessions of the experiment. A single session was divided into three sections for each reporting condition, conjunction report, color report, and letter report, with the order of these counter-balanced across sessions in a Latin square design. The first session of the experiment was treated as a practice session and data were not analysed from this portion of the experiment. In the second session, subjects received 40 practice trials in each condition, and 200 data collection trials in each condition (40 trials per cue delay). This resulted in a total of 40 trials at each cue delay per condition. Subjects responded by entering the first letter of the color name, the letter, or the color-letter combination of the item at the cued location on the IBM AT keyboard. If subjects made a response which was not one of the possible responses, they heard a beep and saw an error message asking them to correct the typing error. They received no other error feedback. Six subjects viewed displays with their heads held against a forehead rest and hood which also shielded the monitor screen from glare. The remaining six subjects did not use this head rest.

Results and Discussion

The two groups of six subjects were compared to test for any significant differences in performance as a result of the different testing conditions (the different cue durations and the presence or absence of the hood). There was no significant difference between the groups ($F[1,10] = 0.05$, $p > 0.80$) so the two groups were collapsed into one for subsequent analyses.

The levels of feature and conjunction reports, and differences between obtained and predicted performance on the conjunction reports are presented below (Figure 3). A repeated measures analysis of variance was performed to establish whether the differences between predicted and observed values were significant for each type of prediction. The difference for the between session prediction was not significant ($F[1,11] = 2.92$, $p > 0.11$) while the difference for the within session prediction was significant ($F[1,11] = 48.81$, $p < 0.001$). The accuracy of feature report across the two conditions, conjunction and separate feature conditions was not significantly different for color reports ($F[1,11] = 1.42$, $p > 0.25$) but was significantly different for letter reports ($F[1,11] = 6.54$, $p < 0.03$).

A comparison of the frequency of different types of errors was made to see whether conjunction errors occur more frequently than one would expect if errors were simply the result of guessing. Following Treisman and Schmidt (1982), error responses were collapsed into conjunction errors (those errors where subjects recombined features from the display, i.e. illusory conjunctions), intrusion errors (errors where an item containing a feature that was not present in the display was reported), and errors where a reported item was present in that combination elsewhere in the display. These latter errors were separated from the other errors as they may have been misperceptions of the cue location rather than illusory conjunctions. The difference between conjunction errors and intrusion errors provides a measure of the number of illusory conjunctions actually being produced by the subject. The number of intrusion errors was corrected (weighted by a factor of 3) to compensate for their reduced probability of occurrence, since only one color and letter were absent from each display and three other (uncued) features were present.

This corrected intrusion error measure was then compared to obtained conjunction errors through a repeated measures analysis of variance.

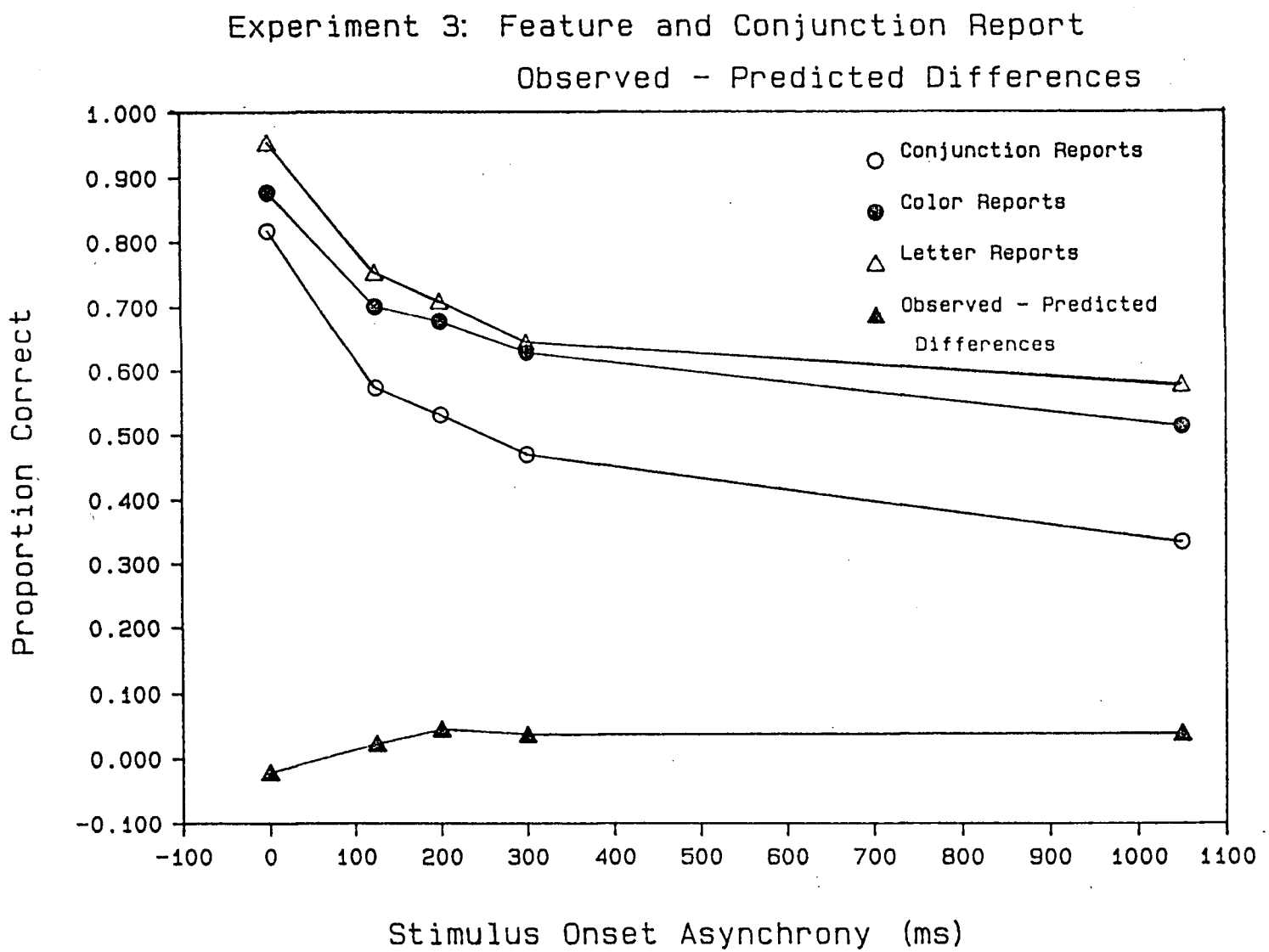
Experiment 3: Feature and Conjunction Report
Observed - Predicted Differences
Proportion Correct
1.000
0.900
0.800
0.700
0.600
0.500
0.400
0.300
0.200
0.100
0.000
-0.100
-100
0
100
200
300
400
500
600
700
800
900
1000
1100
Stimulus Onset Asynchrony (ms)
Conjunction Reports
Color Reports
Letter Reports
Observed - Predicted Differences

Figure 3.

The difference between conjunction errors and intrusion errors was significant ($F[1,11] = 10.96$, $p < 0.01$), with a significant interaction with cue delay ($F[4,44] = 2.85$, $p < 0.05$). This interaction was due to the low frequency of errors at 0 SOA, and when errors at this interval were removed from the analysis the interaction disappeared (see Figure 4).

A final analysis was the construction of confusion matrices for the two dimensions, color and letter. The purpose of this analysis was to determine if subjects were responding to the color or to the brightness of stimuli. These were confounded in the earlier experiments. Responses from all intervals except 1050, where responses include many guesses, were sorted according to what value of a dimension was presented and responded on a given trial. If subjects are responding based on the brightness rather than color of stimuli we should see more confusions within luminance categories than across. It is apparent in Table II that this is not the case. Rather, subjects made the predictable confusions between colors that were most similar (especially between blue and purple, and yellow and green). The numbers in the table represent proportions of responses that fell into each cell. The actual numbers of responses are only

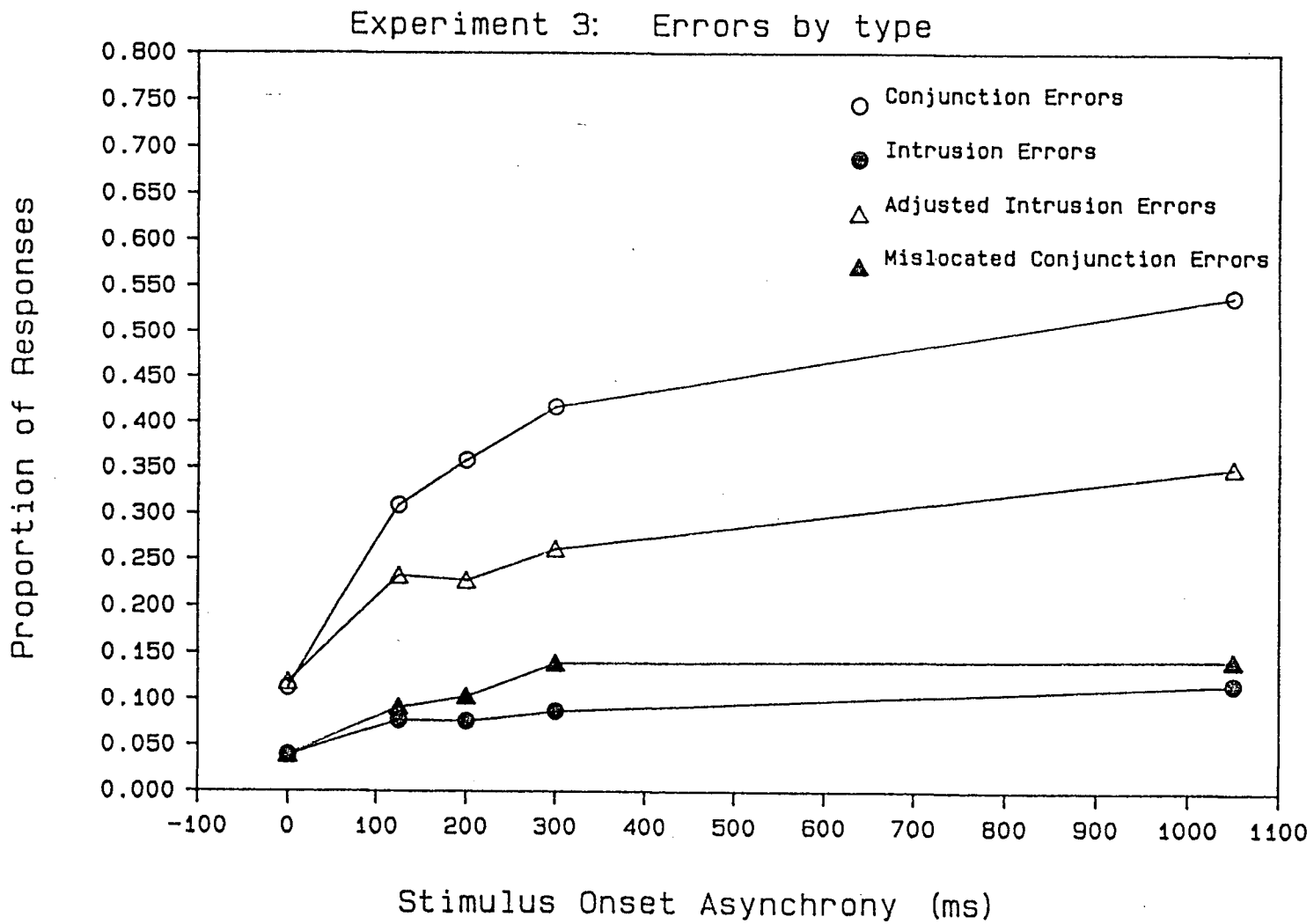
Experiment 3: Errors by type
Conjunction Errors
Intrusion Errors
Adjusted Intrusion Errors
Mislocated Conjunction Errors
Proportion of Responses
0.800
0.750
0.700
0.650
0.600
0.550
0.500
0.450
0.400
0.350
0.300
0.250
0.200
0.150
0.100
0.050
0.000
-100
0
100
200
300
400
500
600
700
800
900
1000
1100
Stimulus Onset Asynchrony (ms)

Figure 4

approximately equal, since cued items were chosen randomly on each trial.

Table II

Confusion Matrix for colors.

		RESPONSE				
		Luminance Categories				
		Category 1			Category 2	
		RED	GREEN	PURPLE	YELLOW	BLUE
STIMULUS	RED	.714	.059	.055	.050	.073
	GREEN	.039	.696	.059	.103	.104
	PURPLE	.032	.085	.530	.059	.295
	YELLOW	.071	.139	.050	.652	.074
	BLUE	.053	.080	.195	.040	.663

A similar analysis was performed on letter responses. Here it is not so clear what pattern of confusions to expect, since letters were chosen to have distinctive features. Still, some unexpected confusions occurred fairly frequently (see Table III). For example, S was more frequently reported as N than reported as O. The main confusions are understandable. X was confused with both N and S, and also was least accurately reported of all the

letters. Perhaps this is evidence that the intersection of the two lines is not a distinguishing feature. The other major confusion was O reported as S, but not S reported as O, which also agrees with data from search experiments that terminators are detected and closure is not (Treisman and Gormican, in preparation)

Table III

Confusion Matrix for letters

		RESPONSE				
		N	O	S	T	X
STIMULUS	N	.714	.073	.090	.063	.065
	O	.067	.694	.135	.067	.037
	S	.082	.055	.768	.047	.047
	T	.052	.041	.053	.808	.046
	X	.105	.065	.096	.078	.655

General Discussion

Summary of Results

What have we learned from these experiments? A review and collection of the results is the first step in answering this question. Table IV is a summary of the analyses of the difference between observed conjunction performance and the predicted performance from the feature and the within session predictions.

We find a consistent, though small, advantage for obtained conjunction performance against both types of prediction. This difference is significant in all three experiments for the within session prediction, and non-significant for the between session prediction. The average advantage for conjunctions across both types of prediction is 3.3 percent (2.0 percent for between session prediction and 4.6 for within session prediction). Collapsing the advantage may provide the best measure of the actual conjunction advantage, since each prediction on its own is problematic.

Table IV
Summary Table for Observed Performance
vs Predicted Performance
(Repeated Measures Analysis of Variance)

Comparison	df	*F*	*p*	% Difference
Experiment 1				
Between session	1,5	1.04	> 0.30	3.0
SOA (between)	6,30	39.29	< 0.001	
Within session	1,5	3.60	< 0.01	5.6
SOA (within)	6,30	34.16	< 0.001	
SOA - Within Session Interaction	6,30	2.95	< 0.05	
Experiment 2				
Between session	1,3	0.16	> 0.70	0.6
SOA (between)	5,15	14.16	< 0.001	
Within session	1,3	16.58	< 0.05	3.9
SOA (within)	5,15	7.80	< 0.001	
Experiment 3				
Between session	1,11	2.92	> 0.10	2.5
SOA (between)	4,44	74.84	< 0.001	
Within session	1,11	48.81	< 0.001	4.3
SOA (within)	4,44	45.97	< 0.001	
SOA - Within Session Interaction	4,44	3.49	< 0.05	

The within session prediction yields a reduced predicted conjunction performance since both prediction and observation were based upon the same trial. Any factor affecting the subject, whether it improves or disrupts performance, acts on both feature and conjunction performance in concert. If the subject blinked or looked away from the display both feature and conjunction were missed. The operation of multiplication in the generation

of the prediction may cause an advantage for conjunction performance simply as an artifact of the correlations in performance built into this measure.

The between session prediction suffers from a different set of difficulties. Because predictions are based upon performance across different blocks, variations in state are much larger. Indeed, we see a larger variance in this condition. A second factor is the lower load in the feature report blocks. Subjects were only required to report one thing about the display instead of the two in the conjunction condition. A check to see whether this made any real difference is to compare the levels of feature performance across these two conditions (see Table V).

Table V
Summary Table for Color Performance
and Letter Performance
(Repeated Measures Analysis of Variance)

Comparison	df	F	p	%Difference
Experiment 1				
Color	1,5	0.98	> 0.36	2.6
Letter	1,5	2.89	< 0.15	2.5
Experiment 2				
Color	1,3	0.03	> 0.88	0.3
Letter	1,3	7.53	< 0.08	2.1
Experiment 3				
Color	1,11	1.42	> 0.25	3.8a
Letter	1,11	6.54	< 0.03	6.0

[a] This difference was in the opposite direction to the others

While color reports between the two conditions are never significantly different (they are even better in the conjunction report condition in one experiment), letters seem to have an advantage when they are reported alone, while colors do not. This contradicts the data of Ceraso (1985), who found that letters were no better reported alone than they were in combination with color. The color-letter difference likely arises from two sources. Color is the more difficult of the two features here, with an overall lower accuracy.

The other source of the difference may be the order of reporting of features in the conjunction report condition. Because it is difficult for English speakers to put adjectives after nouns (e.g. "T red"), subjects always reported the color first and then the letter. If they were reporting from a rapidly decaying trace the information about the letter might have decayed before they could report it. Ceraso (1985), however, found no effect of order of report on accuracy for color and letter report. These two factors would inflate the difference between the two letter conditions while keeping color performance more or less the same. Because of the difference on letters, I cannot be confident that the between session prediction does not over predict conjunction performance (i.e. the measure

underestimates feature performance). For these reasons the combined estimate of conjunction advantage seems the best measure of actual conjunction advantage.

In experiment 1 a clear partial report superiority was obtained up to 600 ms SOA for conjunction performance. While the total report performance from experiment 1 cannot be statistically compared to the partial report in the other two experiments the actual numbers obtained may be roughly compared.

Since the trials from which the total report measure was taken are those where subjects saw the longest duration displays and were the most practiced at the task it is reasonable to assume that total report under any of the other conditions would not exceed that obtained here. If this is the case we have a partial report advantage across all experiments at all SOAs up to 600 ms. Certainly, there is a clear partial report advantage to 300 ms SOA (see Table VI). A partial report advantage may also be inferred from the highly significant effect of cue delay in all experiments (see Table IV).

Table VI

Summary Table for Partial Report Performance on Conjunctions Across Experiments

SOA	average number of items reported	SD
Whole Report	2.05	0.444
Experiment 1		
-200	7.96	0.098
0	6.24	1.979
200	3.11	1.329
300	3.00	1.387
400	2.83	1.440
600	2.88	2.007
1100	2.52	1.823
Experiment 2		
0	6.00	1.876
70	4.80	2.473
100	4.46	1.833
140	3.48	1.447
320	3.12	1.450
1020	2.46	0.790
Experiment 3		
0	6.53	0.996
125	4.58	0.963
200	4.25	1.151
300	3.75	1.086
1050	2.67	0.892

The Conjunction Advantage

The major hypothesis of this thesis was not strongly supported by the results. The advantage for conjunction reports over what was predicted by independent feature reports is 3.3 percent averaged over both types of prediction. The direction of this difference is consistent across experiments and types of prediction, and so it seems likely to be reliable. Subjects are reporting more than they should be able to if all they have access to are separate feature maps, yet at the same time the advantage is quite a small one, and is not significant across both types of prediction.

A possible source of the conjunction advantage (other than the existence of a conjoined representation) may be found in the nature of the task subjects performed. Feature integration theory states that features may be detected without knowledge of their exact locations, while conjunctions must be located to be reported. Attention in this paradigm is selecting items on the basis of their spatial locations, so individual features may not be more available than combinations of features. This would lead to an under-reporting of features relative to conjunctions and relative to their true availability. Since report of a

correct conjunction depends on the correct extraction of both features composing the item, feature report should always exceed conjunction report, and this is true of the present data. However, if features are being under-reported in relation to their true availability, then the prediction of conjunction performance obtained from the product of the observed feature performance actually underpredicts the expected conjunction performance. Since the obtained conjunction advantage was so small, and only one type of prediction was significant, this possibility cannot be rejected.

It is possible that once more permanent representations are established, combinations of features are tied together in a holistic representation. The object files proposed by Kahneman and Treisman (1984) provide a temporary representation that could serve this purpose. Do we have any reason to believe that features do not continue to be represented independently? Ceraso (1985) suggested that letters and colors are always stored independently. He suggested this because letter-color combinations did not show the same pattern of responses as other shape-color combinations when subjects were asked to report them. Recall of letter-color combinations did not differ when presented either in the same object or in two different objects. When subjects performed the same task with

geometrical forms and colors they were more accurate when the features were components of the same stimulus rather than located in separate items. These results support other evidence on the primacy of objects for the geometrical forms (e.g. Kahneman & Henik, 1981), but contradict them in the case of letters. The results hold for both a recall test and for a straightforward report from brief displays (i.e. less than 50 ms). Ceraso suggested that letters have a dual nature. As well as being forms, they have meaning as symbols, and map onto the phonetic structure of the languages that use them. For literate people, letters may be very easily encoded in representations that are simpler to recall than other less meaningful shapes. This representation is different from a feature map, but still may be quite independent of other attributes of the stimulus, since the symbolic value of the item is what is extracted.

There is no evidence in the present data that any conjoined representation of the features in the display becomes more available over time. At least from about 100 ms SOA, there is little change in the advantage of conjunctions over the predicted value. This lack of effect of SOA on the observed-predicted difference is somewhat surprising. The difference is more or less constant over all SOAs except where near perfect performance on all the

reporting conditions reduces the observed-predicted difference to nearly zero. If either a forming or a decaying conjunction representation is assumed, one would expect to see some change in the relative advantage of conjunction reports.

Other evidence indicates that, at least in short-term memory, information about the identity of items and their locations in the visual field are stored together (Cumming & Coltheart, 1969). This data is for the retention of information about single features, however, rather than of conjunctions of features.

So we do not see strong support for the notion that an enduring conjoined representation is established over time. This conclusion does not fit well with phenomenology. We are aware of objects in the environment as coherent wholes, and the same is true of recall of a remembered object.

Errors

The detailed error analysis in experiment 3 allows us to reach several conclusions relevant to current issues in iconic memory. One such issue is the role that identity and location information play in the retention and loss of information from iconic memory.

There is some evidence that identity and location information may not follow the same time course. Clark, (1969) found that her subjects were not able to report colors by their location in a display, but were able to report the presence of items of a certain color. Townsend (1973) found that subjects produced no partial report advantage when a recognition rather than a recall paradigm was used. The use of recognition instead of recall is problematic, since it is unlikely that the same processes are used in both, which makes comparisons between the two difficult to assess. In Townsend's design (letter recognition) it is possible that subjects were matching stimuli to probe letters based upon distinctive features rather than identity, so the lack of effect may show that information about the stimulus is maintained, but not the exact nature of that information. Townsend concluded that spatial and identity information were stored independently and follow a different time course.

In addition, Townsend compared long and short cue durations. A long duration cue (900 ms) harmed, rather than improved, performance when report was cued by location. This was somewhat surprising, since a plausible reason for the existence of location errors is misperception of the cue location exacerbated by short cue durations. The present result (experiment 3), which employed a 500 ms cue for one

half of the subjects and a 50 ms cue for the other half, showed no difference in performance between the 50 ms and the 500 ms cues. This appears to contradict Townsend's result, but 900 ms is a much longer cue duration and may hamper performance for other reasons. The result that was replicated was the lack of an advantage for the longer cue duration.

Is Iconic Memory Pre-categorical or post-categorical? Can the error analysis in experiment 3 shed any light on the debate about where attention intervenes in iconic memory? I believe it can, but only indirectly. If we assume that all items in a display are automatically encoded and identified then all items present in a display should be processed whether or not they are attended. If identity information (post-processing) decays, then we would expect to see the number of intrusion errors increase relative to the number of location errors. Conversely, if location information decays we expect an increase in location errors. This latter position is the one which has been supported by several theorists (Townsend, 1973; Butler, 1980). Treisman and Gelade (1980) found that location information played a different role when a task required detecting a feature in a briefly presented display than it did when the task required the detection of a conjunction. Treisman and Gelade looked at the conditional probabilities for correct detection of a

target given that it was mislocated. Mislocation did not affect the accuracy with which feature targets were reported, but it did affect the accuracy for conjunction targets.

In the present study we can take a fine-grained look at errors subjects are making. Some location errors are indeed mislocations of entire items, where color and letter move together. These, however, occur less frequently than the adjusted intrusion errors. The most frequently occurring error is one where both color and letter are mislocated within the display, but they are independently mislocated. These are conjunction errors, or illusory conjunctions, since they are combinations of features that were not in fact present in the display (see Figure 4).

The presence of this large number of conjunction errors suggests that iconic memory does not contain a unitized and identified representation of the items in a display. Rather, it lends support to the idea that what is stored is an independent representation of the features present and separate information about their locations. This information must be combined at a later point for the conjunction to be correctly reported. It also confirms the result that information about items in a display, be that information about features or the identities of entire objects, has a different time course.

Extensions of the research

Ashby and Townsend (1986) proposed a series of tests that provide converging evidence for perceptual independence. The tests start with a test of sampling independence, and progress to a check of the conditional probabilities of getting a feature correct given each possible value on a second dimension. If the data set passes all of their proposed tests, it is assumed to demonstrate the existence of perceptual independence for the dimensions and process under consideration. Since perceptual independence must be inferred rather than observed only a strong case based on converging evidence can justify a conclusion of perceptual independence. The present analyses reflect a step of Ashby's and Townsend's procedure, but not the complete series.

A logical extension of this research is to re-run the same paradigm using integral features instead of separable ones. If integral features are unitary as is claimed, then we should see a substantial advantage for conjunction report over feature report. If these stimuli are truly holistic, then it may be difficult for subjects to report features independently and feature report may actually be lower in accuracy than conjunction report.

A second extension is a simple check to see if the type of attentional manipulation involved in this task is the only one that would work. Briand and Klein (1987) claim that only drawn attention (orienting), which they contend is automatically attracted to peripherally appearing stimuli, affects feature integration. They found that centrally directed attention (i.e. with a directional pointer) did not improve the conjoining of features, while peripheral attentional cues did. A simple replication of the present results with central cues to report should show no advantage for the conjunction report condition if they are right. If they are not, the results should not differ from those reported here.

Another extension of the present work is to develop a paradigm that will allow the measurement of both feature and conjunction report without necessitating that each of these be located before they are reported. Using a simple recognition paradigm is not the best solution. This has been tried by other researchers, but a direct comparison of recognition and recall is problematic. The creation of a directly comparable situation is difficult; to use the current paradigm one would need equal numbers of feature elements present and absent from a display (otherwise the task can become "match the missing element"). This requirement is impossible to satisfy with colored letters as

stimuli. Either the number of colors required makes them become indiscriminable or the display sizes must be kept so small that the task is much to easy.

A possible paradigm is a modification of the one which Eriksen and Collins (1967) used. They presented subjects with two displays of apparently random dots. If subjects could integrate the two displays, they saw a three letter nonsense syllable. If they could not, they saw only random dots. Stimuli could be created to produce a pattern that is only discriminable when integrated, and this pattern could be composed of either feature or conjunction stimuli. If conjunctively defined items can be integrated to form a pattern, then a unitary representation of the first display stimuli would be implied. This would be strong support for an early integrated conjunction representation.

Conclusion

To end the discussion, I will summarize the conclusions drawn from this research. In the present investigation I found a substantial and significant partial report advantage for both conjunctions and features that were cued by spatial location. This superiority extended to 600 ms SOA, which is longer than is usually the case.

Across the three experiments I observed a slight advantage for conjunction reports relative to the predicted

value from the independent feature reports. The advantage averaged about three percent and was non-significant for the between session prediction, but significant for the within session prediction. The message from these measures is thus equivocal. However, this result is balanced by the observation that subjects are making a large number of conjunction errors relative to intrusion errors and mislocated conjunctions. The hypothesis that is best supported by these results is the compromise position discussed in the introduction. This is the possibility that both independent and holistic representations exist. A puzzle is why conjunctions do not exhibit a clearer advantage than the three percent indicated here. It is possible that subjects are reporting conjunctions from a very early representation, perhaps even from maps that encode combinations of color and orientation.

References

Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. Quarterly Journal of Experimental Psychology, 24, 225-235.

Anderson, N. H. (1981). Foundations of Information Integration. New York: Academic Press.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of Perceptual Independence. Psychological Review, 93 (2), 154-179.

Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. The Bell Systems Technical Journal, 40, 309-328.

Biederman, I., & Checkosky, S. F. (1970). Processing redundant information. Journal of Experimental Psychology, 83, 486-490.

Butler, B. E. (1980). Selective attention and stimulus localization in visual perception. Canadian Journal of Psychology, 34, 119-133.

Butler, B. E., Browse, R. A., & Mewhort, D. J. K. (1986). Do letter features migrate?: It depends on the task. Paper presented at the meeting of the Canadian Psychological Association, Toronto, Ontario, June, 1986.

Ceraso, J. (1985). Unit formation in perception and memory. The Psychology of Learning and Motivation, 19, 179-210.

Clark, S. E. (1969). Retrieval of color information from preperceptual memory. Journal of Experimental Psychology, 82 (2), 263-266.

Coltheart, M. (1980). Iconic memory and visible persistence. Perception and Psychophysics, 27, 183-228.

Coltheart, M. (1983). Iconic memory. Philosophical Transactions of the Royal Society of London (Series B), 302, 283-294.

Cumming, G., & Coltheart, M. (1969). Short-term memory: Are item and position information stored independently?. Psychonomic Science, 15 (4), 216-218.

Davidson, M. L., Fox, M. J., & Dick, A. O. (1973). Effects of eye movements on backward masking and perceived location. Perception and Psychophysics, 14, 110-116.

De Valois, R. L., & De Valois, K. K. (1980). Spatial vision. Annual Review of Psychology, 31, 309-341.

Di Lollo, V. (1977). Temporal characteristics of iconic memory. Nature, 267, 241-243.

Eriksen, C. W., & Collins, J. F. (1967). Some temporal characteristics of visual pattern perception. Journal of Experimental Psychology, 74, 476-484.

Eriksen, C. W., & Schultz, D. W. (1978). Temporal factors in visual information processing: A tutorial review. In J. Requin (Ed.), Attention and Performance VII, Hillsdale, N.J.: Erlbaum.

Garner, W. R. (1974). The processing of information and structure. Potomac, M.D.: Lawrence Erlbaum.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. Cognitive Psychology, 1, 225-241.

Gibson, J. J. (1959). Perception as a function of stimulation. In S. Koch (Ed.), Psychology: A study of a science, 1, 456-501.

Haber, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. The Behavioral and Brain Sciences, 6, 1-11.

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurons in the cat's striate cortex. Journal of Physiology, 148, 574-591.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. Journal of Physiology, 195, 215-243.

Kahneman, D. (1973). Attention and Effort. Englewood Cliffs, N.J.: Prentice-Hall.

Kahneman, D. & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy and J. R. Pomerantz (Eds.), Perceptual Organization. Hillsdale, N.J.: Erlbaum.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman, R. Davies, and J. Beatty (Eds.), Varieties of Attention, New York: Academic Press.

Lockhead, G. R. (1966). Effects of dimensional redundancy on visual discrimination. Journal of Experimental Psychology, 72, 95-104.

Lockhead, G. R. (1972). Processing dimensional stimuli. Psychological Review, 79 (5), 410-419.

Merikle, P. M. (1980). Selection from visual persistence by perceptual groups and category membership. Journal of Experimental Psychology: General, 109, 279-295.

Pashler, H. (1984). Evidence against late selection: Stimulus quality effects in previewed displays. Journal of Experimental Psychology: Human Perception and Performance, 10 (3), 429-448.

Prinzmetal, W., Presti, D., & Posner, M. (1986). Does attention affect visual feature integration?. Journal of Experimental Psychology: Human Perception and Performance, 12 (3), 361-369.

Sperling, G. (1960). The information available in brief visual presentations. Psychological Monographs, 74 (11).

Sperling, G. (1967). Successive approximations to a model for short-term memory. Acta Psychologica, 27, 285-292.

Townsend, V. M. (1973). Loss of spatial and identity information following a tachistoscopic exposure. Journal of Experimental Psychology, 98 (1), 113-118.

Treisman, A. M. (1969). Strategies and models of selective attention. Psychological Review, 76, 282-299.

Treisman, A. M. (1982). Focused attention in the perception and retrieval of multidimensional stimuli. Perception and Psychophysics, 22 (1), 1-11.

Treisman, A. M. (1986). Properties, parts, and objects. In K. R. Boff, L. Kaufman, & J. P. Thomas, (Eds.) Handbook of Perception and Human Performance: Cognitive Processes and Performance, (Vol. 2), New York: John Wiley.

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

Treisman, A. M., & Gormican, S. (in preparation). Feature analysis in early vision: Evidence from search asymmetries.

Treisman, A. M., Russell, R., & Green, J. (1975). Brief visual storage of shape and movement. In P. M. A. Rabbitt & S. Dornic (Eds.), Attention and Performance V, London: Academic Press.

Treisman, A. M., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. Cognitive Psychology, 14 107-141.

Treisman, A. M., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. In S. Dornic (Ed.), Attention and Performance, VI, Hillsdale New Jersey: Lawrence Erlbaum.

Von Wright, J. M. (1968). Selection in visual immediate memory. Quarterly Journal of Experimental Psychology, 20, 62-68.