1989

Subitizing And Counting: Preattentive And Attentive Processing In Visual Enumeration

Lana M. Trick

Subitizing and counting: Preattentive and attentive processing in visual enumeration

by
Lana M. Trick

Department of Psychology

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
December, 1989

List of Tables

Abstract

Subitizing, the process of visual enumeration when there are fewer than four items, is rapid (40-100 msec/item), accurate and effortless. In contrast, counting, the process of enumerating more than four items is comparatively slow (250-350 msec/item), effortful and error prone. Why does this occur? In this paper an attempt is made to incorporate subitizing and counting into a general theory of visual perception and spatial attention, as espoused by Marr(1982), Ullman(1984), and Treisman(1988). In particular, it is argued that the rapid apprehension of number in the 1-4 range is parasitic on a preattentive limited capacity mechanism that individuates feature clusters by assigning spatial reference tokens or FINSTs to them(Pylyshyn, 1989). These spatial reference tokens permit the identities of a small number of items to be maintained though their properties and retinal coordinates change, a capability important for directing the attentional focus and coordinating eye and hand movements. If the subitizing process makes use of such preattentive information, then it should not be possible to subitize when spatial attention is required to compute spatial relations, resolve the item as a whole or discern items to be counted from other distractor items. Thus, it was predicted that the slope of the latency function in the 1-4 range should approximate that of the 5+ range if spatial attention is required to perform a particular enumeration task. In contrast, it was predicted that subitizing should be possible when preattentive information could be used to distinguish the items to be counted from one another, or from other distractor items. Therefore, it was predicted that there should be discontinuities in slopes of the latency function between the 1-4 and 5+ range, as shown by deviations from linearity in trend analysis.

Five experiments were performed. In the first, subjects were shown capable of subitizing when the task was to enumerate items of a particular colour though they were not capable of subitizing when the task was to enumerate items that were connected to each other by a contour. This result was expected because spatial attention is presumed

necessary to compute the connected relation (Ullman, 1984; Jolicoeur, 1988). The second pair of experiments showed that though subjects can easily subitize when items are defined by groups of contours instead of simple edge points, they cannot easily enumerate such items if they are concentric, as would be predicted given that preattentive grouping processes would cluster all the contours into a unit in this case. The fourth and fifth experiments show that subjects can subitize certain target items in a field of distractors, but only if the property that differentiates targets from distractors is a feature, or a property thought to emerge preattentively. In situations where attention is required to form a unified object description by joining different dimensions (e.g., colour and orientation), or by joining different parts of an item (e.g., an O and a stem to form a Q), subitizing was not apparent. Overall, these experiments suggest that the subitizing process relies on preattentive information.

Acknowledgements

First I would like to thank my advisor, Dr. Zenon Pylyshyn, who has provided me with many opportunities for research and the exchange of ideas with people from different universities. Zenon been very supportive over the years as I hacked my way through successive approximations of my thesis, and has always maintained a sense of humour though at times the empirical situation was dismal. I would also like to thank Linda Pylyshyn for the many celebrations she has helped stage for us, and her unusual deserts.

I would also like to thank my external examiners, Dr. Bob Mercer and Dr. Doug Mewhort, for serving on my committee. In addition I would like to thank Dr. Rick Gurnsey and Dr. Keith Humphrey for their help and advice. In particular, I would like to thank Rick for his painstaking attention to detail and discussions about texture segregation research and Keith for his infectuous enthusiasm and broad ranging interests, from Chaos theory to sex differences in olfactory sensitivity to pig sweat. In addition I would like to thank the members of my advisory committee, Dr. Albert Katz and Dr. Jamie Campbell. Albert has always provided me encouragement and sound advice on experimental design and writing, both as advisor on Masters thesis and on the present thesis. I would like to thank Jamie Campbell and Val Thompson for their friendship, many wonderful meals, and all the times they have tried to teach me how to play Bridge.

I would also like to thank the members of the Pylyshyn lab, Jacquie Burkell, Roy Eagleson and Richard Wright, for their many long hours counting in the dark (without pay) as I was developing my experiments, and for providing a lab atmosphere with just the right combination of scientific zeal and goofiness. Also I would like to thank Inese Krammins, who keeps the Pylyshyn lab running, and whose level headedness in the face of pandemonium and crisis has got us all through many deadlines.

There are a number of other people I would like to thank, graduate students and regular patrons of the hospital cafeteria, Ramona Bobocel, Rae Gilchrist, Joanne Harbluk, Sandra Houston, Randy Paterson and B.J. Rye, for years of companionship and jokes about the food. I would especially like to thank Chris Roney, who has heard, over the years, more about subitizing than any human being ever should.

Finally, I would like to thank my parents, Gerry and Lucianne Trick, for their unfailing support despite the fact that at times it seemed like I would never get out of graduate school.

Table of Contents

List of Figures

List of Tables

Chapter One
Introduction

The topic of this investigation is vision, in particular the coordination of parallel and serial, preattentive and attentive, stages of visual analysis. Coordinating these stages has been thought to involve item individuation, the process of distinguishing feature clusters from each other (Pylyshyn,1989). My goal is to learn about item individuation by studying visual counting, a task that by its nature requires individuation because of the need to distinguish items already counted from those yet to count. For this reason it is necessary to think about what goes on when people count.

Consider the following situation. You are seated in front of a display and your task is to say how many dots there are as fast as you can, with accuracy. Vocal response latency and error rate is measured. As a subject you may notice that there is something qualitatively different between the experiences of enumerating small and large numbers of dots. When there are small numbers of dots, as in Figure 1-1a, enumeration seems effortless and immediate; you simply "see" how many dots there are. When there are large numbers, as in Figure 1-1b, enumeration seems slow and laborious; you may be conscious of grouping the dots into clusters, and then moving from cluster to cluster, finding the number of dots, and adding it into a running total (Warren, 1897; Shrager, Klahr, and Chase, 1983).

Not only do the two experiences "feel" different, but there are also different associated latency and error functions. When there are small numbers of items, the slope of the latency function is shallow; for adults each additional item may add a constant between 40 - 100 msec in the 1-4 range, for example. When there are large numbers of items in the display, the slope is large; each additional item may add a constant between 250 - 350 msec when there are 5 or more. See Figure 1-2a. Error functions reveal that subjects rarely, if ever, err when enumerating small numbers of items. People seldom

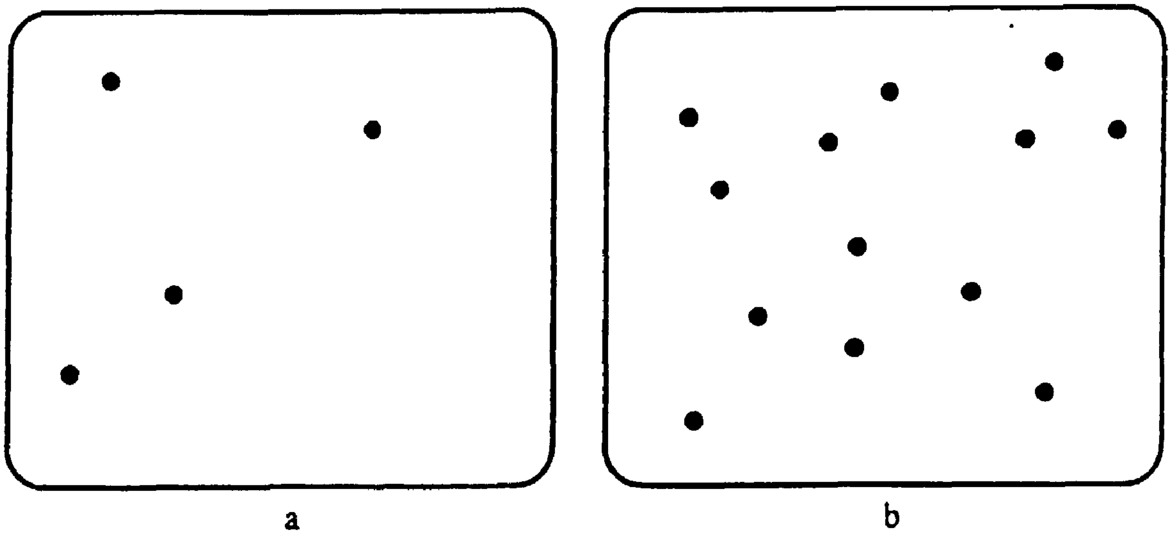

Figure 1-1: Dot displays

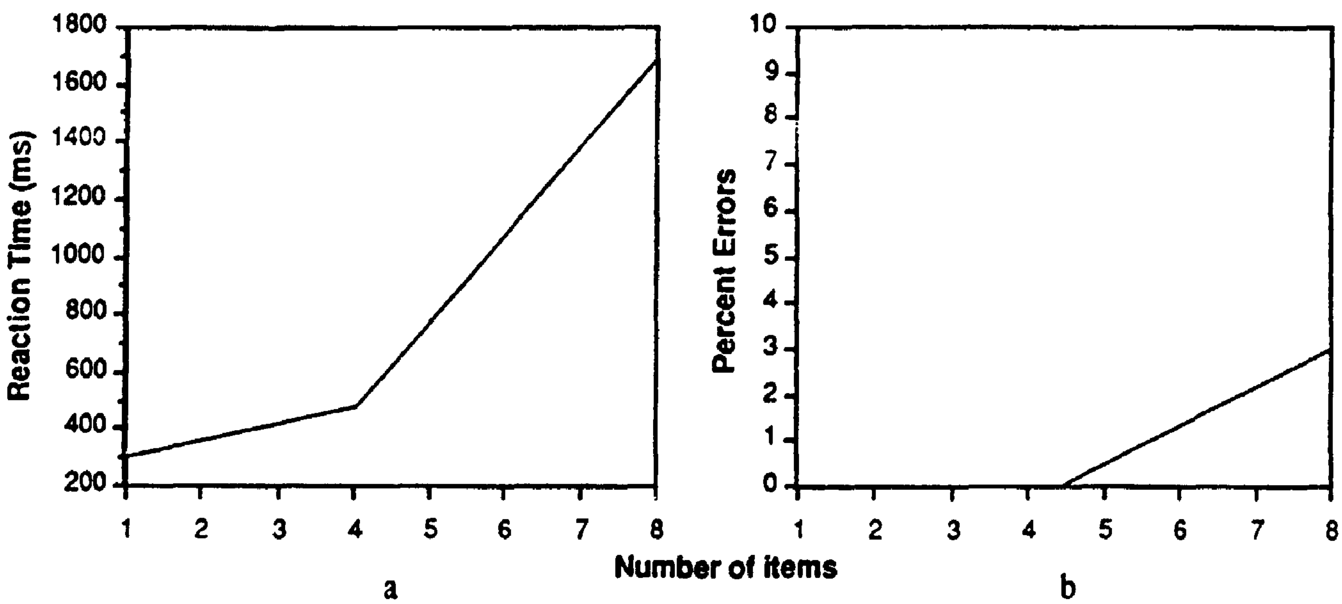

Figure 1-2: Idealized latency and error functions for subitizing and counting

make mistakes counting to 4, for example. When there are more items, however, the error rate gradually increases with number. See Figure 1-2b.

These findings are usually interpreted as evidence that there are two enumeration processes. One process is specialized for small numbers of items and is fast, accurate and effortless. This process has been called *subitizing* (Kaufman, Lord, and Reese, 1949). The other process can handle large numbers of items, but is slow, effortful and error-prone. This process has been called *counting*.[1] The point at which the slope in the latency function changes, the "elbow" in the reaction time curve, is taken to be the boundary between the subitizing and counting ranges. There has been controversy about exactly how many items can be subitized; estimates vary between 1-3 and 1-7, in part reflecting differences in stimuli and the procedures used for pre-processing data (Mandler and Shebo, 1982). There are also individual differences in how high people subitize, however (Akin and Chase, 1978). Nonetheless, the modal estimate from dot enumeration studies is that most adults subitize to 4 (Aoki, 1977; Klahr, 1973a; Oyama, Kikuchi and Ichihara, 1981; Atkinson, Campbell and Francis, 1976; Simons and Langheinrich, 1982). Thus, the subitizing range is most commonly considered to be 1-4 whereas the counting process is thought to take place when there are 5 or more items.

Although the idea that there are two enumerations processes is not new, no one has satisfactorily explained WHY two enumeration processes are necessary. We have a very fast, accurate process for enumeration--subitizing. Why can't we subitize any number of items? In this thesis I will be arguing that subitizing exploits a limited capacity parallel mechanism for item individuation, the FINST mechanism (Pylyshyn,1989). When the capacity of this mechanism is exceeded a serial process is employed, requiring spatial attention.

---

[1] The terminology is inherited from previous research. Unfortunately, it is confusing: "Counting" is used both to refer to a particular psychological process and the task of enumeration in general. When possible I will use "enumeration" to refer to the task of counting and reserve "counting" for the process.

In order to develop this argument I will first talk about vision in general, building a broad theoretical framework. This framework is based primarily on the work of Marr (1982), Ullman (1984) and Pylyshyn (1989; Fodor and Pylyshyn, 1981), but incorporates some of the research on attention, that of Treisman (Treisman and Gelade, 1980), in particular. Second, I will relate this theoretical framework to the phenomena of subitizing and counting. Finally I will present a series of experiments testing the hypothesis that the subitizing process depends on a preattentive mechanism.

A General Theoretical Framework for Vision

Broadly conceived, visual processing can be thought to have two goals. The first is to provide a description of the world that is suitable for the task of perceptual motor coordination; we need a representation of surfaces that will permit us to move around in the world without bumping into obstacles, falling into pits or being hit by projectiles. The second goal is to provide a description of the world suitable for the task of object recognition, classification and naming; we need a representation that will enable us to identify objects in different contexts, so that we will be able to behave appropriately. Thus, although we may never have seen a skunk except in pictures, we will recognize one if we encounter it, and know, for example, that patting it on the head would be a bad idea. Accomplishing these goals involves deducing the three dimensional structure of the world from its two dimensional projection, mapping back from the proximal stimulation on the retina to the distal properties of objects in the world. This task is difficult because in general there is no 1:1 mapping between properties of objects in images and properties of objects in the world. Consider object shape, for example. The same object can have many different retinal projections, and many different objects can cast the same retinal projection (Rock, 1983). Consequently, one pattern of retinal stimuliation may be interpreted in many different ways. The question is, how does the visual system make the right interpretation?

In trying to understand how the visual system functions, I will draw on research from two traditions. The Computationalist tradition offers insights into vision as information processing task: What are the goals of visual processing? What sort of information is available to the system and how can it be used to accomplish these goals? What sort of operations need to be performed on the information to most efficiently accomplish these goals? The second tradition, grounded in empirical research on attention, offers insights into human capacities. The attention research was engendered with the observation that people have little consciousness of information that they are not attending to, and they often seem incapable of responding to unattended information even if they would like to (James,1890). This lead researchers to posit a mechanism called attention that served to screen out irrelevant information. There are different views of the how attention works, however, and these views carry tacit assumptions about the nature of visual processing, and object recognition in particular. According to one, the *Early Selection View*, attention serves to select stimuli for full perceptual processing; only attended stimuli are processed to the point of recognition, and these stimuli are processed one at a time. Object recognition is thus assumed a difficult and time consuming task, one not wasted on the (probably irrelevant) background. Unattended stimuli receive only cursory analysis; only simple physical properties are derived, such as luminance, colour, orientation, size, motion and depth, for example (Treisman, 1985). These preattentively derived properties are called *features*.

Computationalists have broken the task of vision into three stages (Pylyshyn,1989), and these stages correspond, in part, to those implicit in the Early Selection View of attention. See Figure 1-3. The first stage, low level vision, takes information about light at each point on the retina from the photoreceptors and produces a representation that makes edge segments explicit. This type of analysis corresponds to preattentive processing in the *Early Selection* view of attention. The outputs of low level vision have been called the Primal and 2 1/2 D sketch by Marr (1982) or feature maps by attention

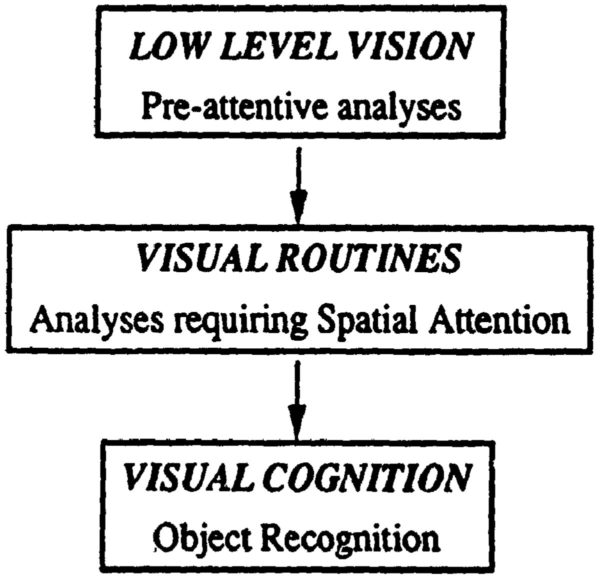

Figure 1-3: Levels of visual analysis

researchers such as Treisman (1988). These representations serve as input for intermediate level analysis, or Visual Routines. Visual routines compute spatial relations, necessary for structural descriptions of objects. It is at this stage that spatial attention operates, according to *Early Selectionists*; specifically, this is the stage at which attention serves to choose particular areas in the visual array for processing (Kahnemann and Treisman, 1984). The final stage of visual processing is highly influenced by goals and beliefs, and is thus called Visual Cognition. It is at this stage that object recognition occurs; structural descriptions derived from low and intermediate level analysis are matched to memory representations for particular objects or classes of objects. Because my interest is primarily in low and intermediate level analysis, I will talk in a little more detail about these stages.

Low level vision

To start, there is a need to discover elements in the image that correspond to boundaries or edges in the world. Boundaries are important both for purposes of recognition and navigation, and need to be made explicit early. Boundaries may be defined in a number of ways, however. For example, a boundary may be defined by a difference in the intensity, colour, texture, or motion of two adjacent areas in an image, as well by the differences in binocular disparity indicative of disparities in depth, as shown by Julesz (1971, cited in Rock, 1983) with random dot stereograms. In this brief discussion, I will be focusing mainly on intensity, colour, orientation and texture.

Low level vision takes from the rods and cones information about the light projected onto every point on the retina. Thus, the proximal stimulus may be thought of as a two dimensional array of values corresponding to the intensity of light projected at every point. From this continuous information a discrete symbolic description of the location, size and orientation of edge points is derived, according to Marr (1982) . The proximal properties that most often correspond to edges are discontinuities in light intensity. For example, in an image a dark area next to a bright one might indicate an edge. The discontinuites that indicate edges occur at many scales, however. Some contours are large and thick while others are fine and thin. It is desirable that the information at a number of resolutions be preserved, because the edges derived at one level of resolution are often independent from those derived at another. The problem then becomes that of registering intensity discontinuities at different resolutions, and determining which ones are produced by boundaries and which are produced by layout, viewpoint, shadows and glare.

Low level analysis is carried out by a network of local parallel processors, according to Marr (1982). Each processor is responsible for a particular type of discontinuity and location on the retina. If this type of discontinuity exists in the area that a given unit is depending on the task and type of cuing used (Ericksen and Schultz, 1977 and Posner, Nissen, and Ogden, 1978, respectively).

Treisman, however, suggested what was going on inside the spotlight (Treisman and Gelade, 1980). She demonstrated that although individual properties such as colour and line orientation are derived preattentively, *combinations* are not; attention was thus required for feature integration, or integrating the different features into an object description. The evidence for her hypotheses came from experiments on search and texture segregation as well as studies of report performance in divided attention tasks (Treisman,1988). Her search studies demonstrated that although a T would *pop out* in a field of X's, (presumably because of the vertical and horizontal lines), and a green letter would *pop out* in a field of brown letters, (presumably because of the colour), a green T would not *pop out* in a field of brown T's and green X's (Treisman and Gelade,1980). In fact, the time to indicate whether there was such a *conjunction* of features varied with the number of items in the display and moreover, the slope for trials in which the target was not present was twice the slope for trials in which the target was present, suggesting serial self terminating search. Similarly, combinations of features don't produce effortless texture segregation; an area composed of blue horizontal lines would not stand out as a separate area in a field of vertical blue and horizontal green lines. Finally, if subjects are asked to list the shape and colour of a number of items and the focus of attention is not at a particular location, subjects may miscombine features; subjects make *conjunction errors* when reporting the characteristics of items at unattended locations. Conjunction errors are errors that result from a miscombination of features in the display. Thus, when presented with a display in which there are red squares, green circles and yellow triangles, subjects may report red circles if their attention is focused on another location by a preceeding peripheral cue (Treisman, 1985), or another task (Treisman and Schmidt,1982). If cues direct attention to the item sought, the probability of conjunction errors drops. Treisman concluded that attention is required in order to ensure correct importance in this type of research. In it, subjects are required to search for a particular target item in a field of distractor items. For example, they may be required to search for a horizontal line in a field of vertical lines. Subjects are required to hit one key if the target is in the display and another if it is not. If the time to make the decision is independent of the number of items in the display, it is assumed that the property that differentiates the target from the distractors must be derived in parallel, i.e., preattentively. These items are said to *pop out* in search (Treisman and Gelade, 1980). If, however, the time to make the decision depends on the number of items in the display, then it is assumed that the property is not derived in parallel. Typically, in these cases, the slope for the trials in which the target is not in the display is twice that of the trials in which the target is in the display. This result is interpreted as showing that attentive processing is serial and self terminating. As soon as the target is located, processing stops: when the target is in the display this occurs on average half way through the items. From this sort of research a number of candidate features has been derived, many more than Marr originally considered. Other candidates include colour (Treisman, Sykes and Gelade, 1977), intersection (Treisman and Souther, 1985), curvature (Treisman and Paterson, 1984), and luminance change (Jonides and Yanstis, 1988; as opposed to colour change, Burkell, 1986).

The second stage of analysis involves organization of the tokens into groupings. Place tokens are grouped on the basis of principles not unlike the Gestalt principles of proximity, similarity, good continuation and common fate, according to Marr (1982). Thus, tokens corresponding to discontinuites that are close together in the image, and similar in contrast, orientation, size, and motion, etc. may be grouped to form a unit. This grouping process is hierarchical and recursive. For example, see Figure 1-4. The place tokens associated with a number of tiny vertical bars may be grouped and assigned one token, in this case a larger horizontal bar. Similarly, tokens associated with a number of tiny horizontal bars may be assigned one token, a larger vertical bar. Then tokens for

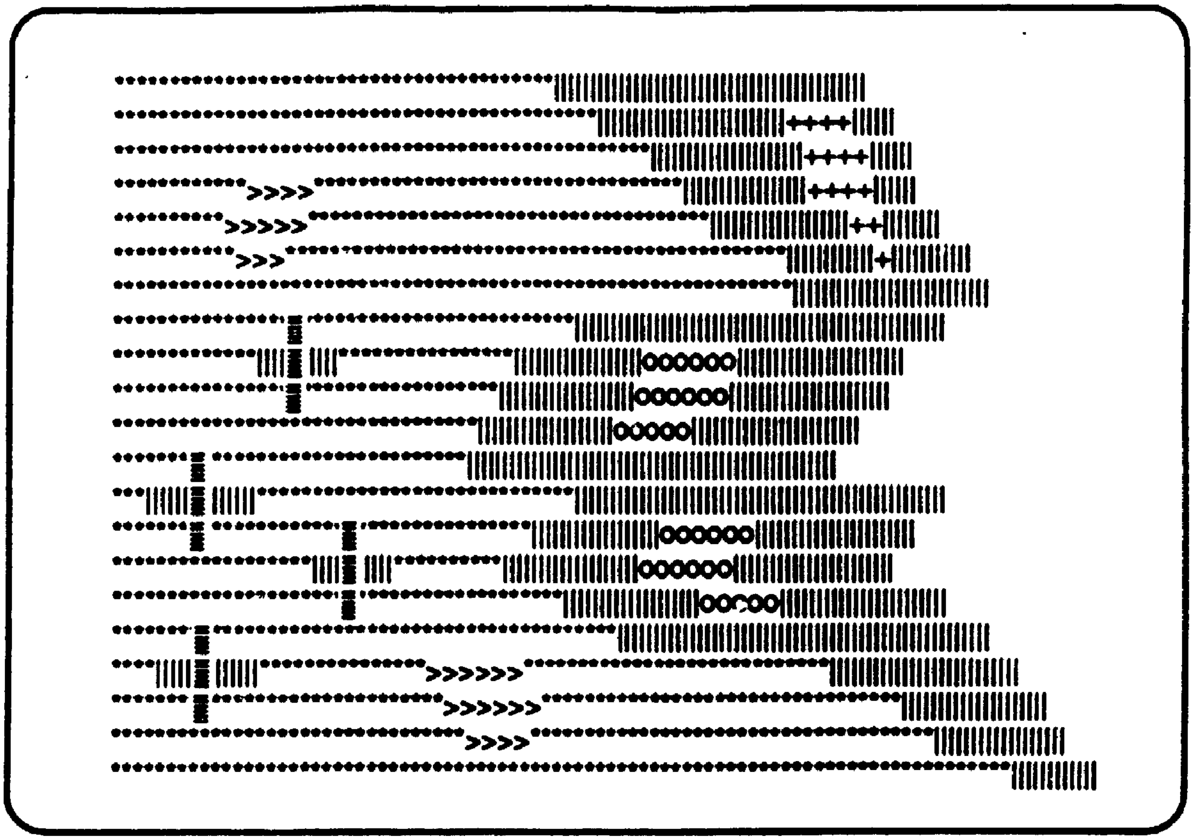

Figure 1-4: Texture diagram

the two intersecting bars may be grouped, by proximity, to form a blob. Because of the tendency to group adjacent similar edges, texture segregation occurs. Consequently, subjects are inclined to see an implicit boundary surrounding items with different features than the rest of the display. Thus, subjects can easily indicate a quadrant that is different from the others in the display; if one quadrant of the display had vertical lines and the other horizontal it would be easy for subjects to locate the dissimilar quadrant because it appears to be surrounded by a subjective contour. In fact, this is another technique used for discovering properties that are derived preattentively (Beck,1982;Julesz,1984).

There is little consensus on the algorithms for edge detection or grouping, and even less on the varieties of place tokens, or the information associated with them (i.e., features). In fact, the entire notion of features is currently under attack in the attention literature (cf., Treisman, 1988). Nonetheless, the important message to be taken from this exercise is that the initial analysis involves both edge detection and grouping, and the computations are thought to be carried out by a network of local parallel units. Because the processing is accomplished by units each responsible for only a tiny area in the retina, the output representation is necessarily retinotopic and pointillistic. Each unit only "knows" about the light projected upon a small area of the retina. Consequently, properties that are defined over the whole display cannot be described at this stage. For example, a unit responsible for one corner of the display could not "know" that the edge it registers is on the same object as another edge picked up by the unit responsible for the opposite corner.

Intermediate Level Analysis: Visual Routines

According to Ullman (1984), the purpose of intermediate level or attentive analysis is the computation of abstract spatial relations, properties defined over objects, not retinal locations. There are many abstract spatial relations. See Figure 1-5. Examples include the topological properties, such as inside, closure and connectedness, and number as well. Spatial relations are important for object description and thus recognition. For example, even babies under 16 weeks can distinguish scrambled schematic faces from ones in which the spatial relations are preserved (Fantz, 1961, cited in Ullman, 1984). Spatial relations among parts are integral to object description; they must be represented so that the stimuli will be seen "as" a face. Similarly, spatial relations are also important for visual motor coordination; it is important that we get the coffee *inside* the cup rather than *beside* the cup, for example. In general, spatial relations are important in the creation a geo-stable representation of the world, one that will not change with our viewpoint. This is because many spatial relations are viewpoint invariant. For example, two stimuli will still appear connected to even if viewed obliquely. Thus, it is important that properties defined over objects, not retinal locations, be made explicit.

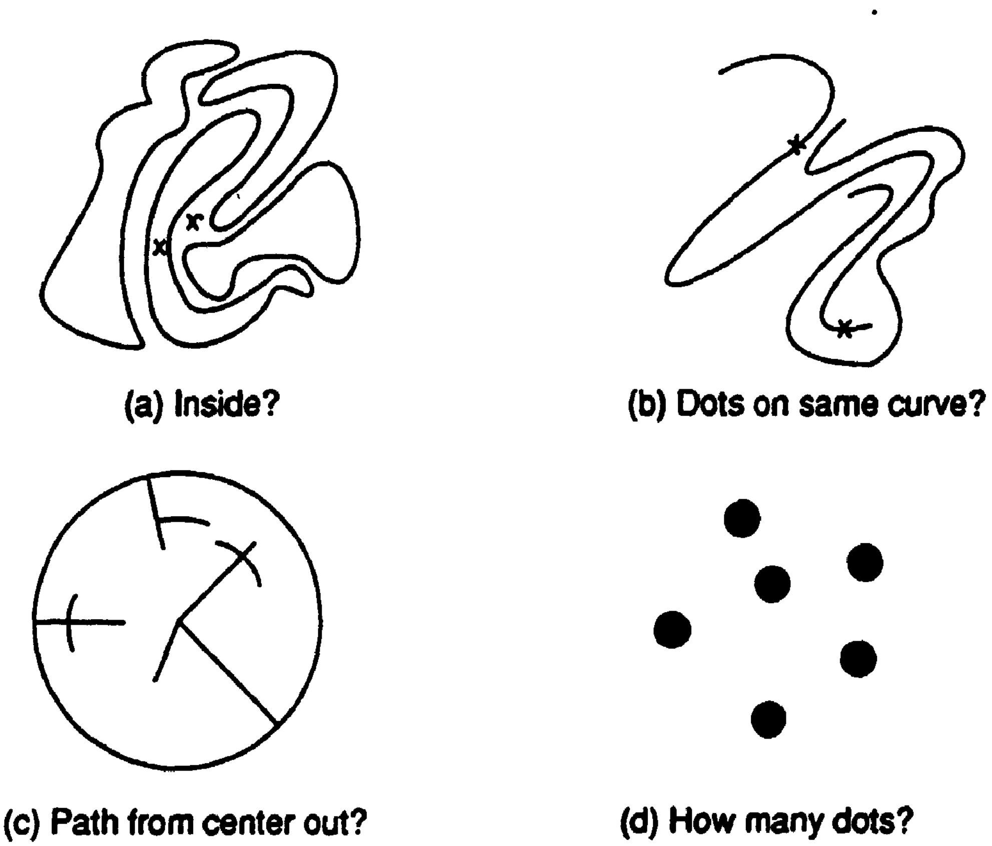

Figure 1-5: Spatial relations that might require a visual routine

People can very quickly derive these abstract spatial relations, often within one half second (e.g.,Wright,1989). Nonetheless, these relations cannot easily be computed by local parallel networks of units, according to Ullman (1984). The reason is that for these relations there is no one configuration to look for, no one location to check. For example, given the innumerable sizes and shapes of bounding contours, and the innumerable sizes and shapes of internal items, there is an unbounded number of ways that one thing can be inside another. Similarly, given the many sizes, shapes and configurations of items, there is an unbounded number of ways that there could be 42 items in a display. Consequently, the simple template matching schemes used in low level vision would not work.[2] On the other hand, a spatially serial mechanism, one in which a processing focus is moved through the visual array, might be easily able to derive these relations, as Ullman demonstrates for the connected and inside relations. The use of serial mechanisms explains why reaction times for these relations, varies with the complexity of the input, contour length and density in the case of connected (Jolicoeur, 1988), area inside or outside the contour for inside (Wright, 1989), and number of items in the case of enumeration (Klahr and Wallace, 1976).

Intermediate analysis is largely goal driven, according to Ullman. Although some computations may be carried out by default (e.g., global shape), the output of intermediate level analysis depends largely on the intention of the viewer. For example, when in a forest you don't automatically count trees. You only count when you want to count. The reason that intermediate analysis is goal driven is that there are too many spatial relations, an infinite number once combinations of relations are considered. Furthermore, if spatial relations were computed between every edge segment and every other edge segment in the visual array, a lot of unnecessary processing would be done.

Ullman proposed that abstract spatial relations are computed by what he calls Visual Routines. Visual routines are programs made up of a few elementary operations, that basically work by moving a processing focus through the visual array. Although several operations may be performed simultaneously at the processing locus, the locus can only be at one place at a time.

This processing focus can be likened to what has been called the spotlight of

[2] It is possible, in principle, to build a parallel network to compute the number of dots in a display, however (Minsky and Papert, 1969). In the general case, when items may overlap, and have different sizes, shapes, colours and brightness, this sort of mechanism will not work, however. Moreover, it seems implausible that we have such a number network, with units corresponding to every number we may someday have to count to.

attention (Posner 1978; Ericksen and Hoffman 1972, 1973; Jonides 1980; Laberge 1983). The idea behind the spotlight metaphor is that spatial expectations act like a beam of light, enhancing perceptual performance in the restricted area where the subjects expect the stimulus to fall. The spotlight of attention research rests on spatial cuing studies. In cuing studies subjects are required to make some sort of perceptual decision, for example, press one key if there is an "M" in the display and another if there is an "N". The stimuli in question are preceded, however, by either a neutral cue that simply warns *when* the stimuli will occur, or a spatial cue that informs the subjects both *when* and *where* the stimuli will appear. There are two types of cuing. A central arrow indicates which side of the display the stimuli will fall in *central cuing* studies (e.g.,Posner, Nissen and Ogden, 1978). In *peripheral cuing* the arrow appears adjacent to where the stimuli will appear (e.g., Tsal, 1983). Regardless, the speed and accuracy of perceptual decisions improve if a spatial cue precedes the stimuli, even if eye movements are prevented. The standard interpretation of this finding has been that the processing focus has been moved through the image as a result of the subject's expectations about where the letter will fall. The processing focus performs perceptual analysis, thus the improvement is due to better perceptual processing rather than changes in decision criterion (Bashinski and Bacharach, 1980; Posner and Snyder and Davidson, 1980).[3] The spotlight cannot be split (Posner, Snyder and Davidson, 1980; Ericksen and Yeh, 1985), though the diameter of the spotlight can be changed (Ericksen and St. James, 1986;Laberge and Brown, 1986). Moreover, by manipulating the amount of time between the presentation of the cue and the presentation of the target, the speed at which this spotlight, or processing focus can been moved has be estimated (Posner, Nissen and Ogden, 1978; Ericksen and Schultz, 1977; Tsal, 1983). Estimates vary between 33.3 msec/degree and 4 msec/degree

---

[3] This conclusion is less controversial in cases where peripheral cuing is used, and in which the perceptual decision is complex, e.g., letter discrimination rather than simple luminance detection. (Shaw, 1984; Briand and Klein, 1987).

depending on the task and type of cuing used (Ericksen and Schultz, 1977 and Posner, Nissen, and Ogden, 1978, respectively).

Treisman, however, suggested what was going on inside the spotlight (Treisman and Gelade, 1980). She demonstrated that although individual properties such as colour and line orientation are derived preattentively, *combinations* are not; attention was thus required for feature integration, or integrating the different features into an object description. The evidence for her hypotheses came from experiments on search and texture segregation as well as studies of report performance in divided attention tasks (Treisman,1988). Her search studies demonstrated that although a T would *pop out* in a field of X's, (presumably because of the vertical and horizontal lines), and a green letter would *pop out* in a field of brown letters, (presumably because of the colour), a green T would not *pop out* in a field of brown T's and green X's (Treisman and Gelade,1980). In fact, the time to indicate whether there was such a *conjunction* of features varied with the number of items in the display and moreover, the slope for trials in which the target was not present was twice the slope for trials in which the target was present, suggesting serial self terminating search. Similarly, combinations of features don't produce effortless texture segregation; an area composed of blue horizontal lines would not stand out as a separate area in a field of vertical blue and horizontal green lines. Finally, if subjects are asked to list the shape and colour of a number of items and the focus of attention is not at a particular location, subjects may miscombine features; subjects make *conjunction errors* when reporting the characteristics of items at unattended locations. Conjunction errors are errors that result from a miscombination of features in the display. Thus, when presented with a display in which there are red squares, green circles and yellow triangles, subjects may report red circles if their attention is focused on another location by a preceeding peripheral cue (Treisman, 1985), or another task (Treisman and Schmidt,1982). If cues direct attention to the item sought, the probability of conjunction errors drops. Treisman concluded that attention is required in order to ensure correct integration of features into an object description. In contrast, the probability of *feature errors*, errors othat occur when subjects add or drop features, is is insensitive to the position of the attentional focus. For example, given a green triangle, red circle and yellow square, the probability that a subject would report a blue triangle should be the same regardless of whether they had focused attention on the triangle or not; attention is not required to determine the colour of items in a display.

If the spotlight of attention or locus of feature integration corresponds to the processing focus in Ullman's Visual Routines, then presumably spatial relations that require a routine, such as connected and inside, should not *pop out* in search because they cannot be derived without attention. This prediction has been borne out in the search and texture segregation research, for the most part. For example, Treisman and Gormican (1988) showed that properties of connectedness and enclosure did not produce "pop out" in search tasks. Similarly, when the items in one quadrant differed from the others due to the spatial relations between the parts of the items, texture segregation did not occur (Beck, 1982;Julesz,1984).

According to Ullman (1984), however, feature integration is not the only operation carried out using the processing focus. In particular, he suggested five elementary operations that might be performed with the focus in order to compute spatial relations.

1. Move the processing focus from one location in the representation to another. Presumbably he is reserving this type of operation for situations such as central cuing studies where there is no visible feature cluster for the processor to move towards. In future this operation will be referred to as *scanning*.

2. Move the processing focus towards a feature cluster, an area in the representation that differs from the background in terms of a low level feature such as colour, depth, orientation, etc. Ullman calls this operation *indexing*. This operation would be important in peripheral cuing and search tasks. Indexing will be of particular relevance in later sections.

3. Move the processing focus along a contour. This operation has been called *boundary tracing* and is thought to be used for computing the connected relation (Jolicoeur, Ullman and Mackay, 1986).

4. Activate an area surrounding the processing focus, stopping at the point of a contour. This operation is of particular importance to the inside relation. Ullman calls the operation *colouring*.

5. Mark areas on the representation that have already been processed. This facility is necessary to prevent infinite loops in enumeration and boundary tracing. This operation is called *marking*.

These elementary operations are composed into simple iterative programs, or Visual Routines. Visual routines can be assembled as needed so that the system can be flexible. If the same relation is computed repeatedly, the routine may become compiled, for greater efficiency, however.

Subitizing, Counting and Visual Routines

It makes sense that number may be computed by a serial goal driven mechanism such as a Visual Routine, because enumeration is not performed automatically, and the process takes place in a time that is dependent on the number of items in the display. In fact, the visual routine for number can easily be created by augmenting Ullman's operations, indexing and marking, with the memory functions required for addition. This counting algorithm corresponds to the common sense procedure taught on Sesame Street. At the beginning, a short term memory counter is set to zero. Then the cycle of indexing to an item, augmenting the counter by one, and marking the item as counted is performed until all the items are marked. The final counter value corresponds to the how many items there are in the display.

This does not seem to be the way people compute number, however. If it were then reaction time should increase monotonically as a function of the number of items in the display; the latency difference between 9 and 10 should be the same as that between 1 and 2. The research on subitizing and counting shows that this is clearly not the case; the slope in the 1-4 range is only 40-100 msec/item whereas the slope for 5 or more is 250-350 msec/item typically. There is no reason, given the discussion so far, and even given the research on mental arithmetic (e.g., Parkman and Groen, 1971) that there should be a sudden discontinuity in slope after 4. How, then, can the subitizing and counting research be reconciled with Ullman's Visual Routines? To answer this question, it 's necessary to consider in detail what it means to move the processing focus.

The problem with Visual Routines

Assume Ullman is correct in his analysis of the problem so far; there are a series of elementary operations for moving a spatial processing focus through the visual array. Thus, there might be an operation, INDEX that would take as its argument information that w : give access to a particular item or feature cluster in the visual array. See Figure 1-6. Say that the attentional focus is currently at the point denoted as *a* in the

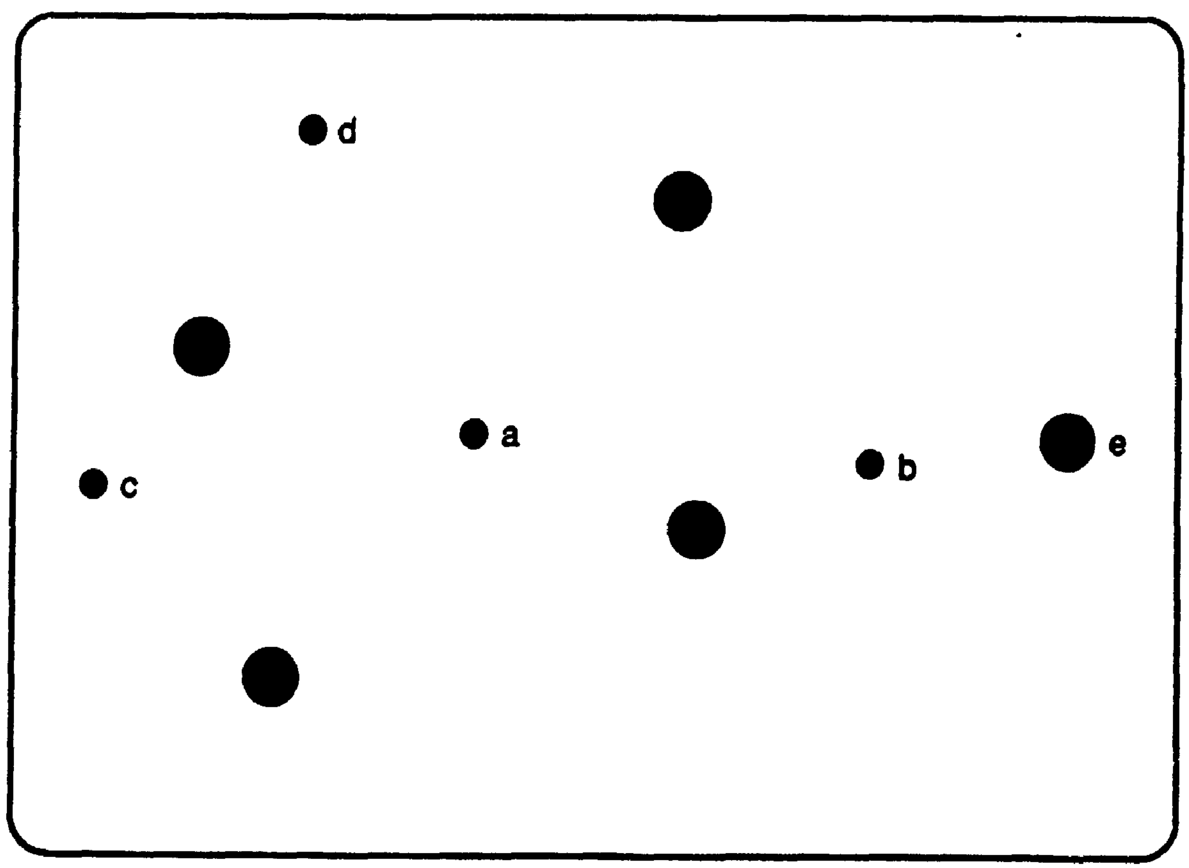

Figure 1-6: Moving the attentional focus

figure. The task is to move the attentional focus to the point denoted *b*. How could this be accomplished? One possibility is that the processor be sent to a location defined by the retinal features of the item, e.g., INDEX (small black dot). This strategy would not work, however, if there were more than one token of the same type., i.e., if there were more than one small black dot. In this example the processor might be as likely to land at the points denoted as *c* or *d*, if a property based method of addressing was used. Moreover, in the world, the retinal properties of an item might change from moment to moment as a result of changes in lighting or projection, or changes in the item itself. A dot might of its own change colour, brightness or size, for example. Thus, the strategy of using properties as arguments in the INDEX operation is doomed.

Another possibility is that a coordinate grid be overlaid, and the INDEX operation take as its argument a retinal coordinate, e.g., INDEX (25,35). The problem with this strategy is that the retinal coordinates of the item change with eye movement. For example, item *b* might fall at the same retinal location as *e* if the person were to move their eyes to the left. Moreover, in the world, items move of their own; things shift, fly, roll and bounce. Thus, the position of item *b* might change even if eye position is constant. To index on the basis of retinal coordinates is to risk sending the processing focus to a place that no longer houses an item, or houses a different item than the one intended.

What is needed is some way of referring to a point, individuating it, tagging it, in the same way that I did when I labelled the points *a*, *b* and *c*. What is more, it is important that these tags stay with their respective items though the properties and position of the items change. In fact, these tags would function in much the same way as pointer variables do in Pascal or C. For example, in C it is possible to define a variable *&b* that has the memory position of variable b. Though the value or memory position of *b* may change, the variable *&b* remains assigned to item b, and the value of *&b* can be appealed to in order to "find" *b*, and thus recover information about b.

Thus, in summary, between low and intermediate stages of visual analysis, there is need for a stage in which certain feature clusters are individuated, or named in such a way that at any time their retinal position can be accessed. It is at this point that Pylyshyn's (1989) FINST model comes in. See Figure 1-7. FINSTs, for FINgers of

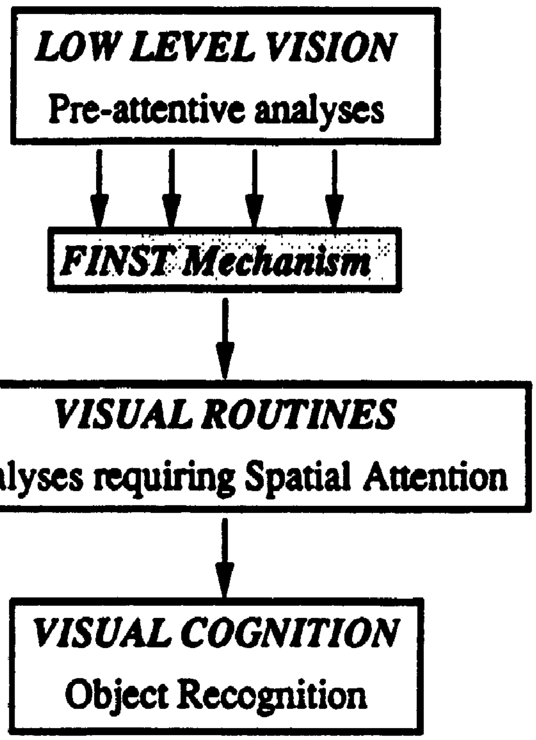

Figure 1-7: Visual processing and the FINST mechanism

INSTantiaion, are spatial reference tokens, or pointer variables, that permit access to certain places in the representation. They provide a way of referring to a feature cluster without specifying properties or coordinates, which may change from moment to moment. Without the ability to individuate feature clusters, *indexing* would not be possible. When faced with the goal of moving the attentional focus, or for that matter, the eye or finger towards a feature cluster it is necessary to know *which* cluster to move toward.

The FINST hypothesis

FINSTs are internal reference tokens that are assigned to a small number of feature clusters after low level processing--edge detection and grouping (Pylyshyn, 1989). Each FINST remains assigned to its respective cluster even if the retinal position of the cluster changes as a result of eye movements or movements by the object. Only FINSTed areas can be further accessed by attentive processes employed by visual routines, or by motor commands, permitting eye or finger movements towards the FINSTed items. There are only a small number of FINSTs. Previous experimentation suggests five at least, however (Pylyshyn and Storm, 1988). Thus the FINST system is a limited capacity parallel mechanism for indexing feature clusters.

FINSTs do not encode the features they point to, they just make it possible to examine the properties of the clusters if needed. Further, there is no way to find out whether two FINSTs refer to features at the same location except by examining the location. Thus, in Treisman's task attention is required in order to find out if an object is both green and square. The FINST mechanism simply performs variable binding; FINSTs are simply names, symbols that are bound to feature clusters. These names permit attentive operations and motor commands to refer to the cluster and access its location in the visual array.

So far only bottom up assignment of FINSTs has been discussed. Pylyshyn (1989) leaves open the possibility that FINSTs may also be assigned in a top down manner, in response to goals and expectations. For example, it may be possible to FINST only certain features on the basis of an intention. Thus, subjects may be able to FINST all the red items and ignore the green. In fact, this prediction was borne out by Egeth, Virzi and Garbart (1984). Their results suggest that subjects simply scan a relevant subset when searching for conjunctions of colour and shape. Thus, if subjects were searching for red O's in red N's and green O's, they only search the red items, for example (cf., Wolfe, Cave, and Franzel, 1987). Similarly, it may be possible to FINST a given item among a set of identical items as the result of an intention, e.g., focus on a particular fly speck on the wall. What seems unlikely is that FINSTs can be assigned in entirely featureless areas such as Ganzfelds, where there are no illumination or colour discontinuities. Subjects even lose a sense of where their eyes are focused when looking into a Ganzfeld.

To date, there are several experiments that support the FINST hypothesis. The first is a multiple target tracking experiment by Pylyshyn and Storm (1988). In it subjects were faced with a number of identical objects. For a brief time a subset of these objects flashed. Subjects were required to treat this subset as target items and the rest as distractors, although the items became identical after the intial moments. Then all of the items were set into random independent motion. After a period of time one of the items changed shape. The subjects' task was to decide if the object that changed shape was a target or distractor. The number of targets and distractors was varied, along with tracking time and the rate of motion. Trials in which there were eye movements were discarded from analysis. It was found that subjects could track up to five objects in a background of distractors even if the objects were moving so fast and so erratically that attentive scanning from target to target was impossible. It was concluded that up to five objects can be tracked in parallel. FINSTs, initially assigned when the items flashed at the start of the experiment, remained "glued" to their respective items although the positions of the items changed. The maximum number of targets that can be tracked is not known; although the accuracy deteriorated with the number of targets, subjects were still responding with 86% accuracy when there were 5 targets and 5 distractors.

The second study is of particular relevance to counting and thus will be discussed in detail. Sagi and Julesz (1984) presented subjects with circular arrays of diagonal line segments containing a small number of horizontal and vertical lines. (These horizontal and vertical lines "pop out" of the background diagonals). The stimuli remained for 5 msec and then were replaced after a variable amount of time by a pattern mask. There were two tasks. In the number discrimination task subjects were simply required to say how many pop out items (i.e., horizontal or vertical lines) there were. There were four possible discriminations: 1 vs 2, 2 vs 3, 3 vs 4 and 4 vs 5. Discriminations were blocked. The stimulus onset asynchrony required for 95% accuracy in discrimination was measured. Sagi and Julesz found that for number discrimination the required stimulus onset asynchrony did not vary substantially with number. (The slope was only 1.9 msec per item). In contrast, when the task was to indicate whether all of the "pop out" items had the same orientation, the required stimulus onset asynchrony varied with the number of objects; the slope was 16.6 msec per item. Thus, it took 16.6 msec longer to determine that four horizontal lines had the same orientation than three, for example.

Sagi and Julesz interpreted their findings as evidence that our knowledge of where stimuli are precedes our knowledge of what they are. "Where" knowledge is all that is required for number discrimination; subjects need only know of a number of different locations that are unlike the background. (It is not enough simply to know of discontinuities, it is necessary to know of *separate* discontinuities). "What" knowledge is required in order to determine if all discontinuous items had the same orientation, however. Given that only up to five pop out targets were used, the Sagi and Julesz' results are entirely consistent with the FINST hypothesis. A small number of feature clusters can be assigned tokens in parallel, but finding out *what* the tokens refer to requires serial attentive processing. Of course, the Sagi and Julesz study does not go far enough for purposes of the FINST hypothesis. It is also necessary to show that number does have an effect on processing time once the capacity of the parallel mechanism is exceeded. Once the number of "pop out" items exceeds the number of FINSTs, seriality should once more be evident in number discrimination. Lorinstein and Haber (1975) showed this to be true in another masking experiment where the task was simple dot enumeration rather than number discrimination. Up to 6 dots could be counted at 4 msec per dot, but thereafter each additional dot required 60 msec.

Subitizing and counting according to the FINST hypothesis

At this point it becomes possible to forge the link between subitizing and the FINST hypothesis. Why can we only subitize four or five items? The reason is that the system that individuates feature clusters by binding them to reference tokens has limited resources; there are only a small number of reference tokens (FINSTs). Consequently, the system is parallel but nonetheless limited capacity. Once the capacity is exceeded a different process must be used.

How then are subitizing and counting performed? It is only possible to speculate at this point. A tentative proposal is possible, however, by drawing on vision research and the empirical and theoretical work on counting. To start, consider subitizing, a case in which the number of items in the display is less than the number of internal reference tokens or FINSTs. Subitizing can be seen to involve two stages. The first stage involves the assignment of reference tokens; one FINST is assigned to every item in the display. In the primal or 2 1/2 D sketch items are indicated by place tokens assigned to intensity discontinuities, or groupings of intensity discontinuities. FINSTing involves tagging a subset of these place tokens. It is further assumed that FINSTs can be assigned to place tokens associated with certain properties, as long as the property emerges preattentively. As long as target locations differ from the others on the basis of a property computed by low level vision, a feature such as colour or curvature, this sort of "top down" assignment of FINSTs is possible.

This first stage of the subitizing process can be thought of as prenumeric because at this stage you are only conscious of "some" items in the display; the number name has not yet been accessed. This sort of information must be available to the system before the attentional processor is moved to an item to check its identity. Otherwise the spatial processor would not "know" when to start indexing or "know" when to stop. In fact, this much is assumed in Jolicoeur's (1988) experiments on boundary tracing. In these studies subjects were required to indicate whether X's were connected by a line. Before this process begins the subjects must know that there are TWO different X's, however. Presumably boundary tracing would not occur if there were only one X in the display. Similarly, in conjunction search, the time to determine whether a particular combination of features exists in the display depends on the number of items in the display. This result is interpreted as showing that the attentional processor is moved from item to item, combining features. If rudimentary information about separate items were not available, the relationship between display size and reaction time would fall apart; subjects might stop short, processing only one or two of the items before responding, or might continue on, looking for items long after all the items in the display had been checked. If this occured, reaction times would not so neatly parallel the number of items in the display (even if only a relevant subset of the items are being checked). Moreover, subjects would miss targets more often, because of failure to check all the locations; subjects rarely err in search tasks though they may be unaware of the number of items in the display (Treisman and Gelade, 1980). A similar sort of "prenumeric" information is necessary when subjects are required to decide whether a display is subitizable, without actually counting items in the display (Mandler and Shebo, 1982; Atkinson, Campbell, and Francis, 1976). The time required to make such a decision is more or less constant with number of items in the display.[4] This sort of information is prenumeric because it is available before the number name is accessed. It is number information in the sense that it can be used to distinguish zero from some, one from more than one, or a subitizable number from a more-than-subitizable number of items.

The second stage of the subitizing process involves using information about the tagged place tokens to gain access to a number name stored in long term memory. The

[4]Reaction times to determine whether 5 items were subitizable were somewhat higher than the others, in both studies. This may be because 5 is sometimes counted and sometimes subitized. It would consequently be hard for subjects to decide such a borderline case.

process of finding the cardinality of a particular display cannot be accomplished by anything like a template match between a visual representation and long term memory representation, however. Not only is the concept of "4" abstract, as discussed in the section on visual routines, but the concept of an item is abstract: Items come in an infinite number of sizes, shapes and colours, and may even be defined by illusory contours, disparities in stereoscopic depth (as in a random dot stereogram), or movement. In fact, the processes involved in number recognition must be quite different from the processes involved in object recognition, because with number there is the problem of embedding: In every 4 item display are 3 item displays, 2 item displays and 1 item displays. In object recognition, this would be equivalent to a case where within every goat, there was a turtle, a fish and a cat. The only thing that differentiates every 4 item display from every 3 item display is that 4 item displays have 1 item, 2 item, 3 item, AND 4 items implicit. Consequently, the process of finding a number name must involve pairing each item in the display with a number name in order of the number names. (Klahr (1973b) makes a similar assumption in his production system for subitizing).

Where then does the subitizing slope come from? Given this rudimentary theory, there are two possibilities: the variable binding stage, or the response choice stage. In the course of this paper two types of enumeration study have been discussed, ones in which the dependent variable is reaction time, mentioned at the beginning of this paper, and ones in which the masking paradigm is used, and the dependent variable is stimulus onset asynchrony to obtain a certain level of accuracy, as in the Sagi and Julesz (1984) study. In addition, two types of enumeration task have been discussed, ones in which there is a wide variety of responses, mentioned at the beginning of the paper, and others in which the number of response alternatives is limited to two, as in the Sagi and Julesz study. By comparing subitizing slopes from studies that differ on these dimensions, dependent variable and number of response alternatives, it is possible to glean information about the origin of the subitizing slope.

Table 1-1 contains a listing of subitizing and counting slopes from a variety of different studies. (Unless otherwise specified the stimuli to be enumerated were dots). (1) Notice that the subitizing slope is primarily evident in reaction time studies and moreover in reaction time studies with a large range of possible responses. When masking methodology is used the subitizing slope all but disappears. For example the slope is a mere 4 msec/item in the enumeration studies of Lorenstein and Haber (1975), 4 -10 msec/item in Oyama, Kikuchi, and Ichihara (1981), as compared to 40 -72 msec/item in enumeration studies where reaction time is measured. Second, the subitizing slope seems to be slightly more pronouced when there is a wide range of responses as in typical enumeration studies, than when there are only two responses. For example, in number discrimination and number matching experiments the slope is typically low, 33 msec/item or 31 msec/item in the 1-4 range in Folk, Egeth and Kwak (1988) and Simons and Langheinrich (1982). For comparable enumeration studies (without eye movement control) in which there is a wide range of possible responses the slope in the 1-4 range is 60 msec/item on average (Klahr, 1973a; Aoki,1975). If a masking paradigm is used AND the number of response alternatives is limited, as in Sagi and Julesz (1984), the slope all but disappears (1.9 msec/item). Consequently, it seems that the subitizing slope arises primarily from the need to choose a response from a range of responses. If the response range is limited, or if the response choice part of the latency is excluded by using a masking paradigm, the subitizing slope decreases. Thus it would appear[5] that the second stage of the counting process is responsible for most of the slope.

The counting process begins when it becomes apparent that ALL the tokens are assigned, i.e., once we learn that we can't subitize (cf., Mandler and Shebo, 1982;

---

[5]Things are not that simple, however. The subitizing slope is also affected by stimulus degradation, as is apparent from the 3 msec duration and masking conditions of Liss and Reeves (1983). This issue will be discussed in greater detail later. In addition eye movements tend to inflate the subitizing and counting slopes, and consequently, reaction time studies that control eye movements, either by dropping trials with eye movements (e.g., Klahr, 1973a), or limiting the exposure duration of the counting display (e.g., Oyama, Kikuchi,Ichihara,1981; Liss and Reeves, 1983) have lower subitizing slopes.

Table 1-1

Subitizing and counting slopes from enumeration studies

ENUMERATION STUDIES: RESPONSE IS NUMBER IN A RANGE

Authors	Range	Slope in msec	DV	Comments
Akin	1 - 3	78	RT	3D blocks-
Chase	4 - 15	354	-	11/14 subjects-
(1978)				
	1 - 4	153	-	3/14 subjects-
	5 - 15	393	-	-
Aoki	1 - 4	42	RT	-
(1975*)	5 - 15	297	-	-
Klahr	1 - 4	66	RT	-
(1973)	6 - 10	268	-	-
	1 - 4	72	RT	-
	6 -10	268	-	-
	1 - 4	25	RT	+
	1 - 4	60	-	-
Liss	1 - 3	35	RT	Duration 20 msec+
Reeves	" "	103	-	Duration 3 msec+
(1983)	" "	163	-	Mask after 50 msec+
	4 - 10	338	-	Duration 20 msec+
	" "	315	-	Duration 3 msec+
Oyama				
Kikuchi	1 - 4	40	RT	+
Ichihara	5 - 15	370	-	+
(1981)				
Lorenstein	1 - 6	4	SOA 50%	+
Haber	7 - 16	60	-	+
(1975)				
Oyama	1 - 4	4 - 10	SOA 50%	+
Kikuchi				
Ichihara				
(1981)				

Table 1-1 (continued)

Subitizing and counting slopes from enumeration studies

NUMBER DISCRIMINATION STUDIES: DISTINGUISHING N from N+1

Authors	Range	Slope	DV	Comments
Folk	1 - 4	33	RT	No distractors-
Egeth	" "	38	-	Distractors-
Kwak				
(1988)				
Sagi	1 - 4	1.9	SOA 95%	+
Julesz				
(1984)				

NUMBER MATCHING: SAME NUMBER vs DIFFERENT NUMBER

Authors	Range	Slope	DV	Comments
Simons	2 - 4	31	RT	-
Langheinrich	4 - 6	97	-	-
(1982)				

+ indicates eye movement control.
* cited in Aoki(1977). The original is written in Japanese.

Atkinson, Campbell, and Francis, 1976). After this initial discovery attentional processing begins; the attentional focus is moved through the representation area by area, as suggested by Ullman. Typically, some sort of scanning strategy is formed. For example, children are taught to work from left to right, top to bottom in order to minimize the probability of "getting lost" in the display. Next, a counting routine is performed; an index, mark and add cycle is executed. As children we progressed item by item, probably because we learned to count by ones before we learned to add. Most adults count by three's or four's, however. Thus, the counting process is just as discussed earlier; counting involves grouping the items into clusters of three or four, subitizing the group, adding the result into a running total kept in working memory, and then moving the attentional focus to the next group. This activity can be conceived as the process of FINSTing a place token at a low level of resolution (a large blob), moving the processing focus toward it, and then reassigning the FINSTS to place tokens associated with the elements within a group (the small blobs representing the black dots in the display).[6] Thus, the counting process is complex. It involves many stages: discovering the display is non-subitizable, grouping the items, forming a scanning strategy, moving the attentional processor from group to group, subitizing the number of items within a group, adding the result into a running total. As a result, a given counting latency is composed of the latencies from many processes. Some of these processes may take place in a time independent of the number of items in the display. For example, the time to decide whether or not the display is subitizable, the time to group items (if the groups that occur in counting arise due to low level grouping processes), and the time to form a

---

[6] At this point it is hard to say whether the grouping process as discussed by Marr would result in the same clusters that people seem to form in the process of counting. People seem to group items primarily on the basis of proximity, rather than other cues when counting. Moroever, they seem to group on the basis of *relative* rather than absolute proximity (Van Oeffelen and Vos, 1982). Thus, although two items might be rather far away from each other, if they are relatively closer to each other than they are to three other items, the two items will be grouped. Further, people seem to prefer to have a subitizable number within each group, so they will divide groups in halves or thirds if there are too many items in them, even though a proximity alone wouldn't dictate such splits (cf., CODE algorithm, Van Oeffelen and Vos, 1982; Shrager, Klahr, and Chase 1982).

scanning strategy may take place approximately as fast when there are large numbers of items as when there are small numbers of items. Other processes would be expected to take longer as the number of items in the display increases. Thus, the more items in a display, the more times the attentional focus would have to be moved, the more addition operations would have to be performed, the more marking would have to take place. Also, number would affect the number of items within a group; the more items in a display, the larger number of items within a cluster, and the longer to subitize the cluster. Similarly, latency to perform additions is thought to be affected by the magnitude of the addends (Parkman and Groen, 1971); if the minimum addend is small, addition takes less time than when the minimum addend is large. Consequently, the counting slope comes from a variety of different processes, some of which take longer the more items there are in the display.

Therefore, subitizing is fast because it is a very simple process, involving two stages, one of which may be close to parallel (1.9 msec/item is close). Because there are few memory requirements, subitizing is also very accurate. Thus, for example, a concurrent articulatory suppression task, designed to fill the articulatory-acoustic memory buffer, did not interfere with enumeration when there were up to 6 or 7 items in a display, but did interfere thereafter (Logie and Baddely, 1987). The counting process is slow because of the many stages involved, some of which take longer the more items there are in the display. The counting process is error prone because of memory requirements; it is possible to forget the subtotal, forget the addition table, or forget which items have been already counted. There are simply more things to go wrong when counting than subitizing.[7]

Overview of studies

The empirical substantiation of the FINST model is beyond the scope of one thesis. In fact, at this point experimentation is needed to help constrain the FINST hypothesis. My goal in this series of experiments is merely to demonstrate, using converging operations, that subitizing relies on a preattentive mechanism whereas the counting process requires spatial attention. Specifically, two conceptions of spatial attention will be explored, Ullman's conception of spatial attention as a processing focus for the computation of spatial relations and Treisman's conception of attention as the "glue" that integrates different features into object representations. In the first set of experiments I will work with Ullman's notion, and try to demonstrate that subitizing is no longer possible when spatial attention is required to compute a spatial relation such as connected, or to resolve items as wholes when preattentive information is misleading. In the second set of experiments I will be working with Treisman's idea of attention. Specifically, I will be trying to show that people can subitize items in a background of distractors, but only if the targets and distractors differ in terms of a feature, or property that induces *pop out* in search tasks.

---

[7] In addition to the relationship between number and response latency, there is a relationship between number and standard deviation of response latencies (cf., Akin and Chase, 1978). Thus, when there are small numbers of items in the display, the standard deviations for latency may vary between 30 and 100 msec for a given individual. As the number of items in the display increases, so does the standard deviation, so a given subject may have a standard deviation of 300 msec when counting 8 items. Moreover, the relationship between standard deviation and number is evident not only within a subject's data, but between subjects. Individual differences are more manifest when subjects are counting large numbers of items than small. The relationship between latency, standard deviation and error is an annoyance because it makes analyzing and interpreting data from enumeration studies difficult. Nonetheless, the relationship between these three variables is an important clue to the sort of processing involved in subitizing and counting. The counting process may be more susceptible to moment-to-moment fluctuations in concentration or memory load (Logie and Baddely, 1987). Similarly, the counting process is more vulnerable to the idiosyncracies of particular displays, and such properties as familiar item configuration (Trick, 1987), symmetry (Howe and Jung, 1987), compactness, planarity and linearity (Akin and Chase, 1978), item heterogeneity (Frick, 1987), and the arrangement of heterogeneous items (Beckwith and Restle, 1966). Consequently, it stands to reason that the standard deviations for latencies are higher for large numbers of items than small: many more factors influence counting than subitizing.

The following studies share a common strategy. I will be looking for situations in which the subitizing process can no longer function. How will I know when this happens? The trademark of the change from subitizing to counting is the increase in slope after 3 or 4 in response latency. This increase causes the "elbow" in the latency function, and is registered as a deviation from linearity by trend analysis.[8] If an experimental manipulation gets rid of this deviation, causing the slope to be constant with number, then there is evidence that the same process is being used for both small and large numbers of items. Moreover, if this slope is high, in excess of the norm for subitizing given the presentation conditions, this is evidence that the counting process is being used for both small and large numbers of items: The subitizing process has been foiled.

---

[8] Notice that it is not the slope per se that indicates whether subitizing is occuring but slope discontinuities between the 1-3 and 5-7 ranges. I am looking for evidence that small numbers of items are being processed differently than large numbers. Thus, although Chi and Klahr (1975) found that the slope in the 1-3 range was 195 msec/item for 5-7 year olds, it is assumed that subitizing is occurring because the slope was much smaller than the slope for the 4-10 range, which was 1097 msec/item. In fact, there is evidence of differential processing of small and large numbers of items even in 22 week old infants: infants can discriminate 2 from 3 but not 4 from 6 (Starkey and Cooper, 1980). With age, the subitizing range increases from 1-3 to 1-4, and the subitizing and counting slopes both decrease. The subitizing slope asymptotes when children are 10 years old whereas the counting slope continues to decrease until children are 14, however (Svenson and Sjoberg, 1983). The difficulty for young children seems to be in learning the 1:1 correspondence between items and number names; children will often count the same item several times or forget to give an item a number name, though they know the number names in order. Thus, the ability to point out each member of a set once and only once is correlated with the ablity to count in 2 1/2-4 year olds (Potter and Levy, 1968).

Chapter Two
Enumeration and Visual Routines: Counting connected items

People routinely enumerate one subset of things while ignoring another subset. For example, a person orchestrating a family reunion might want to find out how many children there are in a crowd of children and adults. Thus, we regularly enumerate certain "targets" while ignoring a number of "distractors". The difficulty of this task may be influenced by how easy it is to determine whether an object is a target or distractor, however. The goal of this experiment is to show that when attention is necessary to compute a spatial relation to determine whether a given item is a target or distractor, subitizing is not possible. In this experiment the spatial relation in question is connectedness.

The idea that computing connectedness requires spatial attention has been supported by studies on search and texture segregation. For example, a connected item (O) will not pop out of a group of unconnected items (C) in search (Treisman and Gormican,1988). Similarly, effortless texture segregation does not occur when one group of textural elements differs from the others only by having connected elements, such as  and  (Beck, 1982;Julesz, 1984). In addition, Jolicoeur has performed a series of studies that show that the latencies to make connectedness judgements vary with the complexity of the display, as would be expected if some sort of serial mechanism were operating (Jolicoeur, 1988). In these studies subjects were required to push one key if two X's were connected by a line, and push another key if the X's were on different lines. See Figure 1-5b. The position of the first of the two X's was always at fixation. The retinal distance between this first X and the second was kept constant while the contour distance, or the distance on the line connecting the points, was varied. Jolicoeur typically found that contour distance had strong effects on how long it took for people to say that two X's were connected. Subjects were much faster at responding when the X's were separated by a short contour distance than a long distance, though in each case the retinal distance between the two X's was the same.[1] This result would be expected if subjects were moving a serial attentional focus along the contour at a fixed rate; 21 msec/degree is typical in Jolicoeur's studies (1988). Moreover, his results replicated even when eye movements were prevented. When displays were shown for only 150 msec, too short a time for initiation of eye movements, the relationship between contour distance and reaction time remained the same (Jolicoeur, Ullman and Mackay, 1986). Thus, the scan rate is a function of the time to move the attentional focus rather than the retinal focus. This relationship between contour distance and reaction time has been taken as evidence that connectedness is computed by a serial spatial mechanism. Specifically, these results have been taken as evidence that subjects perform Ullman's boundary tracing operation in order to determine if two items are connected; the processing focus or attentional spotlight is moved from the starting X along a contour until the second X is encountered.

Given that attention seems to be required to determine whether two items are connected, it would seem unlikely that subjects could subitize connected items in a background of distractor items if subitizing exploits preattentive information. This is because information about which items are connected only becomes available *after* attentive analysis. The mechanism that individuates feature clusters is *preattentive* according to the FINST hypothesis. In contrast, there should be evidence of subitizing if targets are defined by their colour instead of their connectedness. Colour is often cited as an example of a property that is derived preattentively. For this reason, items that differ

---

[1] Though this methodology neatly controls for retinal distance between target X's, it confounds contour distance with the number of curves in a contour: a four unit contour connecting two points would presumably have more bends in it than a one unit contour connecting the same points. The present study also confounds these two factors, unfortunately. Jolicoeur, however, showed that it was not simply the number of direction changes that influences reaction times. He had subjects counting items on the contours defining a semi-circles of different sizes (Jolicoeur, 1988). Both contour distance and the degree of curvature (i.e., how tight the curve was) effected the boundary tracing rate in this experiment, though the number of curves, or number of direction changes in the contour, was constant.

in colour from the others in the display will "pop out" in search; subjects detect an item that is a different in colour from the others in a time independent of the number of items in the display (Treisman and Gelade, 1980). Similarly, if a group of elements in one area of a display is distinct in colour from the textural elements surrounding, texture segregation will occur (Beck,1982;Treisman and Gelade, 1980). If the subitizing process can make use of this preattentive information about the colour of the targets, then the usual slope discontinuity in the enumeration function should be apparent in this condition. Specifically, low level processing should deliver colour information that will distinguish the target items from distractor items. The item individuation mechanism could then assign spatial tokens to items of a particular colour; FINSTs could be assigned "top down" to the desired subset of items. Hence, it should be possible to subitize items of a particular colour, even when displays are crowded with irrelevant contours and distractor items.

If the attentional focus is in fact being moved along the contour when subjects count connected items, contour length should have an effect on counting latencies. For example, say that subjects were given the task of counting all the connected items on the predominantly horizontal winding contour in Figure 2-1, starting from the top left corner. The time it takes subjects to count the connected items should depend, in part, on the *Link length*, or the number of curves in the winding contour before a gap. This prediction follows from Jolicoeur's boundary tracing experiments, that show that the time to determine if two items are connected is dependent on the contour distance between them. In this experiment there are simply more items to be connected, and a counting task is superimposed.

When subjects are required to count items of a particular colour, however, link length should not have an effect. In this experiment there was no need to move the attentional focus along the contour in order to find out how many items of a particular

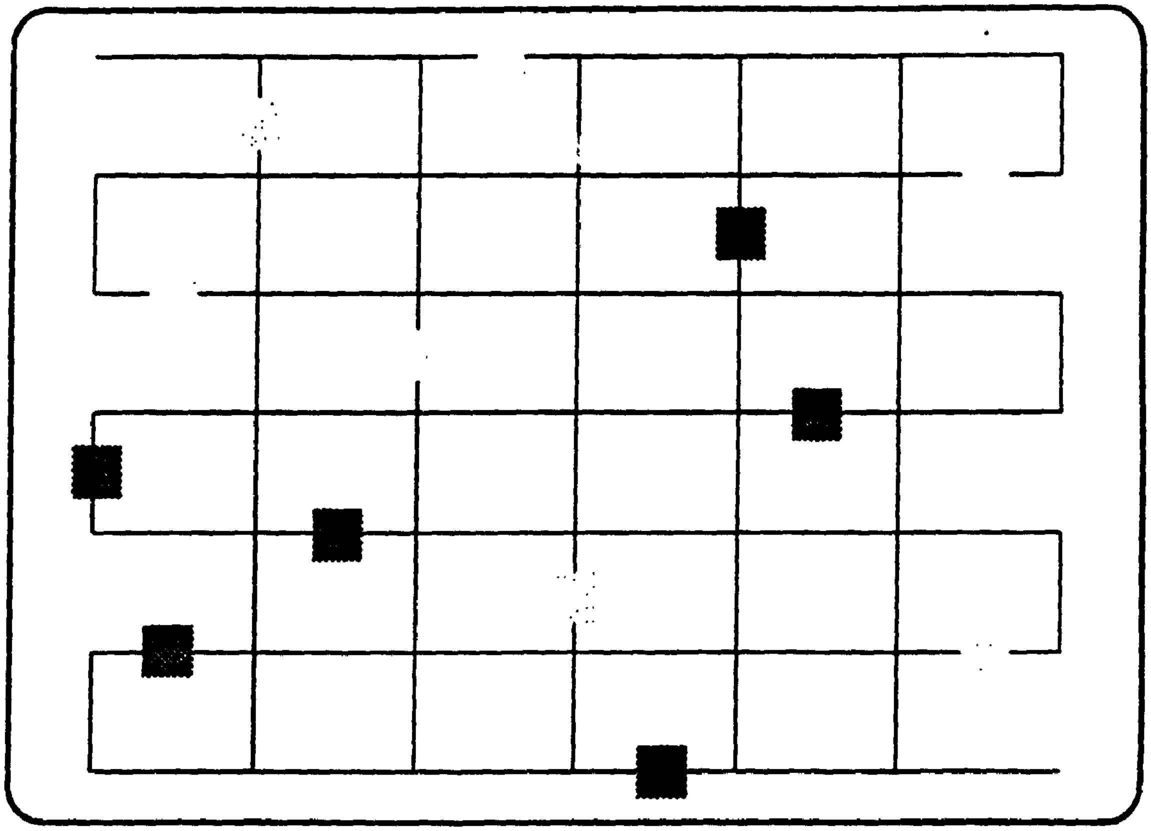

Figure 2-1: Counting items connected by a winding contour

colour there were: Target items were not necessarily located on the same contour, anyway. Link length in this case would be related to the area in which the target items would appear. Thus, target items might appear anywhere in a 4 link area, but fall on either connected or unconnected contours. If any distance has an effect on latencies in this condition, it would be the *scan path* distance, or the retinal (Euclidean) distance between the starting point and the various groups of items to be counted. It is hard to determine apriori what a particular subject's scan path through a display will be; it depends on which items are grouped and the route chosen between groups. Nevertheless, it seems unlikely that there would be a strong correlation between link length and the scan path distance because scan paths are simple straight line routes rather than the circuitous journeys along winding contours.

Thus, in summary, two factors will be of primary interest in this counting study. The first is task; subjects were required to count connected items in one condition and coloured items in another. The second factor was link length, the area in which the target items are located. Link length was included, in part, as a manipulation check, to ensure subjects were in fact moving their attentional focus along the contour when they were counting connected items.

In this, and all the following experiments, the dependent measure is reaction time, the amount of time it takes a subject to (correctly) say how many items there are in a display. Reaction time is by far the most popular dependent measure in counting studies, possibly because reaction time is an easy to gather "continuous" measure. Using reaction time permits several factors to be investigated at once without the need for a large number of observations per subject and a psychophysical design. In the following experiments there are typically one or two other factors in addition to number, so reaction time seemed the best choice as dependent measure. In addition, there were no eye movement controls in the studies, because the majority of the search and boundary tracing experiments don't use them, and I would like to compare my results with these studies. Furthermore, both procedures for controlling for eye movements, limiting exposure duration, or throwing out trials in which eye movements occured, were impracticable given the way the data had to be analyzed, and the length of the experimental sessions.

Thus, to conclude, subitizing is not expected to occur when subjects are enumerating connected items; the deviation from linearity due to the increase in slope after 3 or 4 should not be evident in the reaction time function for the *Connected condition.* In contrast, subitizing should emerge as usual when subjects are counting items of a particular colour. Second, link length should have a significant effect on latencies when subjects are counting connected items, but not when subjects are counting items of a particular colour.

Method

Design

There were three main factors in the study. The first was task. Subjects were either required to count items of a specified colour, (*Colour condition*), or count items connected by a particular line, (*Connected condition*). Task was blocked, so that subjects performed two *Colour* sessions and two *Connected* sessions. Subjects did one session of each type before they did the second session of each condition; otherwise, session order was counterbalanced. The second factor was link length, the area in which the target items could appear. There were three possible link lengths, 4, 5 and 6 links, which corresponded to dense, medium and sparse dispersion of targets in the display. The third factor was the number of target items that subjects had to count, between 1 and 8. The dependent variable was vocal response latency, the time required to say how many target items there were in the display. A randomized block factorial design was employed.

Subjects

Eight subjects participated in the four session study for payments of $40. Five of the subjects were female. The subjects were either graduate students or research assistants at the University of Western Ontario. Of the eight, six had experience in counting studies and two had never been in a counting experiment before.

Apparatus and Stimuli

An Apple II+ computer was used to generate the displays and record the data. The computer was connected to a Gerbrands G1341 voice activated relay that was used to measure vocal response latency.

In this experiment the subjects' task was to count coloured blocks. Each trial involved three displays: a fixation display, a counting display, and a mask. The fixation display was simply a white square with a black asterisk inside, projected on a black background. Subjects were asked to fixate on this point at the beginning of each trial.

Each side of the square measured 0.48 cm. Items of this size occupy a visual angle of 0.25 degrees square when viewed from 110 cm. The fixation point was used to direct subjects to the place where they were to start counting from. It could appear in three places, the top left, top right, or bottom left corner of the display. In the *Connected condition*, the fixation point would direct subjects to the end of the contour that they were supposed to trace. For example, if subjects were required to trace a horizontal contour, the end of the contour might be in the top right corner. In order to keep the conditions equivalent, it was also necessary to direct subjects to the corners of the display in the *Colour condition*. For the *Colour condition*, the fixation simply indicated where to start counting from.

The counting display was made up of coloured blocks and vertical and horizontal white lines, projected on a black background. Subjects were required to count square green or purple blocks. The sides of each block measured 0.48 cm or 0.25 degrees visual angle. There were up to 9 target blocks in each display, and either 2, 5, 6, or 8 distractor blocks. Only latencies for counting up to 8 target items with 6 distractors were analyzed, however. The trials with 9 target items were included as catch trials so that subjects could not simply guess "8" when there were a lot of items. Similarly, the trials with 2, 5, or 8 distractors were included so that item density or the proportion of the field covered by a particular colour could not be used as reliable cues to the number of target items. In the *Colour condition*, half the time the target items were green and the distractors purple, and half the time the targets were purple and the distractors green. In the *Connected condition*, the colours of each target and distractor was determined randomly for each trial and subject.

Item locations were chosen from 84 positions on a matrix. These positions are coded as locations 1-84, in Figure 2-2. The matrix was 11 by 13 cm, or 5.7 by 6.74 degrees visual angle when viewed from a distance of 110 cm. The blocks were located on a grid

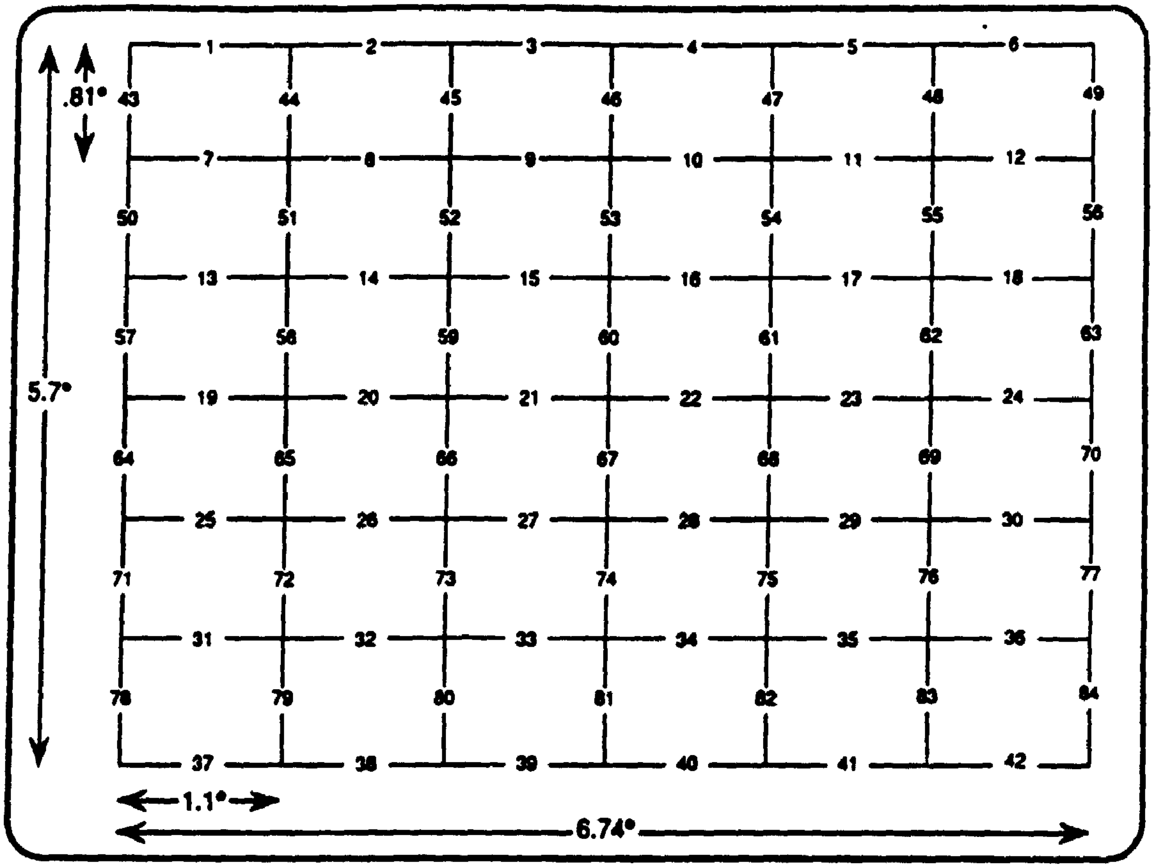

Figure 2-2: Item locations for experiment 1

of white vertical and horizontal lines. These grid lines were organized into two chains. In horizontal chains a chain was formed linking each of seven 6.74 degree horizontal lines to the line below it with a 0.95 degree vertical segment at the left or right end of the line. In the vertical chains the situation was reversed. Seven vertical line sements of 5.7 degrees were connected by 1.12 degree horizontal sements at the top or bottom. The horizontal chain intersected the vertical chain every 0.81 vertical degrees, whereas the vertical chain intersected the horizontal chain every 0.95 horizontal degrees. One of these two chains would be designated as the target chain, and would be complete. The other would be left open at the ends, leaving five parallel lines. See Figure 2-3. Half of the time the horizontal chain would be completed, and the distractor contours would be parallel vertical lines. In these cases, half the time the starting point on the target contour

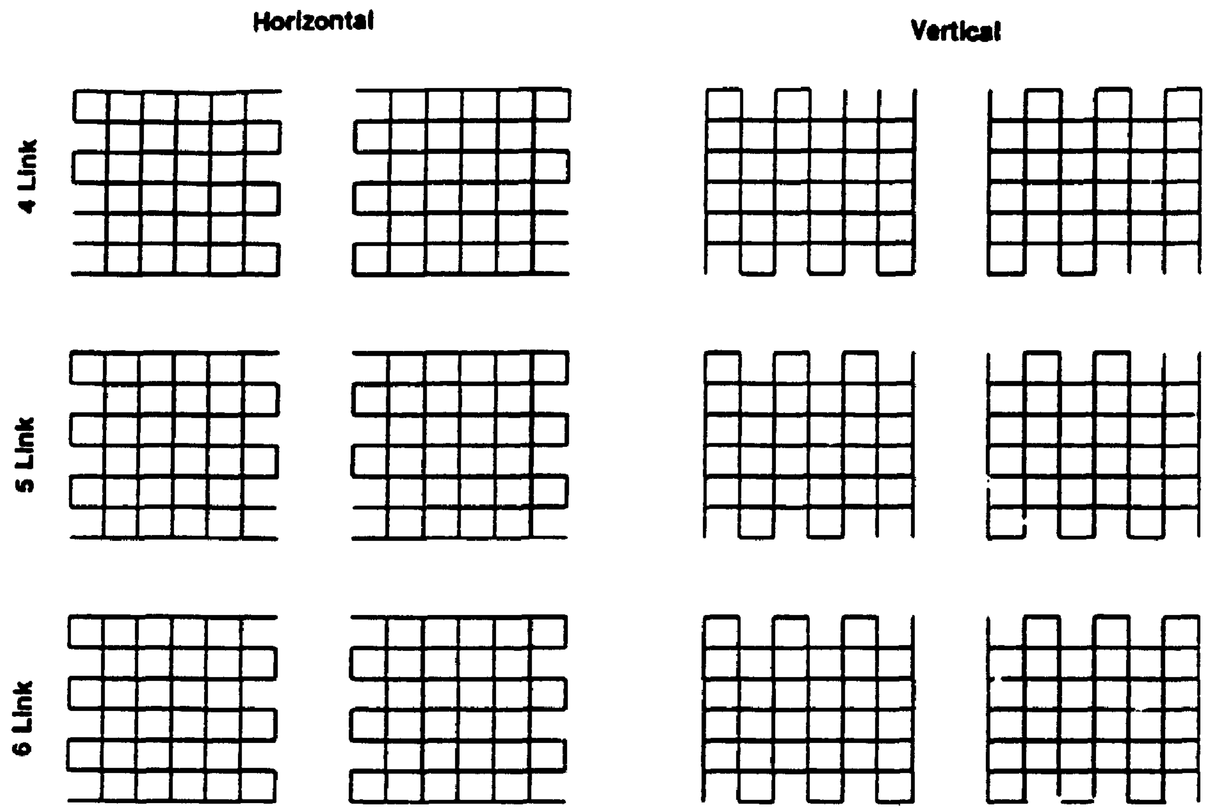

Figure 2-3: Varieties of contours for experiment 1

would be the top right corner, and the other times the starting point would be the top left corner, as indicated by the fixation. Half the time the vertical chain would be completed, and the distractor contours would be parallel horizontal lines. In these cases, half the time the starting point would be the top left corner, and half the time the starting point would be the bottom left corner.

Target items were concentrated in larger or smaller areas of the grid, depending on the variable, link length. Link length simply refers to the number of turns in the vertical or horizontal contour before there was a break, or a segment not connected to the rest of the chain. Targets were concentrated within either 4, 5 or 6 links in the chain, as can be seen from Figure 2-3. In 4 link chains there were four turns in the chain before the chain broke off. This left two lines not connected to the rest of the chain. In 5 link chains there were five corners, with one segment not connected to the chain. Six link chains had six corners, with no breaks in the contour. Link length entailed slightly different things in the *Colour* and *Connected* conditions, however. In the *Colour condition*, link length simply specified the area in which the dots could appear. For example, for a 4 link horizontal chain, target items would appear within a 6.74 by 3.8 degree area whereas for a similar 5 link chain, target items could appear in a 6.74 by 4.7 degree area. For a 6 link chain the full 6.74 by 5.7 degree area was used. In the *Connected condition*, however, link length also specified the length of the contour that had to be scanned in order to count all the connected items, because of the requirement that the target items be connected by the same line. For example, for a 4 link horizontal chain the maximum distance that would need to be scanned to find all the targets would be 37.5 degrees visual angle. For a similar 5 link chain, the maximum distance would be 45.2 degrees, and for a 6 link chain, the maximum would be 52.9 degrees. Vertical chains were slightly shorter: a 4 link chain was 33.0 degrees, a 5 link, 39.8, and a 6 link 46.6 degrees visual angle.

The procedure for positioning distractors in the displays also varied with condition. For the *Colour* condition, the distractors could appear anywhere in the grid except for where the targets were located. For the *Connected* condition, the situation was more complex. Distractors could appear anywhere but on the designated chain before the first break. Notice that this entails that the greater the link length, the fewer possible distractor locations; for 4 link contours there were 52 possible distractor locations, for 5 link, 48, and for 6 link, 42. Both target and distractor positions were chosen randomly for each trial, for each subject, given these constraints.

The masking display was simply a screen of dots of various colours. This random dot mask was used to ensure that signs of the counting display did not linger due to afterimages or gradual phosphor decay after the display went off. In this, and all the following studies, masks of various sorts were used in order to prevent subjects from cheating--setting the voice activated timer off quickly and then counting items at their leisure after the display went off.

Procedure

Subjects were invited into a slightly darkened room and were seated 110 cm from a computer screen, with the computer terminal within easy reach. (The illuminace in the room was 135 mw/m^2 as measured by a Techtronix J6502 photometer). The subjects' task was to say how many blocks of a certain type there were in each display. The type of block to be counted was specified at the beginning of each trial. Subjects were encouraged to report how many target items there were as fast as they could, with accuracy.

Each trial had five phases. First, a message appeared on the screen specifying which items the subjects had to count. In the *Colour* sessions, the message said either "Count the green items" or "Count the purple items". In the *Connected* sessions, the message said either "Count items connected by the vertical chain" or "Count items connected by the horizontal chain". These messages remained on the screen until the subject pushed the carriage return key to initiate the trial. Second, the fixation point appeared. The fixation was used to inform the subject what corner to start scanning the display from. The fixation point remained on for 2 seconds and then the computer beeped to warn the subject that the trial was imminent. The counting display appeared 256 msec after this warning tone, and remained until subjects made a response that set off the voice activated timer. Timing was initiated at the onset of the counting display. Fourth, a coloured random dot mask came on for 512 msec in order to obscure any afterimages from the counting display. Finally, a message appeared on the screen asking subjects to type in the number that they had said or an "M" for mistrial. Subjects were instructed to use the "M" response if the display disappeared before they had a chance to count the items or if the display did not disappear after they first pronounced the number. Error feedback was given after the response was typed in. If a mistake was made the computer beeped five times and reminded the subject which items had been targets on that particular trial. Error trials and mistrials were readministered at the end of each block of 28 trials. Attempts were made to disguise the appearance of readministered trials because subjects very quickly learn to avoid the counting process by using shape cues to derive the cardinality of the display if they are presented the same displays again and again (Mandler and Shebo, 1982).[2] The position of the fixation point and the start of the winding contour was changed. It was hoped that different item groupings might emerge if subjects started from a different point in the display, and thus the subjects might not recognize the display as familiar. Otherwise, the readministered trials were identical to the originals.

There were 224 trials in each of the four sessions. Of these, 192 were experimental trials involving 6 distractors and between 1 and 8 targets. The remaining 32 were catch trials, in which there were 9 targets or 2, 5 or 8 distractors. Before the beginning of each session there were 28 practise trials.

The experiment was run in 4 days. The last session took place within a week of first session.

Results

The latency data were analyzed in three ways. First, analysis of variance was performed. Second, trend analysis was performed on averaged and individual datasets, in order to determine if there was evidence of subitizing in the latency functions for the *Colour* and *Connected* conditions. Finally, the slopes for the subitizing and counting functions were calculated using linear regression analysis. For the sake of clarity, the

---

[2] A similar thing occurs when you learn to use dice. For example, although initially you many need to count six dots on the dice, after awhile you can recognize the number by the configuration of dots, two dense lines of spots.

following discussion will be quite selective; only analyses with implications for the experimental hypotheses will be discussed, although many more analyses could be performed. For more complete coverage of these analyses, as well as a listing of the means and standard deviations for each condition, see Appendix A.

Data pre-processing

Error trials were readministered until subjects got them right, and consequently there were no errors to be considered.[3] In addition, outliers were dropped because there were a minimal number of observations per cell and an atypical latency could have a major effect on the results. Outliers were defined as the reaction time farthest from the mean for each task, link length, number and session for each subject. Thus, for each subject, task, number, link length, and session, the reaction time farthest from the mean was dropped as an outlier. This unusual procedure was employed because within subject analysis was done--the separate analysis of each individual's dataset. Within subject analysis was used for individual trend analysis, in order to determine for each subject whether there was any evidence of subitizing. In these analyses each reaction time was considered a case. It was important that there be the same number of cases (reaction times) in each cell because unequal cell variances were typical; a given subject might have a standard deviation of 60 msec when counting to 1, and 300 msec when counting to 7, for example. When unequal cell variances are accompanied by unequal cell sizes, F statistics, used in analysis of variance and trend analysis, are exaggerated. Thus, the probability of a type I error increases (Milligan, Wong and Thompson, 1987). Unequal cell variances do not pose a problem unless there are also unequal cell sizes, however. Dropping one outlier per cell was a way of ensuring that the number of cases per cell was constant, while trying to prevent extreme latencies from unduly affecting the means. Once this procedure was performed for the within cell analyses, it made sense to use

[3] Nonetheless, the number of errors was recorded. Overall, the average error rate was 4.5%. For further discussion of the error analyses, see Appendix A.

these same reaction times when calculating a mean for each subject, to be entered into the between subject analyses--analyses in which each case represented one subject's mean reaction time. Analysis of variance, regression, and averaged trend analysis were performed on these means.

Analysis of variance

Analysis of variance was performed for two reasons. First, it was used to determine if changing the task from enumerating coloured objects to connected objects affected counting latencies. Second, and more important, it was used to determine if the variable link length had a greater effect on counting latencies in the *Connected condition* than in the *Colour condition*. As predicted, task had a major effect on how long it took to count ($F(1,7)=168.6$, $p < .001$); subjects required an additional second to count connected items over coloured items even when counting to 1. See Figure 2-4. Moreover, there was the

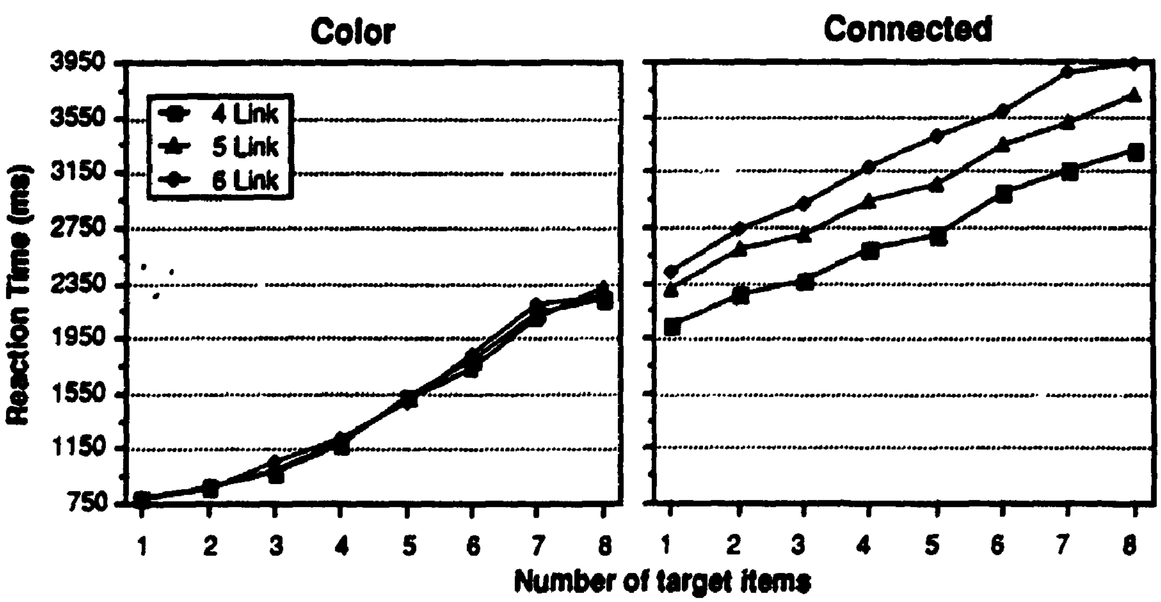

Figure 2-4: Average latencies to count coloured vs connected items

predicted interaction between link length and condition ($F(2,14)=62.3, p<.001$); thus, for example it required 1231 msec longer to count to 1 in the *Connected condition* than the

Colour condition even when the link length was only 4, and this difference became more pronounced as the link length increased, to a maximum of 1630 msec at 6 links. Furthermore, if the *Colour condition* trials are analyzed separately, link length has no effect on latencies ($F(2,14)=0.9$, $p > .1$). Thus, it required no longer to count items dispersed over a large area than it did when the targets were concentrated in a small area. In contrast, link length had a a strong effect on latencies in the *Connected condition* ($F(2,14)=99.2$, $p < .001$). The dispersion of the targets, or more accurately, the length of the contour connecting them, determined how much time was required to count the connected items. With shorter link lengths, subjects responded faster. For example, in order to count to 1, it required 2030 msec on a 4 link, 2307 msec on a 5 link, and 2423 msec on a 6 link horizontal chain.[4]

As can be seen from Figure 2-4, latency functions for the 4, 5 and 6 link conditions are not parallel, however. Specifically, times to count 1 or 2 items in the 6 link condition were somewhat lower than expected given the distance that had to be traced. This may be because in 6 link trials there was a substantial probability that all the targets and distractors would occur long before the end of the contour, as in Figure 2-5. If subjects were to trace to the end of the contour in this display, they would continue long after there were any targets or distractors to check. It seems unlikely that subjects move the focus blindly in this way. For example, in Treisman's search studies, subjects don't persist in looking for conjunctions when there are no more items to check: This is why reaction time, display size, and response are related as they are in search studies. There is evidence that subjects may have stopped boundary tracing short of the end of the

[4]In fact, a series of analyses were performed, correlating each subject's counting latency for a given trial with five distance measures: Euclidean distance between the fixation point and the farthest target, Euclidean distance between the fixation and the farthest target or distractor, maximum contour distance between the fixation and the farthest target, maximum contour distance between the fixation and the farthest target or distractor, and finally link length, the distance along the contour before there was a gap. None of the variables correlated significantly with counting latencies in the *Colour condition*, though there were significant correlations between counting latencies and link length in the *Connected condition*.

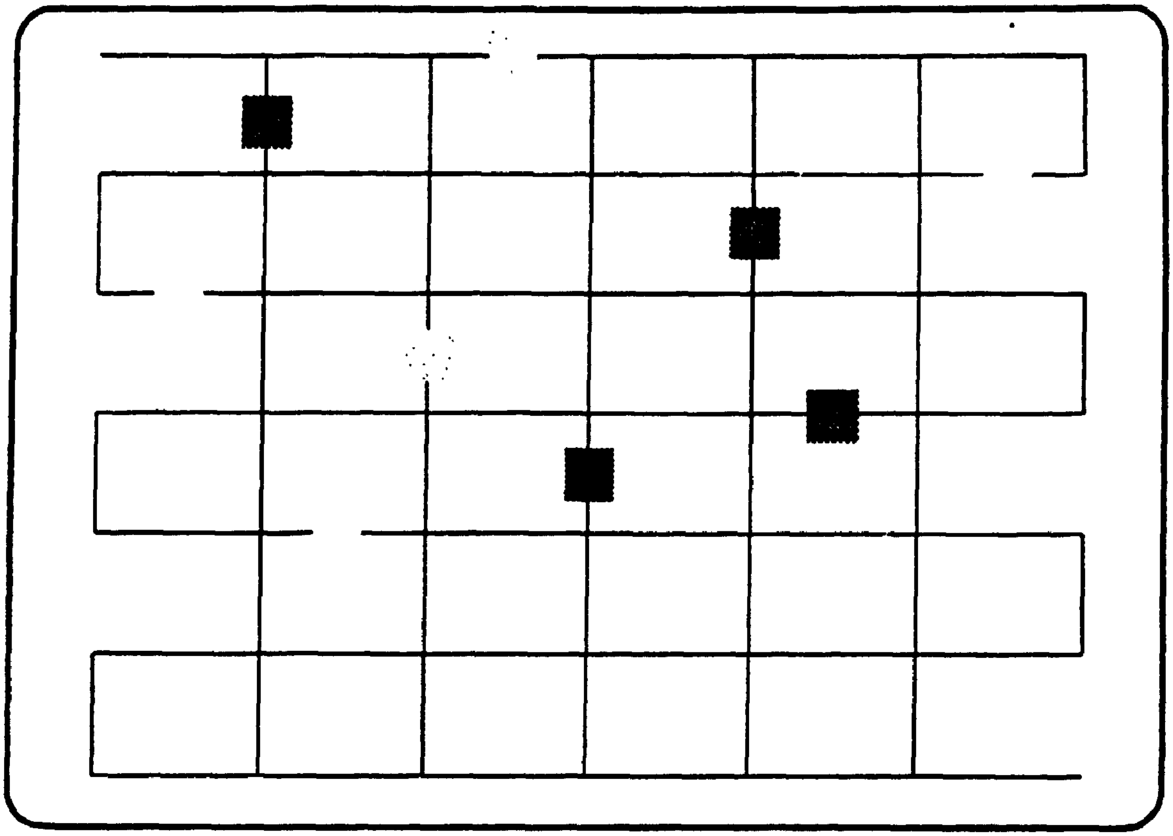

Figure 2-5: Display in which targets and distractors run out before the contour does

boundary in displays such as 2-5. These analyses are presented in Appendix B, as is an explanation of why it was more probable that all the targets and distractors fall in an area before the end of the contour in 6 link trials.

By comparing the time to count 1 item between the 4 and 5 link trials, it is possible to calculate the speed with which the attentional focus was moved, although such an estimate is no doubt inflated by eye movements given the displays in this experiment. On average, it requires 277 msec longer to count 1 item on a 5 link contour than on a 4 link contour, averaged across session 1 and 2. The difference in length between 4 and 5 link contours is 7.26 degrees, averaged between horizontal and vertical chain trials. Consequently, subjects require 38.2 msec/degree to scan contours in this experiment.

To summarize, then, as predicted, task had a major effect on counting latencies; subjects were much faster at counting items of a particular colour than items that were connected. Moreover, although the dispersion of the targets had no effect on counting latencies in the *Colour condition*, the dispersion of the targets, specifically the length of the contour to be traced to locate the connected targets, had a major effect on latencies in the *Connected* condition. This would be predicted if subjects were moving the spatial processing focus along the winding contour at a fixed rate in order to count the connected items. Latencies were also influenced by the area that target and distractors items occupied in the *Connected* condition, however, as discussed in Appendix B.

There was one unexpected finding that, though unrelated to the major hypotheses of the study, has marked effects on the way the balance of the analyses will be reported. When the *Colour* condition trials were analyzed, it became apparent that the colour of target items had a major effect on latencies ($F(1,7)=63.9$, $p < .001$). Subjects were much slower at counting purple items than green ones, and moreover this effect interacted with number ($F(7,49)=3.5$, $p < .005$). See Figure 2-6.[5] As a result of this finding, care was taken to analyze green and purple trials separately in the *Colour* condition.

Trend analysis

The primary source of evidence that different processes are being used for small and large numbers is the sudden increase in slope in the reaction time function after 3 or 4 items. If this increase is large enough, trend analysis will register this change as a significant deviation from linearity. Thus, in this section, the focus will be whether there are deviations from linearity in the latency functions. Although trend analysis is often used in counting and search studies (Akin and Chase, 1978; Chi and Klahr, 1975; Francolini and Egeth, 1980; Klahr and Wallace, 1976; Oyama, Kikuchi and Ichihara,1981; Treisman and Gelade, 1980), and seems to be the best statistic available for this task, it is not ideal, or for that matter, specifically designed for within subject

---

[5] In order to perform within subject analyses, each colour had to be analyzed separately. Consequently the reaction time farthest from the mean for each subject, number, link length, and COLOUR was dropped as an outlier.

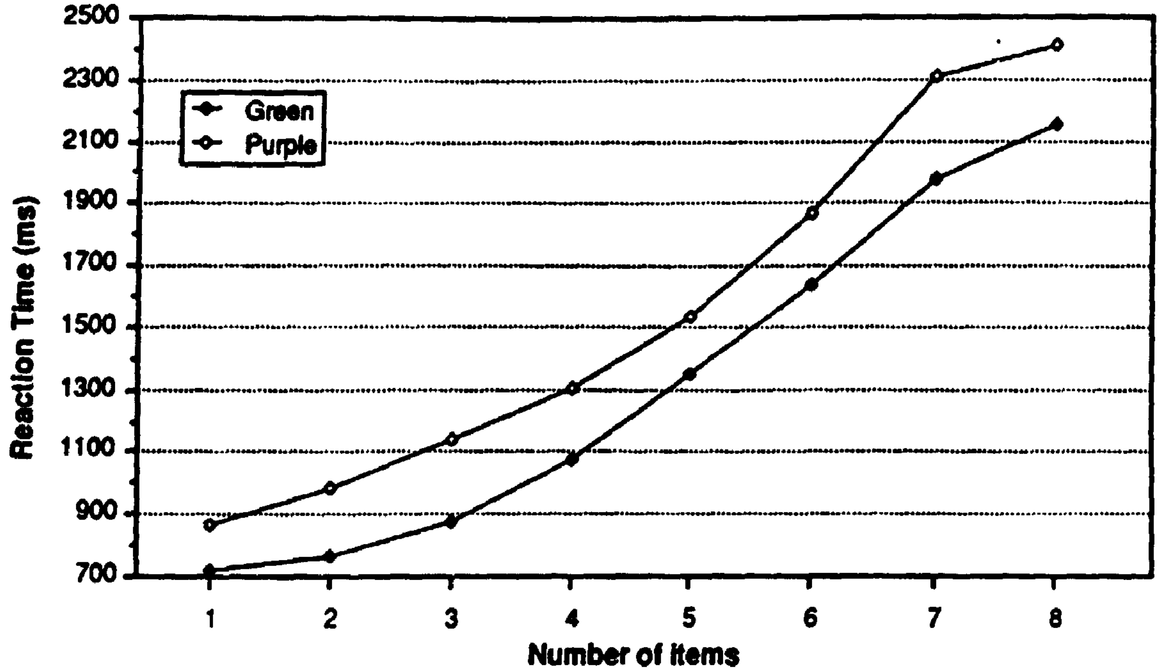

Figure 2-6: Average counting latency for green and purple items in the *Colour Condition*

comparison. Further, trend analysis cannot distinguish deviations from linearity that result from an increase in slope, such as observed in subitizing, from deviations caused by sudden drops in reaction time, or plateaus. In order to confirm that each deviation from linearity resulted from an increase in slope, rather than a drop in slope, it is necessary to check each graph at the point where the deviation from linearity first becomes significant. This involves performing trend analysis various subranges of items, 1-3, 1-4, 1-5, 1-6, until the deviation from linearity becomes significant.[6] Then, if this deviation came about because of an increase in slope there would be evidence of subitizing, especially if the slope remains more or less constant after that point. In

[6]In addition, it is important that this deviation from linearity remain significant after the number where the deviation first arises. For example, sometimes a significant deviation from linearity would emerge in the 1-4 range but there would be no significant deviation from linearity in the 1-5 range. In this case the deviation from linearity in the 1-4 range was judged to be an product of an atypical latency at 4 and not the result of the change from the subitizing to counting process. For a deviation from linearity to be reliable it should be apparent in the data when larger subranges are examined.

addition, trend analysis, like all analysis of variance based statistics, is sensitive to within cell variance. Thus, in the same way that a given difference between means is more likely to be registered significant if the variance within cells is minimal, any systematic trend (linear, non-linear deviation, quadratic) is more likely to be registered significant if the within cell variance is minimal. As a consequence, deviations from linearity may emerge earlier in "easy" conditions, in which there is little variability simply by virtue of variance alone, and it may appear that subjects can subitize more items in difficult conditions than easy.[7] For this reason standard deviations for counting latencies were also monitored, as were the slopes of the latency function at various stages.

Trend analysis was first performed on the averaged datasets, in which each case represents the average reaction time for a particular subject. In this experiment subjects came back for four sessions, and were no doubt well aware that there were usually no more than 8 items in the display. This awareness may have contributed to the end effect that can be seen in the latency data; there is a "terminal drop" in the counting slope between 7 and 8. For example, see Figure 2-3, *Connected condition* 6 link trials in particular. This drop could be registered as a deviation from linearity. To avoid mistaking this deviation for one that arose from a significant increase in slope as a result of the change from subitizing to counting, only the 1-7 range was analyzed. Otherwise, the registered deviations from linearity might reflect the significant downward arch of the reaction time function after 7, while the increase in slope at 4 or 5 is negligible.

Linear trends were apparent in all conditions, but deviations from linearity were only evident when subjects were required to count items of a particular colour. Thus, when subjects were required to count connected items, there were no significant deviations

[7] To complicate matters, high standard deviations are often indicators of the counting process (Akin and Chase, 1978). Thus, for example, if the standard deviations for the latency remain constant and high through the 1-8 range, there is evidence that the same process is being used for the range of numbers, and that process is the counting process.

from linearity ($F(5,105)= 0.3$, $p > .1$; $F(5,105)= 0.2$, $p > .1$; $F(5,105)= 0.1$, $p > .1$ for 4, 5 and 6 link trials respectively). In contrast, when subjects were required to count items of a particular colour, significant deviations from linearity occured, regardless of the colour of the items ($F(5,161)= 6.1$, $p < .001$; $F(5,161)= 3.6$, $p < .005$ for green and purple respectively), or the density of the items ($F(5,105)= 2.9$, $p < .05$; $F(5,105)= 2.8$, $p < .05$; $F(5,105)= 2.9$, $p < .05$ for 4, 5 and 6 link trials respectively), though the scale in Figure 2-4 serves to obscure most of these discontinuities. On the basis of these averaged data, it would seem that subjects could subitize up to 4 for both green and purple items ($F(3,115)=3.9$, $p < .05$; $F(3,115)=3.8$, $p < .05$). When the 4, 5 and 6 link trials are analyzed separately, in all three cases subjects appear to be subitizing to 5, however ($F(4,90)=3.0$, $p < .05$; $F(4,90)=2.2$, $p = .08$; $F(4,90)=2.1$, $p = .09$).

It is possible that averaging across subjects might obscure changes in slope, however, given that there are individual differences in how high people can subitize (Akin and Chase,1978). For this reason, trend analyses were also performed on individual datasets, in which each case represents one trial for a given subject. See Table 2-1. All eight subjects showed appropriate deviations from linearity in the *Colour condition*, regardless of the density of the target items (link length) or colour of the target items, as can be seen from the graphs of individual data presented in Appendix A. Surprisingly, given how noisy the individual datasets were, there were only three cases in which significant deviations from linearity occured in the *Connected condition*. Two occured as a result of a precipitous increase in latency to a given number, n, followed by a precipitous drop in latency at $n + 1$, as would be expected if an outlier was missed. There was only one case in which the deviation from linearity looked like the sort of deviation that occurs when the subitizing process gives way to the counting process. For all other subjects and link lengths, no deviations from linearity were evident, as would be expected if the same process was being used to enumerate both small and large numbers of items. There were significant differences in the proportions

Table 2-1

Trend analysis of individual datasets for connected study

NUMBER OF SUBJECTS SHOWING DEVIATIONS FROM LINEARITY

COLOUR counting condition	
Overall	8/8
Green items	8/8
Purple items	8/8
CONNECTED counting condition	
Overall	0/8
Horizontal	0/8
Vertical	0/8
4 link	0/8
5 link	1/8
6 link	0/8

NUMBER OF SUBJECTS WITH DEVIATIONS FROM LINEARITY

	Overall	Green items	Purple items
# subitizing to 2	1	1	-
# subitizing to 3	5	4	1
# subitizing to 4	2	2	6
# subitizing to 5	-	-	1
# subitizing to 6	-	1	-

of subjects showing deviations from linearity in the *Colour condition* and the *Connected condition* (χ^2 (1)=16.0,p<.01 for the 4 and 6 link trials and χ^2 (1)=12.04,p<.01 for the 5 link trials).

In the *Colour condition*, the only condition in which there was evidence of subitizing, it was possible to find out how high each subject subitized by looking at where the deviation from linearity emerged. Overall, most subjects subitized to 3 in colour counting condition, as can be seen from Table 2-1. Looking at the breakdown by colour, however, it seems that most subjects subitize to 4 when counting purple items and 3 when counting green, however. This difference probably reflects greater variability in latency to count 1-3 items in the purple condition than the green, as was apparent in the data of individual subjects.

Regression analysis

Regression was performed on the averaged datasets in order to determine the slopes for the enumeration functions. As can be seen from Table 2-2, slopes in the 1-3 range varied between 86 msec/item to 135 msec/item in the *Colour condition*, depending on link length.[8] When the data for green and purple trials were analyzed separately it became apparent that the subitizing slope for purple items was almost twice that for green, 141 msec/item as opposed to 77 msec/item in the 1-3 range, however. Moreover, these slopes were significantly different from one another; each fell outside the other's 95% confidence interval. In the *Connected* condition the slopes in the 1-3 range were larger than the same slopes in the *Colour condition*, ranging between 170 msec/item to 239 msec/item in the 4 to 6 link range, though the differences between the slopes in the 1-3 range for the *Colour* and *Connected* conditions are not significant.

[8]A conservative estimate of the subitizing range was employed in order to avoid inflating slopes with latencies from trials in which subitizing did not occur. In addition, the counting slope was calculated using the 5-7 range, in order to avoid deflating the counting slope with the terminal drop.

Table 2-2

Regression analysis for connected study

Slopes are in milliseconds per item.

	Slope	95% C.I.	R
SUBITIZING RANGE (1-3)			
COLOUR CONDITION: Counting green or purple blocks			
4 link chain	86	9 - 162	.32
5 link chain	99	28 - 171	.38
6 link chain	135	63 - 207	.49
CONNECTED CONDITION: Counting connected blocks			
4 link chain	170	47 - 292	.38
5 link chain	191	31 - 351	.33
6 link chain	239	63 - 414	.37
COUNTING RANGE (5-7)			
COLOUR CONDITION: Counting green or purple blocks			
4 link chain	300	177 - 423	.59
5 link chain	295	187 - 402	.63
6 link chain	353	229 - 477	.64
CONNECTED CONDITION: Counting connnected blocks			
4 link chain	239	108 - 370	.47
5 link chain	233	67 - 398	.38
6 link chain	233	50 - 416	.35
Effect of item color: Colour condition only			
OVERALL			
1-3	109	64 - 154	.37
5-7	308	236 - 379	.58
GREEN items			
1-3	77	25 - 130	.33
5-7	310	226 - 394	.66
PURPLE items			
1-3	141	82 - 201	.49
5-7	305	208 - 403	.60

The 5-7 range slopes in the *Colour condition* were significantly different from their respective subitizing slopes, regardless of the colour of the items. Slopes varied between 300 and 353 msec/item as a function of link length. Further, the slopes for the counting functions for green and purple items were almost identical, in the *Colour condition*, 310 msec/item as compared to 305 msec/item respectively. Thus, it appears that either the colour or luminance of the items to be enumerated had a significant effect on the subitizing slope, but not on the counting slope. In the *Connected condition*, the 5-7 range slopes are somewhat smaller, varying between between 233 msec/item to 239 msec/item. The counting slopes in the *Connected condition* are lower than most in the literature, probably because addition is part of the adult counting process (Klahr, 1973b; Woodworth and Schlosberg, 1954), and addition latencies depend on the magnitude of the addends (Parkman and Groen, 1971): the smaller the minimum addend, the faster the addition can be performed. In the *Connected condition*, presumably subjects had to go through the display, item by item, to determine how many connected items there were. Thus, subjects were only adding by ones. Repeated additions by one are faster than additions by any other number. Nonetheless, the reduced slope in the 5-7 range no doubt contributed to the finding that there were no significant differences in slope between the 1-3 and 5-7 ranges.

Discussion

As predicted, even when displays were crowded with irrelevant lines and distractor items, subitizing was evident in the *Colour condition*; there were significant deviations from linearity, and moreover these deviations seemed to be caused by sudden increases in slope after 3 or 4. In contrast, there was little evidence of subitizing in the *Connected condition*. Given that subjects would presumably want to subitize if they could, because that would allow them to respond faster, it seems that subjects *cannot* subitize connected items. I would like to argue that this is because the mechanism for individuating feature clusters, the FINST mechanism, is preattentive, and spatial attention is required to derive the connected relation, as suggested by studies on search, texture segregation and boundary tracing (Treisman and Gormican, 1988; Julesz,1984; Jolicoeur,1988). As a result, in the *Connected condition* subjects can't use the FINST mechanism to selectively individuate targets.

As predicted, link length only affected latencies when subjects were required to count connected items. This result makes sense in light of Jolicoeur's finding that the time required to determine if two items are connected is a function of the link length, or contour distance between them. In this experiment subjects were simply required to enumerate the connected items. Jolicoeur's findings were taken as evidence that subjects move the attentional focus at a fixed rate along a contour of a particular complexity. His scan rates are faster than mine (21 msec/degree as opposed to 38 msec/degree), probably because the contours in this study were longer and more convoluted, and were regularly intersected by distracting lines which would be expected to interfere (Jolicoeur, 1988).

One other finding was not specifically predicted, however, and had to do with the effect of the particular colour of the target items in the *Colour condition*. A pilot study had shown that subjects required approximately 55 msec longer to locate a purple item in a field of green and white distractors than to locate a green item in a field of purple and white distractors. (See Appendix C). Thus, a main effect of colour was not unforeseen. The surprise was the interaction between colour and number. Specifically, the subitizing slope for green items was approximately half that of the purple items, 77 msec/item as opposed to 141 msec/item. In contrast, the counting slopes were almost identical: 310 msec/item for green items and 305 msec/item for purple items in the 5-7 range.

Why did this interaction occur? Consider the differences between green and purple items. First, and most obvious, green and purple items differed in colour, as determined by the various wavelengths and intensities of light radiated from each target item. In this experiment the conditions were not ideal for colour vision, however. The displays were large, 6.74 by 5.7 degrees, and most colour receptors are crowded into a 2 degree foveal area of the retina. Moreover, the experiment was conducted in a slightly darkened room to prevent reflections and glare from the video screen, and though the conditions could not properly be called scotopic, there is the possibility that sensitivity to long wavelenth light might suffer because of the Purkinje effect. Consequently, subjects might be less sensitive to purple light than usual.

A related factor is luminance, or the amount of light radiated from a stimuli. Luminance was measured using a Techtronix photometer, fitted with a J6523 1 degree narrow luminance probe, adjusted to match the human spectral sensitivity function. Measured was the luminance of a single target item situated near the centre of a 6 link horizontal chain, surrounded by six distractors of the opposing colour. At a 1 meter viewing distance, under the lighting conditions employed in the experiment, the black background radiated 1 cd/m^2. A single green item radiated 55 cd/m^2. A single purple item, in the same location radiated only 41 cd/m^2, a value very similar to the luminance of an area containing a white contour line on a black background, 38 cd/m^2. The minimal difference between the luminance of a purple item and the luminance of a similar sized area with a white contour might explain why subjects often commented that they purple items seemed to "hide" in the contour lines.[9] Therefore, both colour and simple luminace would serve to decrease the contrast between the purple items and the background and background matrix. Purple items contrasted less with the background than green because under mesopic conditions sensitivity to long wavelength light is reduced. Similarly, purple items contrasted less with the background contours than the

[9] Subjects made significantly more errors when counting purple items than when counting green ($F(1,7)=9.4$, $p < .05$). The colour effect did not interact with the number of items in the display, though ($F(7,49)=1.5$, $p > .1$). Error rate must be interpreted with caution in this study, however. Errors arose from mistakes in typing as well as mistakes in counting. In addition, subjects sometimes counted the distractors instead of the targets in the *Colour condition* because they forget the which items they were supposed to count in a particular trial.

green items, because the luminance of a purple item was almost identical to a similar sized area containing a white contour on a black background. The finding that contrast-related factors affect the subitizing slope but not the counting slope may indicate that subitizing is a data-limited process whereas counting is more resource-limited (Norman, 1975). Subitizing would be data limited because the subitizing slope is affected by the qualitity of the stimuli. The counting process is less affected by the quality of the stimuli because the principal limitation in counting is access to the processing focus resource. Thus, the enumeration process in the 1-3 range might be data limited whereas therafter the process becomes resource limited.

Results from two other studies also suggest that there is a relationship between contrast and the subitizing slope. Hunter and Sigler (1940) first showed the relationship between contrast, number and the required exposure duration for 50% accuracy in enumeration for small numbers of items. (In their study "small" was less than 8, whereas in this and other studies "small" is less than 4. This discrepancy highlights the perils of comparing studies with 50% accuracy criterions with those with 100% criterions). More recently, Liss and Reeves (1983) showed that with limited exposure duration or backward masking the subitizing slope can increase markedly when reaction time is measured, though the slope in the counting range is relatively unaffected. (Both manipulations would affect stimulus quality, reducing the contrast between the target items and the background). In this study, however, the relationship between colour and slope might simply reflect a greater tendency to perform eye movements in purple than green trials. Subjects were encouraged, by instruction and aversive error messages, to be very accurate. Because it was difficult to discriminate purple items from the background on the basis of either colour or luminance, subjects might routinely perform eye movements in the purple trials and only perform eye movements to distant items in the green trials.

[10]Eye movements have been shown to inflate subitizing slopes (Klahr, 1973a). It is not clear at present why the colour of the items exerted their effect on the subitizing range and not in the counting range. Before this question can be adequately addressed there is a need to replicate this experiment while monitoring eye movements.

Limitations

The displays in this study were made large to accomodate a large number of items, and a considerable variety of link lengths. Further, subjects were directed to start their search in the corner of the display rather than the centre. Consequently, eye movements were inevitable in this study. Nonetheless, it seems unlikely that eye movements alone could explain why there were no deviations from linearity in the latency function in the *Connected condition* while there were in the *Colour condition*. First, eye movements inflate both the subitizing and counting slopes, given that subitizing seems to be a part of the counting process (Klahr and Wallace, 1976). For this reason, it seems unlikely that eye movements somehow obscured the difference between subitizing and counting slopes in the *Connected condition*. Second, given that retinal sensitivity to colour is minimal outside a 2 degree area, eye movements probably occured quite frequently in the *Colour condition* as well as in the *Connected condition*. Subjects would have to make eye movements to ensure that the target and distractor stimuli were projecting on an area of the retina that was sensitive to colour. As a result, eye movements should also contaminate the subitizing and counting slopes in the *Colour condition*.

---

[10]If this were true then there should be a stronger correlation between the Euclidean distance between the fixation and farthest item in the green condition, than in the purple condition in the subitizing range. This prediction was confirmed ($F(1,7)=18.8$, $p < .005$). As expected, there were stronger correlations between response latency and the euclidean distance of the farthest target for green targets than purple in the 1-3 range. The average correlation between distance and response latency was approximately .19 for green items in the 1-3 range, but was only .02 for purple items in the 1-3 range, as would be expected if the distance to farthest target had little effect on latencies because subjects were moving their focus regardless in purple trials. For both green and purple trials there was little correlation between distance to the farthest target and response latency in the 5-7 range, .05 and .09 respectively. Presumably, this particular distance measure had little effect on latencies in the 5-7 range because the length of the scan path rather than the distance to the furthest target influence latencies in the counting range. On the other hand, the sizes of these correlations are not impressive, at best. At this point it is hard to say whether eye movements could explain the slope differences between the green and purple items.

Similarly, it seems unlikely that the effects of link length in the *Connected condition* could be completely explained by eye movements. Jolicoeur (1988) has already demonstrated that the relationship between contour distance and reaction time in connected decisions remains strong even if eye movements are prevented by keeping the exposure duration short. Thus, although subjects no doubt made all sorts of eye movements in this experiment, the results of this experiment could not be completely explained by them. Nonetheless, the study should be replicated with better control for eye movements.

In conclusion, despite various problems, it seems that subjects have greater difficulty enumerating small numbers of items when they have to distinguish connected items from others that are not connected, a process thought to involve spatial attention (Jolicoeur,1988). The deviations from linearity that are the trademark of the switch from the subitizing to counting process are notably absent in the *Connected condition*, and the slope in the 1-3 range was somewhat high for a subitizing slope. In contrast, when subjects are required to enumerate items of a particular colour, there was evidence of subitizing. Deviations from linearity were apparent whether subjects were counting purple or green targets. This result makes sense if it is assumed that subitizing exploits a preattentive mechanism. When targets and distractors differ by virtue of a property that emerges preattentively, subitizing items in a field of distractors is possible.

Chapter Three
Enumeration and Visual Routines: Resolving items as wholes

Usually when we enumerate, we count objects rather than points of light. Objects may be defined by many contours or intensity discontinuities, and at times objects may partially occlude each other. Moreover, objects may even move. Nonetheless, we seem to be able to subitize anyway; we have a sense of the rapid, accurate, effortless apprehension of number when there are fewer than 4 or 5 items in the visual scene. Consequently, we can easily know that there are 3, when we see three variegated macintosh apples in a bowl, or we may immediately sense 4, when we see four holsteins gamboling across the feedlot. How is this possible? According to Ullman (1984) many of the properties that define an object as a whole do not become available until after the application of a visual routine. Thus, properties such as global shape as well as the spatial relations between parts of an object do not become explicit until after attentive analysis. If the subitizing process can only make use of the information that is available preattentively, then how is it possible to subitize whole objects? I would like to argue in this section that it is not always possible to subitize objects, and in cases where the low level processing is misled by the configuration of objects into grouping contours from different objects, or assigning the same place token to several different objects, it is difficult to subitize.

When is low level information an unreliable cue to number of objects? If Marr's (1982) analysis of the problem is correct, place tokens are assigned to intensity discontinuities, and then tokens are grouped by principles such as those known by the Gestalt psychologisits as proximity, similarity and common fate. Thus, items that are close together, or similar in contrast, orientation, depth, brightness, size, or motion, may be grouped together and assigned a place token. Presumably, a place token would be assigned to the centre of such a clustering. Then a group of place tokens may itself be assigned a place token, and the process could repeat. Thus tokens could be assigned at a number of different levels of resolution. Tokens at any level may be enumerated, although enumeration at some levels may be more natural and effortless than others (cf., global to local processes, Navon,1977). Thus, for example, Figure 3-1 might be seen as 372 horizontal line segments, 186 equal signs, or 5 boxes.

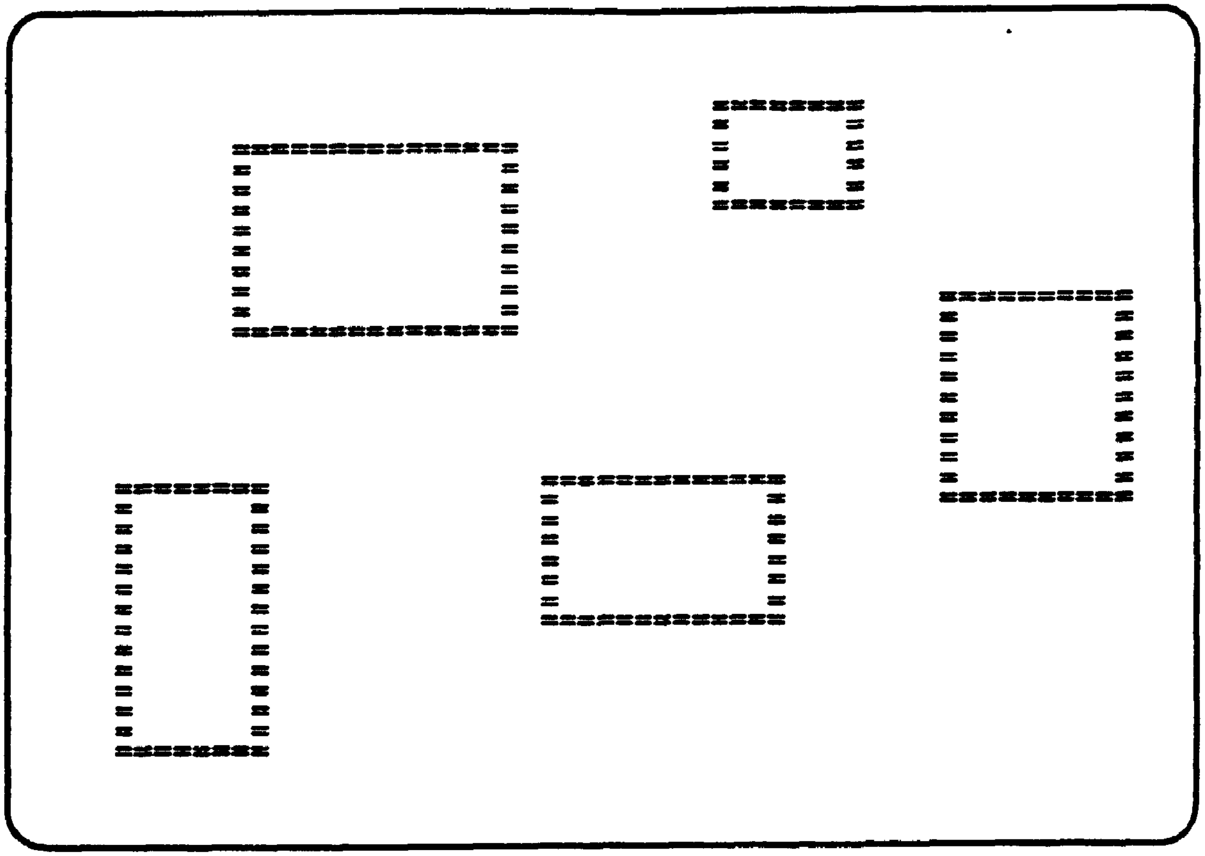

Figure 3-1: Grouping contours into units

Concentric items might be especially difficult to enumerate given this sort of system. The reason is that concentric items all have the same centre, or focus. A place token might be assigned to the centre of a group of lines that form the sides of an item, but if another larger item surrounds this item the place token might be located on the same place. In fact, in such a situation the system might simply assign one token to all of the contours.

Given this expectation, it is interesting that one of the few studies that failed to show evidence of subitizing had subjects counting concentric circles. Saltzman and Garner's (1948) study was unusual in that it had subjects counting figures, rather than dots or points of light (which are still the stimuli of choice in counting studies). They found that under some conditions when subjects enumerated concentric circles, latencies increased steadily with number (Figure 4, p. 235, cf. Woodworth and Schlosberg, 1954; Allport, 1975). There was no sign of the slope discontinuity that augurs the transition between subitizing and counting. The authors did not pursue this result, though later in the same paper they present a dot enumeration study that showed clear evidence of discontinuities. At the time trend analysis was not in use, and the authors were busy demonstrating that there was a subitizing slope in the first place.

There are reasons to suspect the validity of Saltzman and Garner's results, however. They gave their subjects ample opportunity to use cues other than number to decide the cardinality of the display. First, they set up the displays so that the largest ring was always 14 degrees in diameter, and the distance between rings was the radius of the smallest ring in the display. This would make it possible for the subject to avoid the counting task, and with practise, simply make the number judgment on the basis of the distance between rings, or the diameter of the innermost circle. Moreover, they gave their subjects many exposures to the same stimuli, and with repeated exposure subjects may begin to use their memory for form to avoid the counting task (Mandler and Shebo,1982). The objective of this study is to replicate Saltzman and Garner's finding, setting up the displays in such a way that the cardinality of the display could not be decided except by counting or subitizing the items. Moreover, I wanted to ensure that grouping contours by relative proximity was possible, given that grouping by relative proximity seems to be a natural component in the counting process (Van Oeffelen and Vos, 1982). Finally, I wanted to ensure the result would replicate even if the items were presented within the foveal area of the retina.

3.1. Experiment 2: Counting concentric objects

The goal of the present study is to discover, by the process of elimination, why there was evidence of subitizing in Saltzman and Garner's dot enumeration task but not in their concentric circle task. The tasks differed in three ways. First, circles seem more complex than dots; circles are larger, and are defined by a contour surrounding an interior that is the same colour as the exterior of the item. Dots on the other hand, are so small that they might be conceivably registered as an edge point, and are uniformly coloured, and different from the background. Second, in the concentric task, all the circles were of different sizes, whereas the dots were of a uniform size. Third, the concentric circles all shared a common focus, whereas the dots all had different centre points. I will be trying to establish which of these factors caused Saltzman and Garner's results. Consequently, in this experiment there will be three conditions. See Figure 3-2. In the *Same size*

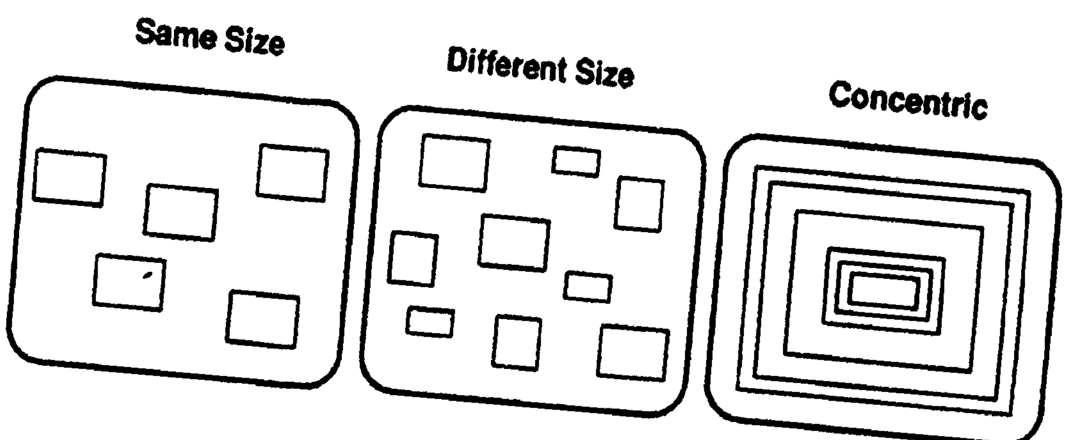

Figure 3-2: Displays for experiment 2

condition subjects were required to count rectangles of the same size spread across the screen. The *Different size* condition was similar except at least one of the rectangles was a different size from the others. Finally, in the *Concentric* condition subjects were required to count concentric rectangles. If Saltzman and Garner's concentric results came about simply because subitizing only occurs for points of light rather than objects, subitizing should not be evident in any of the three conditions. If the result came about because concentric items are necessarily of different sizes, there should be evidence of subitizing in only the *Same size* condition. Finally, if it were the fact that the concentric items by definition had a common centre then subitizing should be evident in all but the Concentric condition

It was expected that subitizing would be evident in *Same size* and *Different size* conditions but not the *Concentric* condition. Why would this be the case? Consider a display of white outline rectangles on a black background. Low level processing would deliver a representation in which illumination discontinuities were grouped into clusters on the basis of the Gestalt grouping principles, primarily proximity in this case. The centre of each cluster would be assigned a place token, and this place token could in turn be FINSTed. When objects are distributed throughout the display, as in the *Same size* and *Different size* condition, these clusters would correctly reflect the number of objects. Edges that were closest typically come from the same object. Thus a FINST could be assigned to each cluster and subitizing could carry on as usual. If the squares were concentric this would not be possible. The edges that were closest together are inevitably from different objects when items are concentric, and moreover, these immediately adjacent edges and corners also have the same orientation. Thus, there would be a tendency to group the wrong contours, on the basis of both proximity and similarity. Moreover, because all of the edges radiated around a common focus, one place token might be assigned to the bunch of them; the dominant impression would be of oneness. A visual routine would be required to properly establish which edges belong to which objects; perhaps boundary tracing would be required. Of couse, this laborious process could be short cut if the subject simply moved the attentional focus outwards from the centre and counted edge crossings, and forgot about trying to resolve each object as a whole. Regardless, subjects would need to move the attentional focus in order to count concentric objects. Consequently, subitizing should not be possible in this case.

Method

Design

There were two factors in this study. The first was display type, that had to do with the kind of stimuli that had to be counted. In the *Same size* condition subjects were required to count rectangles that were all the same size, whereas in the *Different size* condition at least one of the rectangles was a different size from the others. Finally, in the *Concentric* condition the rectangles were one inside another, centred at fixation. The second factor was number. Subjects were required to count 1-8 items. The dependent measure was reaction time. A randomized block factorial design was used.

Subjects

Twelve undergraduate psychology students participated in the study for course credit. Five of the subjects were male. Each subject participated in every condition of the experiment.

Apparatus and Stimuli

As in the previous study, an Apple II+ computer was used to generate the displays and record the data and Gerbrands G1341 voice activated relay were used.

Displays were made up of up to 8 white outline rectangles on a black background There were three types of display. In the *Same size* condition all the items in the display were rectangles of the same size. There were three possible sizes. When subjects were seated 110 cm from the video screen the rectangles subtended 0.26 X 0.16, 0.60 X 0.42, or 1.01 X 0.78 degrees visual angle; the largest rectangle was approximately four times the size of the smallest. Rectangles could be located in any of 24 positions, coded as locations 1-24 in Figure 3-3. Assuming squares of the largest size, the closest horizontal and vertical neighbours were 1.2 and 0.94 degrees from each other. The minimal distance between diagonal neighbours was 0.18 degrees, however. The maximal distance between items was 8.33 degrees for small squares in diagonal corners of the matrix. At

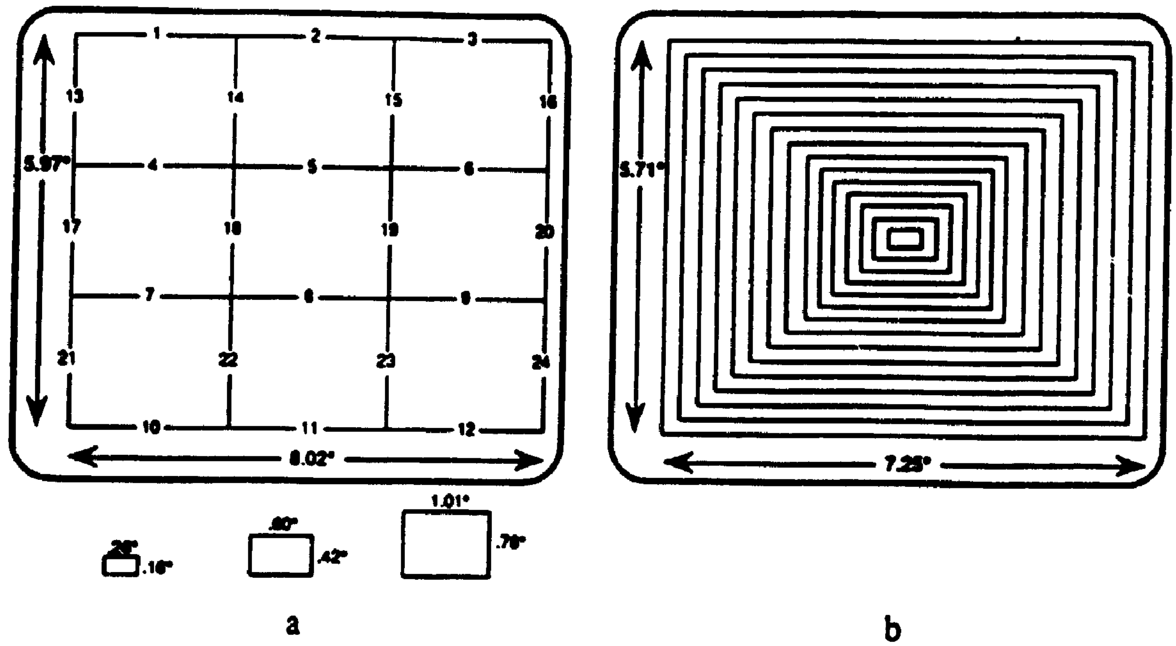

Figure 3-3: Item positions and sizes in experiment 2

most, the entire display would occupy 8.02 X 5.97 degrees visual angle. The size of items and their positions were chosen randomly for each subject and display.

In the *Different size* condition, at least one of the rectangles in the display was different in size from the others. Once again there were three possible sizes and 24 potential item locations. Item sizes and positions were chosen randomly for each display and subject.

Subjects were required to count concentric rectangles centred at fixation in the *Concentric* condition. Rectangles came in 15 sizes, ranging from 0.26 X 0.16 to 7.25 X 5.71 degrees visual angle. The 15 ring sizes are presented in Figure 3-3. For the inner six rings the minimal distance between items was 0.21 degrees horizontal and 0.16 degrees vertical. For the outer rings the distance was made larger because acuity decreases towards the periphery. Thus for the outer rings the minimal distance was 0.29 and 0.21 degrees respectively. The maximal distances between rings was 3.49 horizontal and 2.71 vertical degrees. The sizes of concentric rectangles were chosen randomly for each subject and display.

Procedure

The experiment was conducted in a slightly darkened room. (As in the previous study the illuminance within the room was 135 mW/m^2, as measured by the Techtronix J502 photometer, set to measure incident lighting conditions). Subjects were seated 110 cm from a video screen, with a computer keyboard within easy reach. Their task was to report the total number of rectangles in each display as fast as they could, with accuracy. The latency of their vocal response was measured using the voice activated timer.

Each trial had four phases. First subjects were required to fixate on the central area of a white screen for 608 msec. The computer then beeped to indicate the start of the trial. The counting display came on 256 msec later with up to eight white rectangles. The display remained on the screen until the timer was activated, at which point the screen went white. Fourth, after a pause of 512 msec the subjects were prompted to type in the number they had said or an "X" if something had gone wrong in the trial. The "X" response was reserved for situations in which the timer failed to go off the first time a response was made, or went off before the response was made. These "misfire" trials were readministered at the end of each block.

There were 240 experimental trials. At the beginning of the session subjects were given 24 practise trials.

Results

Data pre-processing

This study was one of the first performed, and as a result, lacks some of the methodological refinements of the later experiments. Specifically, in later experiments error trials were readministered, and in this study they were simply dropped. (Generally, in counting studies error trials are dropped from analysis because of the likelihood that subjects were not actually counting, but guessing or estimating when they make a mistake). Although the average error rate was only 1.95%, the practise of dropping error trials resulted in unequal numbers of cases per cell for the within subject analyses, and as mentioned earlier, this problem coupled with unequal cell variances escalates the probablity of a Type I error for analysis of variance type statistics (Milligan, Wong and Thompson, 1987). Given that I was primarily interested in showing that the *Concentric condition* latencies showed NO significant deviations from linearity, an inflation in the Type I error rate would actually work against my hypothesis, i.e., it would cause trend analysis to register significant deviations from linearity when there was none. Thus, if anything unequal cell sizes would force me to err on the conservative side given my experimental hypotheses. Nonetheless, in order to avoid exacerbating the problem by increasing differences in the proportion of missing cases per cell, latency analyses were performed on raw data for within cell analysis: No outliers were dropped for within subject analysis. It made sense to use the same latencies when calculating averages for each subject for the between subjects analyses.[1]

Analysis of variance

A fixed factors analysis of variance was performed on the averaged latency data. (For means and standard deviations, as well as a discussion of the error rate, see Appendix D). Overall, subjects were fastest at counting objects in the *Same size* condition, and slightly slower when objects were of different sizes. Concentric rectangles took the longest to count, as can be seen from Figure 3-4 ($F(2,22)=103.1$, $p<.001$).

From 2 on there are significant differences between latencies in the three conditions

[1]These analyses were also performed on averages in which the latencies beyond two standard deviations of the mean for each subject, condition, and number were dropped. The results were about the same, as reported in Trick and Pylyshyn (1988).

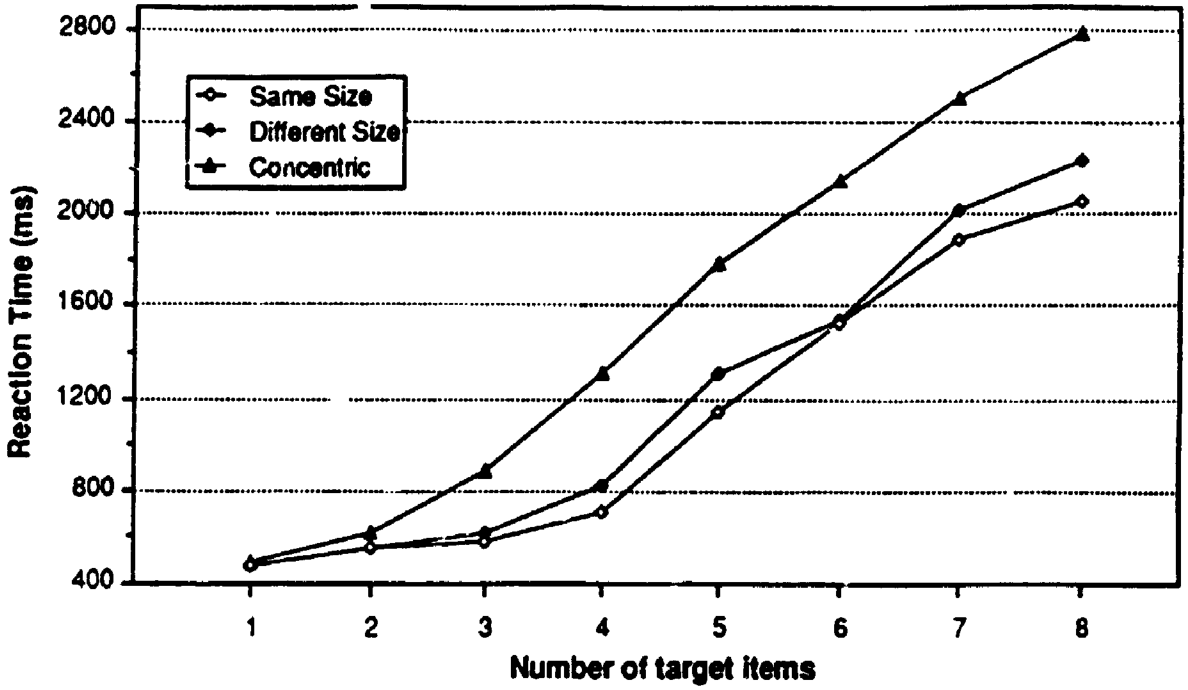

Figure 3-4: Average counting latencies for experiment 2

($F(2,22)=18.4, p<.001$ at 2). Newman Keuls revealed that latencies for the *Concentric* condition are greater than the other two conditions ($p<.05$) when subjects are counting to 2, although there are no significant differences between *Same* and *Different size* latencies. Number yielded its expected effects, with higher numbers of objects taking longer to count ($F(7,77)=276.2$, $p<.001$). Finally, there was also an interaction between display type and number ($F(14,154)=12.6$, $p<.001$).

Trend analysis

Subitizing ranges were determined empirically through trend analysis. First, in order to determine if subitizing occured, trend analysis was performed on the entire range (1-8) to find out if significant non-linear trends emerged.[2] If the reaction time function showed no significant deviation from linearity then it was assumed that subitizing did not occur.

---

[2] Given that there was only one session and none of the subjects had ever been in a counting experiment before, the full range was analyzed because it seemed unlikely that end effects would be a problem.

However, if there were significant deviations from linearity it was necessary to find out where the trend emerged and if it was in the right direction. The point at which the reaction time function first began to show significant non-linear trends was judged to be the boundary of the subitizing range.

Trend analysis on latencies revealed significant linear trends in all conditions. Only the non-concentric conditions showed any significant deviations from linearity, however (non-linear deviation $F(6,88)=6.8$, $p < .0001$ and $F(6,88)=7.5$, $p <.0001$ for *Same size* and *Different size* conditions respectively versus $F(6,88)=1.4, p>.05$ for the non-linear deviation for the *Concentric* condition). Given that non-linear trends are indicative of the change from the subitizing to counting, it would seem that the same enumeration process is being used for both small and large numbers in the *Concentric* condition. In fact, given the magnitude of the latencies it seems probable that the counting process is occuring.

Further analysis of the *Same size* and *Different size* data revealed differences in where the deviation from linearity emerged. For the *Same size* condition the non-linear trend first emerged at 5, indicating subjects on average subitized up to 4 items (non-linear $F(3,55)=13.5, p<.0001$). When items were different sizes on average subjects could only subitize to 3, as can be seen from the Figure 3-4 (non-linear deviation $F(2,44)=6.4, p<.01$).

Averaging across subjects may obscure slope changes in the latencies so consequently analyses were also performed on individual datasets. All subjects had non-linear trends in the in *Same size* and *Different size* conditions. See Table 3-1. Graphs of each subject's latencies are available in Appendix D. Although graphs for individual subjects all show a slight bend at 2, trend analysis revealed significant non-linear trends for only 2 of the 12 subjects in the *Concentric* condition. Consequently, a significantly smaller proportion of the subjects showed evidence of non-linear deviations in the

Table 3-1

Trend analysis of individual datasets for concentric study

NUMBER OF SUBJECTS SHOWING DEVIATIONS FROM LINEARITY

Same size	12/12
Different size	12/12
Concentric	2/12

NUMBER OF SUBJECTS WITH DEVIATIONS FROM LINEARITY

	Same size	Different size	Concentric
Total	(N=12)	(N=12)	(N=2)
# subitizing to 2	1	3	1
# subitizing to 3	3	2	1
# subitizing to 4	7	6	-
# subitizing to 5	1	1	-

Concentric condition than in the *Same size* or *Different size* condition ($\chi^2(2)=27.7, p<.01$)

Further analyses were performed on the individual datasets to ascertain where the deviation from linearity emerged. For both the *Same size* and *Different size* conditions, most subjects subitized to 4. A greater proportion of subjects only subitized 2 or 3 in the *Different size* condition, however. The two subjects that showed evidence of subitizing in the *Concentric condition* were unable to subitize more than 3 in that condition.

Regression analysis

Regression was performed on the averaged data in order to calculate slopes. Although most subjects subitized to 4 when there was evidence of subitizing, subitizing slopes were calculated for the 1-3 range for purposes of comparison. Including latencies for trials in which subitizing did not occur would inflate the subitizing slope, and a greater proportion of subjects in the *Different size* condition only subitized to 3. Thus the slope in *Different size* condition would be disproportionately inflated if the 1-4 range were analyzed. As can be seen from Table 3-2, the slopes for the 1-3 range were 56 and 68 for the *Same* and *Different* size conditions, whereas the slope in the same range was 202 msec/item in the *Concentric condition*, 212 msec/item when the two subitizing subjects were removed. The slope in the *Concentric condition* was significantly greater than the other two, lying outside the 95% confidence intervals for the two conditions.

In contrast, slopes for the 5-8 range in all three conditions fall within 30 msec of each other, with the fastest at 306 msec/item for the *Same size* condition and the slowest at 335 msec/item for the *Concentric* condition. All counting slopes fall within each other's 95% confidence interval. Only in the *Same size* and *Different size* conditions are the slopes in the 1-3 range significantly different than those in the 5-8 range, however.

Table 3-2

Regression analysis for concentric study

Slopes are in milliseconds per item.

SUBITIZING RANGE (1-3)

	Slope	95% C.I.	R
Same size	56	33 - 79	.64
Different size	68	44 - 93	.70
Concentric	202	152 - 253	.81
(N=10)*	212	151 - 273	.80

COUNTING RANGE (5-8)

	Slope	95% C.I.	R
Same size	306	221 - 391	.73
Different size	326	250 - 402	.79
Concentric	335	233 - 436	.70
(N=10)*	310	192 - 427	.65

*Only the data from subjects who showed no evidence of subitizing were included in this analysis.

Discussion

As predicted, subitizing was only evident when items were distributed across the screen. When subjects were required to enumerate concentric rectangles, the slope of the reaction time function was constant and high, suggesting first, that the same process was being used for both small and large numbers of concentric rectangles and second, that the process was the counting process. The results of this study are consistent with Saltzman and Garner's (1948) and moreover show why their results were so different from those of dot enumeration studies. It was the fact that items were concentric, rather than that they were objects of different sizes that produced the constant slope.

The results of this study are consistent with the idea that subitizing is only possible when items can be individuated on the basis of preattentive information. Thus, presumably subitizing was difficult in the *Concentric* condition because moving the attentional focus was required to discover which edge belonged to which object. Low level processing does not deliver the information necessary for enumeration in that condition; grouping on the basis of proximity and similarity will deliver the wrong number of clusters, probably "one" for the centre of the radiating pattern. The tendency to assign a token to the focus of a group would in this case stand in the way of enumeration. Of course, one would have to wonder what signals the subject that attentive processing is necessary. There are several possibilities. First, viewing the concentric rectangles sometimes produced a sensation of depth. In addition, the parallelism of all sides and edges might signal the subject that attentional processing is required. Both of these cues might emerge preattentively and thus would be effective in directing processing.

Subitizing was evident in the *Same size* and *Different size* conditions because low level analysis delivered clusters each of which corresponded to an item. Edges relatively close to each other belonged to the same item, typically. Grouping by similarity was not in evidence because the similar corners were relatively far away from each other, thus proximity cues overode. Because low level grouping processes delivered a number of feature clusters that corresponded to the number of objects, the FINST mechanism could be exploited to accomplish enumeration of small numbers of objects. Consequently, moving the attentional focus from location to location in the proximal stimulus was not necessary in these cases. Enumeration could be accomplished simply by checking the number of assigned reference tokens.

Limitations

It was difficult to design regular displays that could house a respectable number of objects, but would be neither too large, too small, or too predictable. In this experiment there were two problems that evolved because of attempts to keep the display roughly foveal, i.e., no more than 7 or 8 degrees visual angle. The first problem was a partial confound between the condition and the number of item sizes in the display. Ideally, there would have been as many different item sizes in *Different size* trials as there were in *Concentric*; in fact the dimensions of the largest item in the *Different size* condition were only 4 times the dimensions of the smallest whereas in the *Concentric condition* the dimensions of the largest items were approximately 30 times the smallest. The problem was that there was no way to put a lot of large items in the display without having them overlap in the *Different size* condition. Consequently, it is possible that subjects in fact have great difficulty enumerating items of different sizes, but such a small range of sizes were tested that I was unable to tap this difficulty.

Second, there was an unavoidable confound between condition and the distance between contours defining the edges of the rectangles. Of course, in all conditions item sizes and item positions were chosen randomly, so it was possible that contours actually might be farther apart in a given *Concentric* condition trial than a given trial in the other two conditions. Nonetheless, there was a definite crowding problem with concentric displays, and this problem grew worse as the number of items increased. On average adjacent contours were closer together in the *Concentric* condition. As can be seen from Figure 3-3, in the *Concentric condition* the minimum distance between contours was 0.16 degrees horizontal and 0.21 degrees vertical. Thus, in the *Concentric condition* there was approximately a 1/7 (.14) probability that one rectangle would be within 0.16 degrees of the next. Compare, for example, the display configuration in the *Same size* condition. The minimal distance between edges was 0.18 degrees diagonal which occured on the 1/3 of the trials in which the rectangles were of the largest size. There were 24 possible item positions and consequently the probability that two items would be the minimal distance apart when they were the largest of the three sizes was 3/23 on average, given that half of the items had two immediate neighbours and half had four. Therefore, in the *Same size* condition, the probability that two items be within 0.2 degrees was 1/3 * 3/23 (.04). It was much less probable that two neighbouring rectangles be extremely close in the *Same* and *Different size* conditions than the *Concentric* condition.

A distance confound such as this might prove serious given that it might affect low level processing, particularly edge detection. Higher resolution analysis might be required, and the necessary contrast between each item and its background would increase as the density of the items increased. In some counting studies measures are taken to ensure that no item is within .5 degrees of another to prevent such lateral interactions (e.g.,Liss and Reeves, 1983; Sagi and Julesz, 1984), though Atkinson, Campbell and Francis (1976) showed subitizing accuracy does not begin to suffer until items are closer than 0.1 degree. In this study, however, the requirements of keeping the stimuli foveal, and at the same time leaving a wide variety of item positions, made it necessary to have contours close together sometimes, and consequently lateral interference between contours might have occured. There was little evidence that the close proximity of contours from the *same* object interfered with enumeration, however. Recall that there were three possible item sizes in the *Same size* condition. For small rectangles the distance between horizontal edges was 0.16 degrees and the distance between vertical edges was 0.26 degrees whereas for the largest rectangle the distance between horizontal edges was 0.78 degrees and the distance between vertical edges was 1.01 degrees. If simply having contours nearby slowed the process of edge extraction, presumably subjects would be slower at enumerating small rectangles than large, because the contours are closer together, closer even than the suggested 0.5 degree limit (Sagi and Julesz, 1984). When an analysis of variance was performed, comparing latencies for small, medium and large rectangles, there were no significant differences ($F(2,18)=0.4$, $p > .1$). In fact, there was only 11 msec between the average latencies for counting small and large rectangles, and the small rectangles were enumerated *faster*. Consequently, it seems that nearby contours do not impede enumeration if the contours are from the same item. There is also interference from contours belonging on different rectangles to worry about, however. What is needed is a way of equating the distance between items in the different conditions. Otherwise it is difficult to decide if subitizing was not evident in the *Concentric condition* because of lateral interaction between contours, or the concentric configuration of the items.

In order to redress these two problems, a control study was performed. Subjects were required to enumerate non-concentric items that were as crowded together as the corners and sides of the concentric rectangles.

3.2. Experiment 3: Counting parallel lines and parallel corners

In this experiment subjects were required to enumerate either parallel lines or corners. These corners and lines were the same distance apart as the sides of concentric rectangles in the *Concentric condition* in the previous study. If subitizing did not occur in the previous study because of the minimal distance between edges from different items then there should be no evidence of subitizing in this experiment. There would be similar lateral interactions between edges when people had to enumerate parallel lines or parallel corners. In contrast, if subitizing did not occur in the previous experiment because concentric items shared a common focus, then subitizing should occur in this experiment.

Although the proximity and similarity of the parallel items might lead to grouping, at a lower level each item could also be considered as a unit because place tokens could be assigned to different locations. FINSTs could be assigned to these different locations and subitizing could occur as usual.

Method

Subjects

Six subjects participated in the study for payments of $10. Four were male and two were female. All were graduate students or research assistants at the University of Western Ontario, and all but two had been in a counting experiment in the past.

Materials and Stimuli

An Apple II+ computer was used to display the stimuli and record the data while vocal response latencies were measured using a voice activated timer.

The spacing of the stimuli was identical to that of items in the *Concentric condition* in the previous experiment. In this experiment, however, subjects were only shown parallel lines corresponding to one side of the boxes, or parallel right angles corresponding to the corners of the boxes, instead of the concentric boxes. See Figure 3-5. Horizontal parallel lines ranged in length from 0.26 or 7.25 degrees whereas vertical lines ranged from 0.16 to 5.71 degrees visual angle. Each corner was a right angle formed from connecting a 0.42 degree vertical sement with a 0.42 degree horizontal segment. The positions of these corners varied with the corners of the smallest (most central) to the largest (most peripheral) box in the *Concentric condition* in the previous study.

As before, the positions of the items were determined randomly. The top, bottom, left and right sides of the boxes were displayed equally often, as were the top-right, top-left, bottom-right and bottom-left corners. There were up to 9 parallel lines or corners

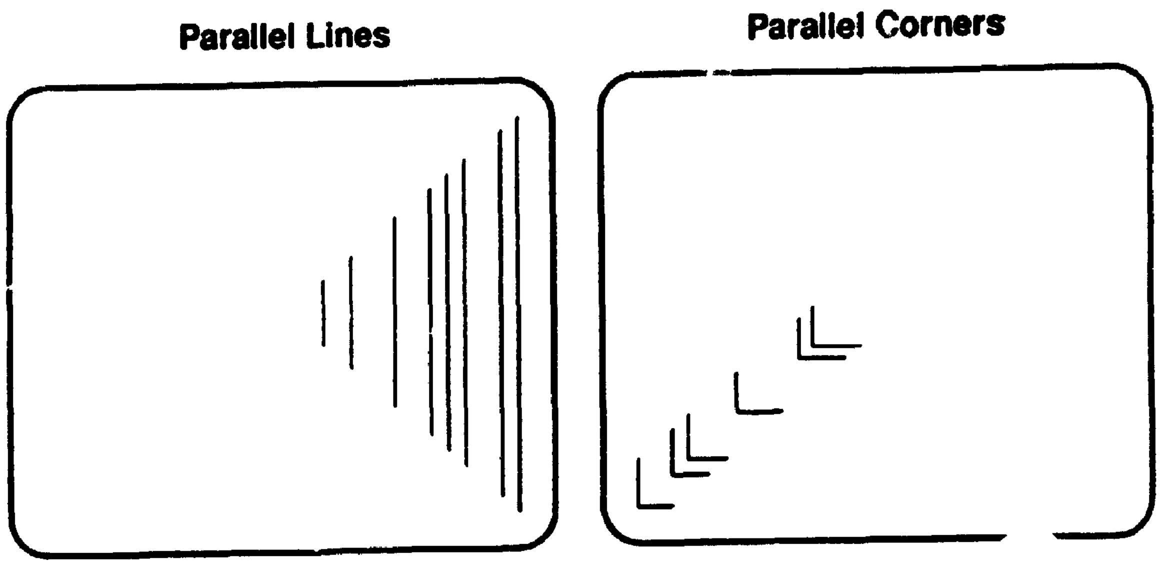

Figure 3-5: Displays for experiment 3

in each display. Only latencies for up to 8 items were analyzed, however. The 9 item displays were used as catch trials to dissipate the end effect typically observed once subjects become aware of the maximum number of items in the display.

A random dot mask with a 0.26 degree square dot in the centre was used as fixation and between trial display.

Procedure

The subjects' task was to say how many items appeared in the display, whether the items were parallel lines or corners. Vocal reaction time was measured. Each trial had four stages. First, subjects were required to fixate on the black square in the centre of a random dot mask. The computer beeped to warn the subjects that the display was about to appear 512 msec before the counting display appeared. Second, the counting display appeared, in which there were either parallel lines or parallel corners. The display remained on until the subject said how many items there were, or made some noise that set off the voice activated timer. Third, for 512 msec a black and white random dot mask was displayed. Finally, the computer prompted the subject to type in the number of items that they had seen, or type in an "X" if something had gone wrong in the trial, such as the timer going off prematurely, or failing to go off when it should. These trials were considered mistrials. If the number that the subject typed in did not correspond to the number of items in the display, the computer beeped five times and printed an error message. Error trials and mistrials were re-administered at the end of each block. The appearance of the re-administered trial was changed by having the items appear on a different side of the screen.

At the beginning of the session subjects were given 20 practise trials to get accustomed to the task. The experimental session required 40 minutes. There were 312 experimental trials to a session.

Results

Data pre-processing

Error trials were readministered until the subjects got them right; consequently there were no error trials to consider. For each subject one outlier was dropped for each number and condition. This outlier represented the latency furthest from the mean in that particular cell for that particular individual.

Analysis of variance

Fixed factors analysis of variance was performed on averaged response latencies of each subject. The type of stimuli had no significant effect on response latencies ($F(1,5)=1.2$, $p > .1$); subjects required approximately as long to enumerate to uniform corners as parallel lines of vastly different sizes. Number had a significant effect on latencies, with higher numbers of items taking longer to count ($F(7,35)=138.8, p < .001$). There was no significant interaction between number of items and condition ($F(7,35)=1.0, p > .1$). See Figure 3-6.

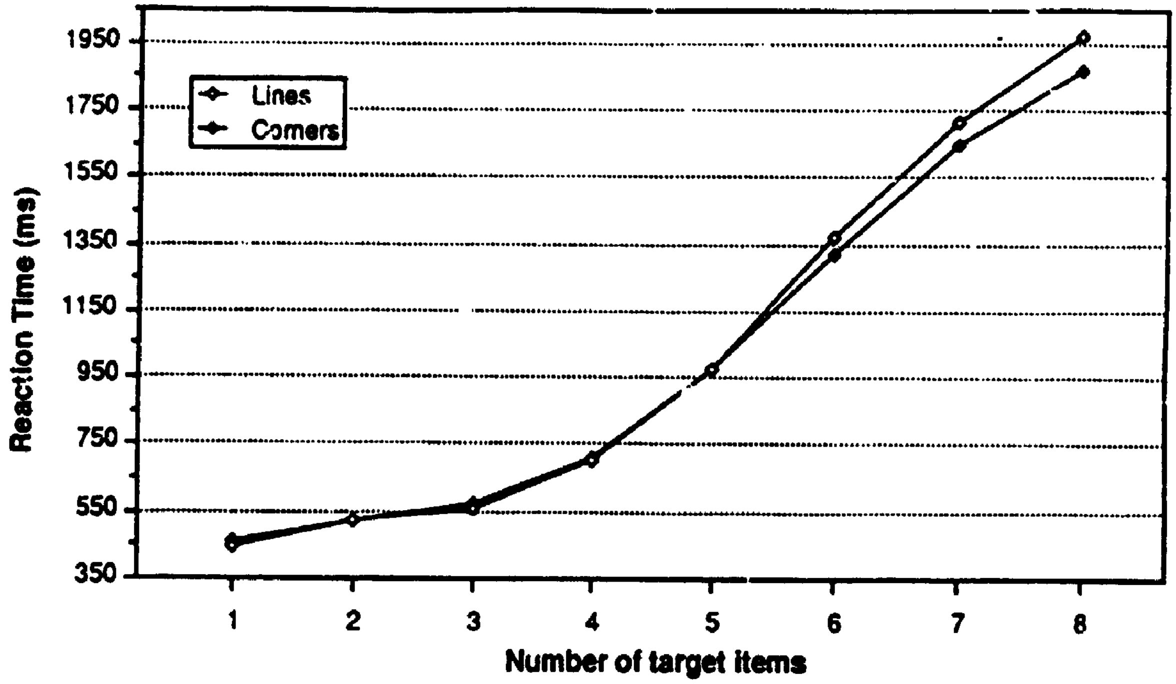

Figure 3-6: Average latency to count parallel lines and parallel corners

Trend analysis

Trend analyses were performed both on the averaged data from all subjects, and on the datasets of each individual subject. There was evidence of a pronouced terminal drop in the data, perhaps because most of the subjects had partipated in a counting study before, and were consequently aware that there were rarely more than 8 items in a display. As a result of this terminal drop only latencies in the 1-7 range were analyzed. (The same analyses were performed on the 1-8 range also, however. It makes little difference to the results, either for the average or within subjects trend analysis). There are significant deviations from linearity in both the parallel lines and parallel corners conditions ($F(5,35)= 12.6, p < .001$; $F(5,35)= 6.3, p < .001$). In both cases, the deviation from linearity occured at 5, indicating that subjects on average, subjects subitize to 4 in the parallel corner and parallel lines conditions ($F(3,25)=6.8, p < .005$; $F(3,25)=6.3, p < .005$).

When analysis of individual datasets is performed a similar picture emerged. In all cases there were significant deviations from linearity observed in the counting latency data. See Appendix D. As can be seen from Table 3-3, most people seem to be subitizing to 2 or 3 in this experiment, but on closer examination of the graphs it seems that this may be due to a plateau between 2 and 3 rather than a sudden increase in the slopes after 2.

Regression

Regression was used to calculate the slopes for the subitizing and counting functions when subjects are counting parallel lines and parallel corners. (See Table 3-4). The subitizing slopes are very similar, 54 msec/item for parallel lines and 56 msec/item for parallel corners in the 1-3 range. Notice that these slopes are almost identical to the 56 msec/item slope for subitizing uniformly sized boxes, in Experiment 2. The slope for the counting function in the 5-7 range is slightly higher for the parallel lines condition than the parallel corner condition, 368 msec/item as opposed to 333 msec/item, though the difference is not statistically significant. In this study, however, slopes in the 5-7 range are slightly higher than comparable slopes for the *Same size* and *Different size* condition in the previous study (306 msec/item and 326 msec/item).

Discussion

There was clear evidence of subitizing when subjects were required to enumerate parallel lines or corners although the edges of different items were as close as the sides of concentric boxes which could not be subitized in the original study. Thus, it seems that the problem in subitizing was peculiar to the task of enumerating concentric boxes. This is not to say that contours from different items weren't grouped: Parallel contours were grouped in this study. Parallel lines were grouped into quite compelling illusory triangles or trapezoids. Parallel corners were grouped to form illusory arrows. Nonetheless, because the contours could be each assigned a place token that did not occupy the same

Table 3-3

Trend analysis of individual datasets for parallel lines and corners study

NUMBER OF SUBJECTS SHOWING DEVIATIONS FROM LINEARITY

Parallel lines	6/6
Parallel corners	6/6

NUMBER OF SUBJECTS WITH DEVIATIONS FROM LINEARITY

	Parallel lines	Parallel corners
Total	(N=6)	(N=6)
# subitizing to 2	3	2
# subitizing to 3	2	2
# subitizing to 4	1	1
# subitizing to 5	-	1

Table 3-4

Regression analysis for parallel line and corner study

Slopes are in milliseconds per item.

	Slope	95% C.I.	R
SUBITIZING RANGE (1-3)			
Parallel lines	54	18 - 89	.63
Parallel corners	56	20 - 93	.63
COUNTING RANGE (5-7)			
Parallel lines	368	273 - 463	.90
Parallel corners	333	212 - 453	.83

location, and this token could be then assigned a FINST, subitizing went on as usual. Second, it is apparent that subjects were capable of subitizing even though the size of the items to be enumerated varied radically, with the longest line being over 30 times the length of the shortest. In fact, the variety of item sizes didn't seem to impede subitizing in the least; subjects were as capable of subitizing different sized lines as uniformly sized corners.

Consequently, it would seem that though subjects are incapable of subitizing concentric rectangles, they can easily subitize the parallel lines and corners that make up concentric rectangles. This leaves us with an interesting question. Given that subjects easily subitize parallel corners or sides, why didn't they simply focus on one side of the display of concentric rectangles, and enumerate lines? There are two possibilities. Either the subjects did not learn strategy of focusing on one side of the display in the *Concentric* condition, or they *could not* focus on one side of the displays, given the layout of the stimuli. It may be quite difficult to focus on one side of the boxes and ignore all the other contours, when all the contours are roughly within the foveal area, or within the attentional focus.[3] Regardless, it is apparent that enumerating small numbers of concentric items is more difficult than enumerating small numbers of items that are laid out across the screen. There is little evidence of subitizing when subjects are enumerating concentric items in these experiments. This might be because low level processing does not deliver clusters of contours that each correspond to an item.

---

[3] Two recent studies suggest that the latter interpretation may be the correct one. See Appendix D for further details.

Chapter Four
Subitizing and search: Treisman's Feature Integration Theory

The goal in this series of studies is to show that when spatial attention is required to distinguish target items from distractor items, subjects can no longer subitize in a field of distractors. According to Treisman's Feature Integration Theory, attention is necessary to combine different parts of an object (e.g., a stem and O to form a Q), and to combine features from different dimensions, (e.g., colour and shape), into a unified object description (Treisman, 1988; Treisman and Gelade, 1980). Search experiments are the empirical backbone of Feature Integration Theory. In search studies subjects are required to find a particular target in a field of distractor items. If the target is present (positive trials) subjects push one key while if the target is absent (negative trials) they push another. Manipulated are the properties of the target and distractors, and the number of items in the display (display size). Reaction time is measured. Treisman typically finds that if target and distractor items differ by a single property such as colour or shape, subjects are capable of finding the target in a time more or less independent of the number of items in the display. Thus, for example, subjects can locate an O in a background of X's in a time independent of the number of items in the display because O's are round and X's are not. Similarly, subjects can find a green O in a field of red O's in a time more or less independent of the number of O's in the display (Treisman and Gelade, 1980) because the target and distractors differ in colour. She calls this phenomenon *pop out*, and interprets these results as showing that properties such as shape (linear vs round) and colour (red vs green) are derived preattentively, or in parallel across the display.

In contrast, if targets and distractors differ only in the combination of parts or features, *pop out* is not apparent; the time required to find a target in a field of distractors begins to increase markedly with the number of items in the display. Consequently, search latency for targets with a subset of the features of the distractors (e.g., an O in a field of Q's) is dependent on display size (Treisman, 1985) because attention is necessary to determine if each item in the display has both an O and a stem.[1] Similarly, if subjects were required to find a red O in a field of green O's and red T's the time to find the target is a function of the number of items in the display (Treisman and Gelade,1980). These results are interpreted as showing that serial processing is necessary in order to combine features or parts into a unified object description. An attentional focus must be moved to each item in turn to search for an item with a subset of the possible parts, or an item with a conjunction of features. Moreover, if reaction time is plotted as a function of display size, and the positive and negative trials are compared, typically the slope of the negative trials is twice that of the positive; it seems that this serial search is self-terminating. In positive trials, in which the target is present, subjects find the target on average half way through the display. Consequently the slope of the positive trials is half that of the negative.

In these experiments I will try to amalgamate Treisman's search studies with an enumeration task. Subjects will be once again enumerating items in a field of distractors. Generally, it is predicted that subjects will be able to subitize targets in a field of distractors, whenever the targets and distractors differ by a property that will produce "pop out" in search. The reason is that it should be possible to assign spatial reference tokens to target items as long as the information that distinguishes targets from distractors is available preattentively, because the FINST mechanism is preattentive. If the targets and distractors do not differ by a property that produces "pop out", subitizing should not be possible because the FINST mechanism cannot be used to selectively individuate

---

[1]In constrast, if the task were simply to indicate whether a stem is present, as would occur if subjects were searching for a Q among O's, *pop out* is once more evident, however. Treisman interpretted this asymmetry as showing that the presence of a particular feature, e.g., the stem in the Q, can serve as a feature whereas the absence of a given feature cannot.

target items. The FINST mechanism is preattentive and the information needed to distinguish targets from distractors presumably requires attention; there should be no evidence of subitizing in this case. Second, I will be manipulating the number of distractors in order to determine the effect of having to screen out irrelevant items. In conditions where target items differ from the distractors by a preattentive feature, there should only be a small cost in latency associated with adding a few distractors. This cost can be likened to Treisman's "Cost of filtering" (Treisman, Kahnemann and Burkell, 1983; Kahneman and Treisman, 1984). According to this idea "pop out" locations compete for the conscious focus; the cost of choosing to report on one "pop out" target item in a field of dissimilar items is small, but increases with the number of distractors. (This cost can be eliminated if subjects know where to focus their attention beforehand). In contrast, in conditions where spatial attention is necessary to distinguish target items from distractors, the cost of adding distractors should be large. The addition of each distractor should noticeably increase counting latencies. This sort of cost might be more properly called the Cost of Feature Integration.

4.1. Experiment 4: Counting O's in distractor letters

Treisman and Gelade (1980) demonstrated that when subjects were required to find an item that differed from the others in terms of a feature, for example, a round O in a field of X's, search time was independent of the number of items in the display: O's *pop out* of X's. In contrast, search latency for an item that has a subset of the features of the others (e.g., an O in a field of Q's) is dependent on the number in the display (Treisman, 1985); O's do not pop out of Q's. This finding has been interpreted as showing that spatial attention is required to combine features at a certain location; simple feature registration is preattentive and parallel, in contrast.

In this study subjects were required to count target items while ignoring a number of distractor items. The purpose was to demonstrate that subitizing is only possible in these cases when the targets "pop out" of the background items. Thus, there are two predictions. First, there should be evidence of subitizing with stimuli that produce display size independent latencies in search; distractors should not prevent subitizing if they differ from the targets on the basis of a feature such as shape. Consequently, subjects should be capable of subitizing O's in a field of X's because O's and X's differ in roundness. The second prediction was that subitizing should not be possible when targets and distractors differed only by the combination of features. Focal attention is assumed necessary to distinguish a stimulus (Q) from one with all but one of its features (O). Because the attentional focus is required to distinguish targets from distractors, subitizing should not be observed.

In addition, the number of distractors was manipulated. If subitizing and search are related, the number of distractors should have only a slight effect on latencies to count 1 item when subjects are enumerating targets that differ from distractors in terms of a feature. (The latency to count 1 is being used instead of the y-intercept because the y-intercept is not very meaningful measure for discontinous functions). Thus, all the counting latencies should be slightly increased when there are X distractors in the display. This prediction follows from the fact that number of distractors has only a small effect on latencies in feature search. In the pilot study documented in Appendix E, slope for positive search trials was approximately 8 msec/item when subjects were searching for an O in a field of X's. In contrast, the number of distractors should have a major effect on latencies when targets possess a subset of the parts that the distractors possess, and thus attentional processing is required. This prediction follows from the finding that the number of distractors has a pronounced effect on the time required to find an O in a field of Q's. (The corresponding slope was approximately 117 msec/item in the pilot study).

Method

Design

There were three factors. The first was the number of targets to count. Subjects were required to count between 1 and 8 O's. The second factor was condition. In the *OX* condition the target letters, O's, popped out of the background letters, X's. In the *OQ* condition the background letters were Q's and consequently the target O's did not pop out. The third factor was number of distractors. There was either 0, 2 or 4 distractor letters.

A randomized block factorial design was used. Condition was blocked so that subjects did the *OX* session on one day and the *OQ* on the other. Session order was counterbalanced. All other factors were randomized.

Subjects

Twelve subjects, eight female, participated for payments of $20. The subjects were graduate students and research assistants from the University of Western Ontario. Eight were experienced in counting studies while four had never participated in experiments of this kind. Each subject participated in every experimental condition.

Apparatus and Stimuli

The stimuli were displays of up to 15 white letters from the standard Apple character set projected on a black background. Letters each occupied 0.70 X 0.40 cm or 0.36 by 0.21 degrees visual angle when viewed from 110 cm. Subjects were required to count letter O's in a background of distractor letters. (See Figure 4-1). There were, in most cases, 1 to 8 O's and either 0, 2 or 4 distractor letters. Forty catch trials were also included in which there were no O's or 9 O's, or 1, 3, or 7 distractor letters. In the *OX* condition the distractor letters were X's; a pilot study documented in Appendix D showed that subjects could search for O's in a background of X's in a time more or less independent of the number of items in the display. In the *OQ* condition the distractor

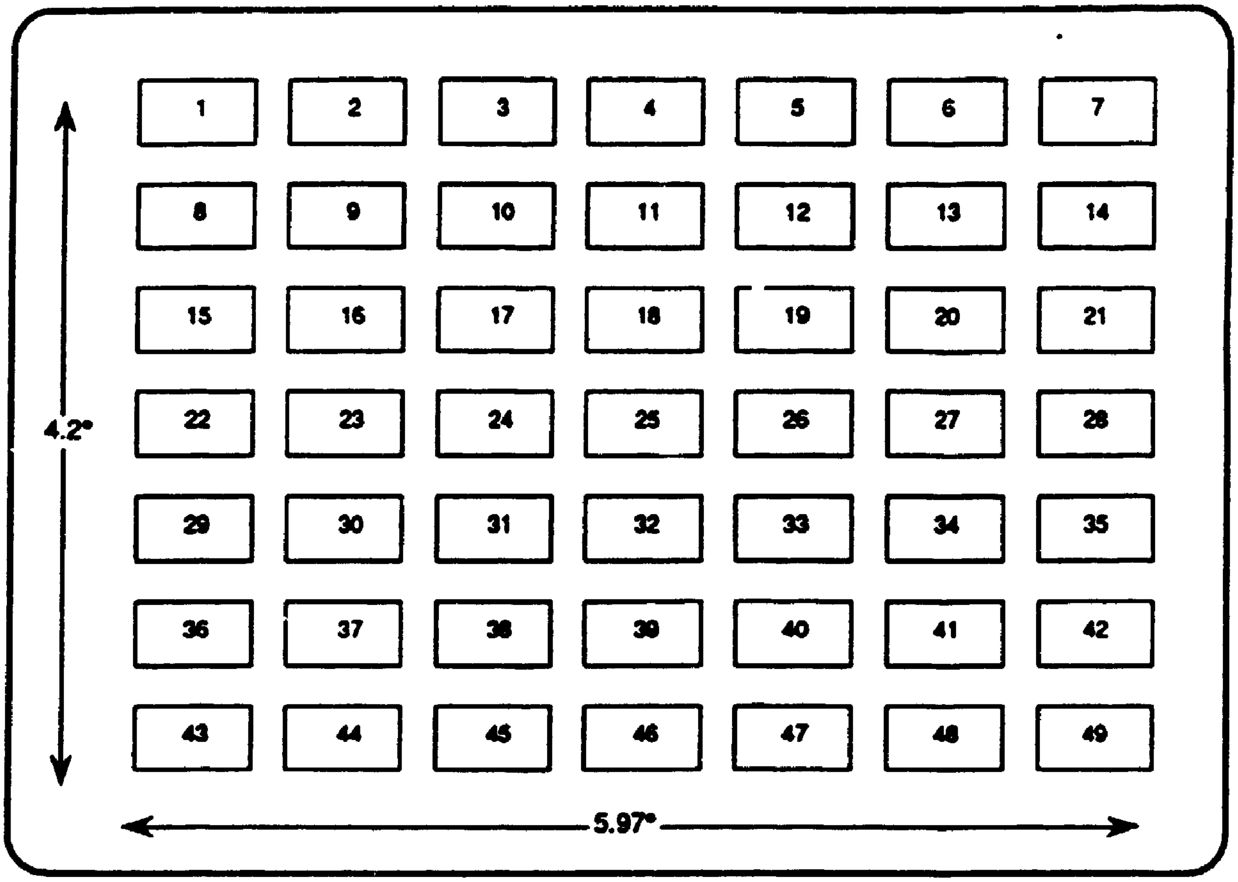

Figure 4-2: Item locations for experiment 4

Procedure

The subjects' task in each of the two sessions was to enumerate the O's in the display. Once again, the experiment was conducted in a slightly darkened room: the illuminance was 135 mW/m^2. Subjects were encouraged to respond as quickly as they could, with accuracy.

Each trial had four stages. First, during the 512 msec pre-trial interval, the screen went white and subjects were required to fixate on a green fixation cross. The computer then beeped to signal the start of the trial. After 256 msec, the letter display appeared and remained on until subjects made a vocal response. Third, as soon as the timer registered the response, the display disappeared and the screen turned white. Finally, after 512 msec subjects were prompted to type in the number that they had said, or an "M" for mistrial if the timer went off prematurely or failed to go off the first time a response was made. These mistrials were readministered at the end of the block. The appearance of the readministered trials was disguised; the positions of items in readministered trials were a mirror image of the original rotated 90 degrees. Accuracy feedback was given after the response was typed in. If the number typed in did not correspond to the number of target items in the display the computer beeped five times and reminded the subject what the target items were.

The experiment was run in two sessions of 328 trials. At the beginning of each session subjects were given 60 practise trials.

Results

Data pre-processing

This study, like experiment 2, was one of the first studies in this series. Consequently, it lacks the methodological refinements of the later experiments; error trials were simply dropped in this study instead of re-administered. The average error rate was only 1.83%, but nonetheless the practise of dropping error trials resulted in unequal numbers of cases per cell for the within subject analyses coupled with the unequal cell variance could inflate the probablity of a Type I error for analysis of variance type statistics (Milligan, Wong and Thompson,1987). Given that most of the errors occured in the *OQ* condition (2.42% as opposed to 1.24%), and given that I was primarily interested in showing that the *OQ* condition latencies showed NO significant deviations from linearity, an increase in the Type I error rate would actually work against my hypothesis, i.e., it would cause trend analysis to register significant deviations from linearity when there was none in the *OQ* condition. Unequal cell sizes would force me to err on the conservative side given my experimental hypotheses, if anything. To avoid aggravating the problem by increasing differences in the proportion of missing cases per cell, latency analyses were performed on raw data for within cell analysis.[2]

Analysis of variance

As can be seen from Figure 4-3, both the number of O's and the number of

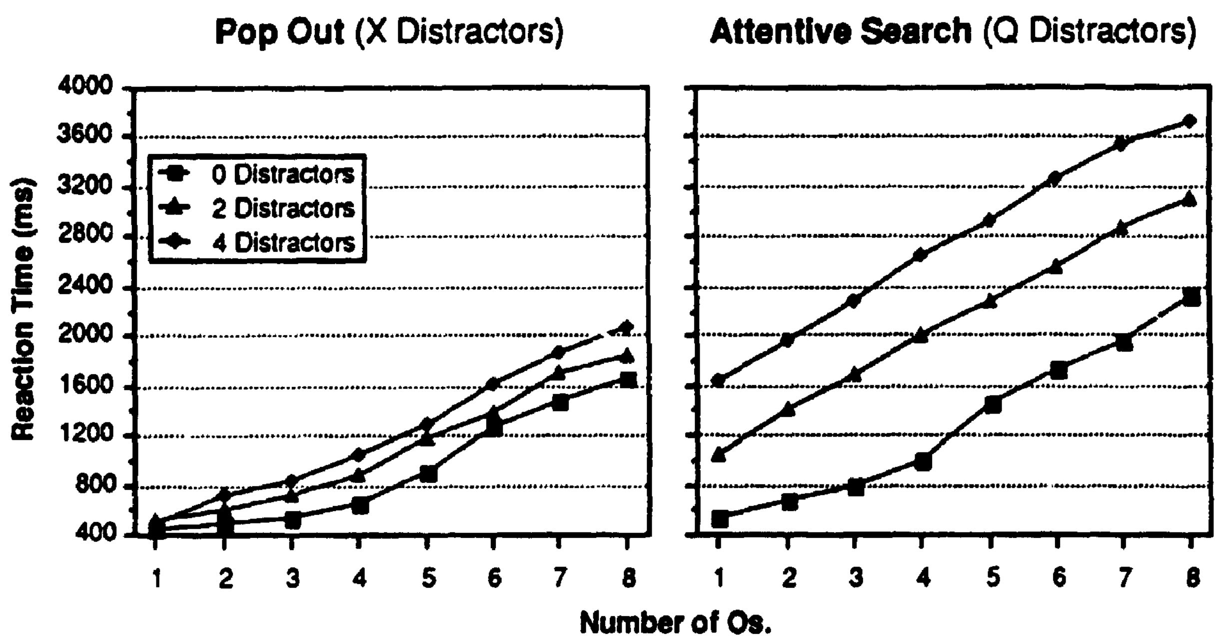

Figure 4-3: Average latency to count O's in distractor letters

distractors affected reaction time. In fact, fixed factors analysis of variance revealed all main effects and interactions to be significant. Overall, counting took longer in the session in which subjects were required to count O's in the background of Q's ($F(1,11)=618.7, p<.001$); subjects were 960 msec slower at counting in the *OQ* condition than the *OX* condition. Reaction times increased as the number of O's increased ($F(7,77)=768.5, p<.001$). These increases were greater in the *OQ* session than the *OX* session ($F(7,77)=57.0, p<.001$), however. The difference between latencies for counting 1

[2]These analyses were also performed on averages in which the latencies beyond two standard deviations of the mean for each subject, condition, and number were dropped (Trick and Pylyshyn, 1989). The results are not qualitatively different.

and 8 objects was 2059 msec on average in the *OQ* condition as opposed to 1339 in the *OX* condition. Further, number of distractors also affected reaction time; the more distractors, the longer it took to count ($F(2,22)=867.6, p<.001$). Once again, though, number of distractors had a greater effect when the distractors were Q's than when they were X's ($F(2,22)=1071.8, p<.001$). It is interesting to compare the effects of increasing the number of distractors on the time to count one object in the two conditions because this situation is analogous to search when the target is present. Every additional distractor adds 250 to 300 msec in the *OQ* condition but only around 36 msec in the *OX* condition. Finally, the deleterious effect of distractors varied according to the number of O's and the type of distractors (F(14,154)=...54,p<.005). For means and standard deviations see Appendix E.

Trend analysis

Linear trends were apparent in every condition in this study. More interesting are the deviations from linearity. First, trend analysis was performed on the 1-7 range to determine whether non-linear trends emerged. Latencies to count to 8 were not included in this analysis because there was evidence of a terminal drop, an effect that often occurs when people have knowledge of the range (Klahr and Wallace, 1976).[3] When there were 2 or 4 Q distractors in the display no significant nonlinear trends were apparent (nonlinear deviation $F(6,88)<1$; nonlinear deviation $F(6,88)<1$ for 2 and 4 distractors respectively). In contrast, when there were 2 or 4 X distractors in the display, subitizing of O's was observed; deviations from linearity were apparent (nonlinear deviation $F(6,88)=3.5, p<.01$; nonlinear deviation $F(6,88)=2.3, p<.01$ for 2 and 4 distractors respectively). In the two distractor condition the deviation from linearity emerged at 5 whereas in the four distractor condition the deviation didn't emerge until 6. Thus, it seems that subjects are capable of subitizing up to 4 O's when there were 2 X's and up to

---

[3] Analyses were also performed on the 1-8 range and are presented in Trick and Pylyshyn (1988).

5 O's when there were 4 X's. From Appendix F, however, it seems that this difference may be an artifact of the higher variability among subjects in the 4X condition. Nonetheless, the results support the idea that subitizing items with distractors in the background is possible when the targets and distractors are discriminable on the basis of a feature such as curvature in this case.

As expected, subitizing was clearly shown whenever there were no distractors. In both the *OX* and *OQ* conditions there were deviations from linearity (nonlinear $F(6,88)=8.5$, $p<.001$; nonlinear $F(6,88)=7.7$, $p<.001$ for *OX* and *OQ* conditions respectively). In both cases the deviation from linearity emerged at 5, indicating that subjects could subitize up to 4 O's when there were no distractors.

In order to ensure that averaging did not disguise deviations from linearity in the slope, analyses of individual datasets were performed. Results were consistent with findings from the averaged data, as can be seen Table 4-1. None of the subjects showed non-linear trends counting O's in a field of Q's. In contrast, most showed non-linear trends when counting O's in X's, 11/12 and 9/12 in the 2X and 4X conditions. Of the subjects who failed to subitize with distractors in the *OX* condition, none failed with both 2 and 4 distractors. Chi square analysis revealed that significantly more subjects showed evidence of subitizing with distractors in the *OX* than the *OQ* condition for both 2 and 4 distractors ($\chi^2(1)=20.3, p<.01$; $\chi^2(1)=14.4, p<.01$ respectively).

There were individual differences in where the deviations from linearity emerged, however; some people subitized larger numbers than others. Although the majority of the subjects subitized to 4, there was some variability, and what's more, this variability increased with the number of distractors.

Table 4-1

Trend analysis of individual datasets for counting O's study

NUMBER OF SUBJECTS SHOWING DEVIATIONS FROM LINEARITY

	OX condition	OQ condition
0 distractors	12/12	10/12
2 distractors	11/12	0/12
4 distractors	9/12	0/12

NUMBER OF SUBJECTS WITH DEVIATIONS FROM LINEARITY

	0 distractors	2 distractors	4 distractors
POP OUT CONDITION: Counting O's in X's.			
Total	(N=12)	(N=11)	(N=9)
# subitizing to 3	3	2	1
# subitizing to 4	8	7	4
# subitizing to 5	1	2	2
# subitizing to 6	-	-	2
ATTENTIVE SEARCH CONDITION: Counting O's in Q's			
Total	(N=10)	(N=0)	(N=0)
# subitizing to 3	1	-	-
# subitizing to 4	8	-	-
# subitizing to 5	1	-	-

Regression

Slopes of the functions were revealed through regression analysis. (See Table 4-2). In order to avoid inflating the subitizing slopes with latencies from trials in which subitizing did not occur, subjects who did not subitize were dropped from the regression analysis.[4] Similarly, a conservative estimate of the subitizing range was employed. Analysis of individual datasets revealed that subjects subitized to 3 at least so the slope was calculated for the 1-3 range although most people subitized to 4 or more. In the *OX* condition there seemed to be a steady increase in the subitizing slope as the number of distractors increased. Each pair of items added to the slope, from 44 to 94 to 108 msec/item, though the addition of the first two distractors had the greatest effect, more than doubling the slope. This effect was not statistically significant however, as the 95% confidence intervals for the slopes overlapped. In contrast, in the *OQ* condition the slopes in the 1-3 range were considerably larger, 323 msec/item for the 2 distractor condition and 314 msec/item for the 4 distractor condition.

Surprisingly, when the subitizing slopes of the two no distractor conditions were compared, there were significant differences, despite the fact that the tasks were identical: Subjects simply had to count O's in an empty background. The subitizing slope was significantly greater when subjects in the *OQ* session than the *OX* session. In the *OX* session the slope in the 1-3 range was 44 msec per item whereas the slope was 129 msec per item in the *OQ* condition when there were no distractors.

As expected, there were few systematic trends in the counting (5-7) range slopes. In all cases, counting slopes fell within each other's 95% confidence intervals, though the slopes in the *OQ* condition were on average slightly higher than the slopes in the *OX* condition.

---

[4] This practise was recommended by Siegler (1987) when faced with subjects that seem to be using different strategies on the same task.

Table 4-2

Regression analysis for counting O's study

Slopes are in milliseconds per item.

	Slope	95% C.I.	R
SUBITIZING RANGE (1-3)			
POP OUT CONDITION: Counting O's in X's			
0 distractors	44	22 - 66	.57
2 distractors	94	61 - 127	.72
4 distractors	108	65 - 150	.72
ATTENTIVE SEARCH: Counting O's in Q's			
0 distractors	129	101 - 156	.87
2 distractors	323	274 - 372	.92
4 distractors	314	257 - 371	.88
COUNTING RANGE (5-7)			
POP OUT CONDITION: Counting O's in X's			
0 distractors	276	190 - 362	.74
2 distractors	262	182 - 343	.75
4 distractors	286	208 - 364	.79
ATTENTIVE SEARCH: Counting O's in Q's			
0 distractors	255	181 - 330	.77
2 distractors	295	211 - 379	.77
(1-8)	293	278 - 309	.97
4 distractors	310	225 - 396	.78
(1-8)	306	289 - 323	.96

Discussion

Given that the presence of non-linear trends in latency is the trademark of the change from subitizing to counting, and given that the slope did not deviate from linearity in the *OQ* condition, it seems that subjects were unable to subitize when enumerating O's in a background of Q's. In search tasks O's do not "pop out" of Q's, and this display size dependent search has been taken by Treisman (1985) as evidence that spatial attention is required to combine features and thus distinguish an item (Q) from another with a subset of its features (O). Consequently, the results of this study are consistent with the idea that subitizing cannot be carried out when the target is a subset of the distractor figure because combining item parts requires spatial attention, and the subitizing process relies on *preattentive* information.

When subjects were counting O's in the background of X's, subitizing was apparent; there were significant differences between the slope in the 1-3 and 5-7 ranges. Subjects can also search for O's in a background of X's in a time independent of the number of X's. This finding is usually interpreted as showing that O's can be distinguished from X's on the basis of the same low level preattentive information that mediates texture segregation (Treisman and Gelade, 1980; Julesz, 1980; Beck, 1982). Consequently, the results of this study are consistent with the idea that subjects can subitize targets in a field of distractors provided the targets and distractors are discriminable on the basis of a feature extracted through low level visual analysis.

The counting process proper was relatively unaffected by the type of distractor item: at least there were no significant differences between slopes in the *OX* and *OQ* conditions. Because enumerating over 4 items requires spatial attention anyway, it would make little difference whether attention was also required to disinguish targets from distractors.

In all conditions there were costs associated with having to filter out distractor items, however. These costs manifest themselves in two ways. First, the latencies to count 1 item were affected. When targets differed from distractors by a pop out feature this cost was small, an added on average 36 msec per distractor when subjects were counting an O in a background of X's, as predicted by the cost of filtering (Treisman, Kahnemann and Burkell, 1983; Kahneman and Treisman, 1984). When spatial attention was required to distinguish targets from distractors in the *OQ* condition, the cost was large, on average 280 msec per distractor. This might be more appropriately called the cost of feature integration. The attentive processing focus had to be moved to each item in turn so that subjects could tell whether the item was an O or a Q. The second way that these costs manifest themselves was in the subitizing slopes in the *OX* condition. Each pair of X distractors added to the slope in the 1-3 range; the first two distractors added 50 msec to the slope, and the next two added another 14 msec. The simple presence of distractors may slow the subitizing process, though the particular number of distractors may be less important.

Condition had an effect on subitizing slopes even when there were no distractors, however. The *OQ* condition had a significantly higher subitizing slope than the *OX* condition even when there were no distractors in the field. There are a number of possible explanations for this finding. First, subjects may haved adopted a cautious strategy in the *OQ* condition because the O-Q discrimination was difficult and error feedback was given. In addition, because the stimuli were small and eye movements might have been required to ensure adequate acuity to make the O-Q discrimination. As a consequence, more eye movements may have occured in the *OQ* condition, and eye movements have been shown to inflate subitizing slopes (Klahr and Wallace, 1976).

In summary, subitizing was not possible under conditions when targets items do not "pop out" of distractors in search, and are thus not easily discriminable. Presumably attention was required to determine whether a given item was a target or distractor in the *OQ* condition, and consequently subitizing was not possible. When attention was not required to distinguish targets from distractors, subitizing was evident. For this reason there were deviations from linearity in the counting latencies when subjects were enumerating O's in X's. This result is consistent with the idea that subitizing relies on a mechanism that individuates the preattentive clusters that are computed by low level processing.

Limitations

Although the present results are compatible with the idea that subitizing is not possible when attentional processing is required to combine parts, there is an alternate interpretation of the data. Distinguishing O's from Q's probably requires more high resolution information than distinguishing O's from X's. (The differences between O's and X's are apparent even after "blurring" your eyes whereas the differences between O's and Q's are not). Consequently, the results may simply suggest that subitizing is not possible when high resolution information is required to distinguish targets from distractors. Therefore, all that can be concluded is that in situations where the targets and distractors are not very discriminable, serial processing is required, and consequently, subitizing is not observed. For this reason another experiment was performed. In it, stimuli varied on two dimensions, colour and orientation, and Treisman's conjunction methodology was used. If the present results replicate with coloured lines, then it seems less likely that the only reason that subjects could not subitize in the *OQ* condition was the need for high resolution analysis.

4.2. Experiment 5: Counting disjunctions and conjunctions of colour and orientation

Subjects were once again required to count items in a field of distractors. In this experiment, stimuli varied on two dimensions, colour and orientation, rather than one (shape) as in the previous experiment. Although the question of what is and what is not a feature is always controversial given that there is no principled way of making the distinction other than empirically, colour and orientation seem on firm ground as separable dimensions on the basis of search and texture segregation studies (Beck, Prazdny and Rosenfeld, 1983). Moreover, there is even evidence that the two properties are processed in different areas of the primate brain (Zeki,1978). Consequently, in this experiment, subjects will be required to enumerate target items of a particular colour and orientation.

There were two conditions in the experiment. In the *Conjunction condition* subjects were required to enumerate items with a particular combination of features (i.e., white vertical lines) in a field of distractors that shared one but not both features with the target (i.e., white horizontal and green vertical lines). In the *Disjunction condition* subjects were required to enumerate targets that differed from background items on the basis of a single feature; thus subjects were required to enumerate white lines in a field of green, or vertical lines in a field of horizontals. Notice that there were two types of target in the *Disjunction condition*, white horizontal or green vertical, and hence the name disjunction. It was important to control for memory requirements given that certain types of memory load interfere with the counting process (Logie and Baddeley, 1987), and consequently a disjunction task was used instead of a simple feature search. That way both "white" and "vertical" had to be kept in mind regardless of condition. Treisman often uses this type of control in her search studies (e.g., Treisman and Gelade,1980).

Given that combining features presumably requires attention, or serial processing of items, it was expected that there should be no evidence of subitizing in the *Conjunction condition*. This prediction follows from the results of the previous experiment, but is also consistent with the idea that subitizing exploits a preattentive item individuation mechanism. It was predicted that subitizing would be possible when subjects were required to enumerate in the *Disjunction condition*, because in any given display the targets and distractors will differ by a feature. Because the information that distinguishes targets from distractors is preattentively derived, it should be possible to individuate only target items, and thus use the item individuation information to perform subitizing. There is already partial support for this claim from another study, although the study differed in that the task was simple feature search, and only a small range was tested, and the total number of items in the display was always less than 6. Francolini and Egeth (1979,1980) demonstrated that latencies for counting 0-3 red letters were independent of the number of black items in the display, even if the items were numbers. The present study is an improvement on the original in that Francolini and Egeth never ventured beyond the subitizing range, and were more interested in category and Stroop effects than counting per se. Consequently their methodology was unnecessarily confusing for the subjects for a simple enumeration study.

Recent studies have called into question several aspects of Treisman's theory. First, evidence is gradually accumulating that subjects don't scrutinize every item in the display, but only the ones with the target colour when performing a conjunction search (Egeth, Virzi and Garbart, 1984; Wolfe, Cave, and Franzel, 1987). Second, there is evidence that the 2:1 ratio of negative to positive search slope is not evident until there are more than 8 items in the display (Pashler,1987). Both of these findings have dictated gradual modifications to Treisman's Feature Integration Theory (Treisman, 1988) though neither is particularly damaging to FINST theory. In fact, the ability to pre-select a subset for attentional access is what is meant by top down assignment of FINSTs. Similarly, given the assumption that FINSTs can be assigned to items that are different from the items in the immediate background, it is not difficult to explain Pashler's (1987) results that suggest we have the ability to scan a homogenous group of items for a conjunction target in parallel, providing that the target differs from the distractors immediately surrounding it by a feature. I am not interested in defending or challenging Feature Integration theory, however, but rather in showing that subitizing and search are linked. I want to show that whenever target and distractor items are so discriminable that the target "pops out" in search, subjects will also be able to subitize the target in a field of distractors. Conversely, if there is evidence of a pronounced slope in the search latencies, and moreover, if the slope for negative trials is twice that of positive, subitizing should no longer be evident.

Method

Design

A randomized block factorial design was used. There were three factors. The first was task; there were two types of search task in this enumeration study. In the *Conjunction condition* subjects were required to enumerate white vertical lines in a field of green vertical and white horizontal lines. In the *Disjunction condition* subjects were either required to count white horizontal lines in a field of green horizontals, or vertical green lines in a field of green horizontals. See Figure 4-4 for examples of the three types of display. The second factor was the number of distractors in the display; there were

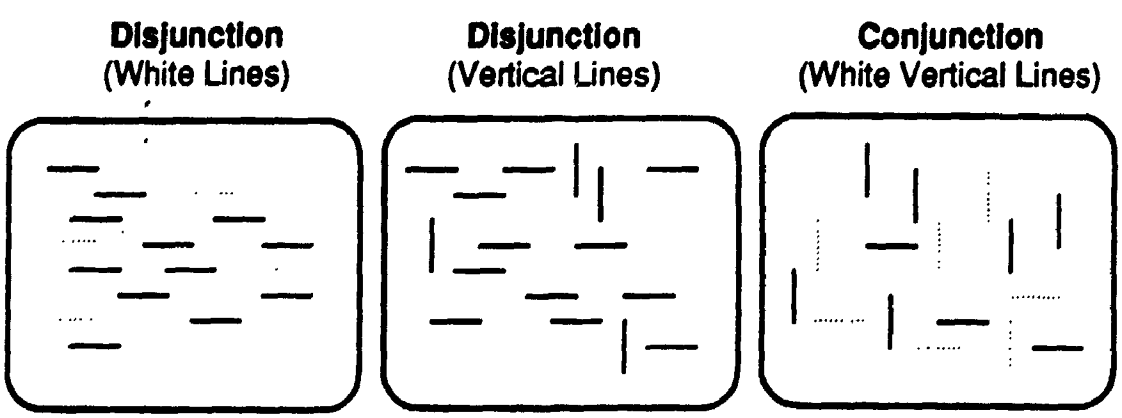

Figure 4-4: Displays for experiment 5

either 0, 12 or 20 distractors.[5] Finally, the number of target items was manipulated. In a given display there was between 1 and 8 items to count.

The task factor was blocked so that subjects performed the *Conjunction* session on one day and the *Disjunction* sessions on the other two days. Half of the subjects started with the *Conjunction* session and half started with the *Disjunction* sessions. Of course, negative transfer between sessions was a problem; the cost seemed minimal in comparison to the additional power bought using a within subjects design, however. Counterbalancing session order at least ensured that negative transfer wasn't confounded with condition, especially given that between condition comparisons were made using only the first of the two *Disjunction* sessions. The second *Disjunction* session was included so that there would be adequate observations to compare latencies for counting white lines with those for counting vertical lines. There was reason to suspect from previous experiments that colour would "pop out" better than orientation in search.

The dependent variable was response latency, or how long it took the subject to say how many target items there were.

Subjects

Ten subjects participated in the three session study for payments of $30. Five of the subjects were female. Subjects were graduate students or research assistants from the University of Western Ontario. Of the 10, 6 had participated in counting experiments before while the remaining 4 were neophytes.

---

[5] A search study was conducted to determine the optimal number of distractors necessary to produce the pattern of results associated with attentive search in the *Conjunction* condition. The results of this study are presented in Appendix F. There was no evidence of the 2:1 ratio of negative to positive slope until there were more than 8 items in the display. Pashler (1987) had similar results.

Apparatus and Stimuli

In this experiment the subjects' task was to enumerate lines of a particular colour or orientation. Each trial involved three displays, a fixation display, a counting display, and a mask. The fixation display was a coloured random dot mask with a 0.5 cm black square in the centre. The subjects were seated 110 cm from the computer screen so consequently the fixation point occupied approximately 0.26 degrees visual angle. Subjects were instructed to fixate on this central black square before the onset of the counting display. The final display, the mask, was a coloured random dot display projected on the screen after the counting display disappeared. The mask was used to prevent afterimages or phosphor decay from effectively increasing the exposure duration of the counting stimuli.

The counting displays were composed of 1-29 line segments. These line segments could be vertical or horizontal, green or white. Each segment was 0.5 cm, or 0.26 degrees visual angle. In each display there could 0-9 target items and 1, 3, 7, 12, or 20 distractors. Trials with 0 or 9 targets or 1, 3, or 7 distractors were used as catch trials, in order to prevent subjects from using item density as a cue to the cardinality of the display. The appearance of the target and distractor items was dictated by the condition. In the *Disjunction* condition trials, targets were either white horizontal lines or green vertical lines, and distractors were green horizontal lines. In *Conjunction* trials the targets were white vertical lines and the distractors were white horizontal or green vertical lines. There were approximately equal numbers of white horizontal and green vertical lines in the display when there were distractors.

At this point two things should become apparent. Although in each condition there were three different types of stimuli, in the *Disjunction* condition there were two types of targets and one type of distractor whereas in the *Conjunction* condition, there were two types of distractor and one type of target. This leads to a sort of double confound. First, there was a confound between condition and the number of types of target item. In the *Disjunction condition* there are two different types of target (white horizontal and green vertical) whereas in the *Conjunction condition* there is one (white vertical). The need to equate memory load between the feature and conjunction conditions by using a disjunction task has already been discussed, however. There was also a confound between condition and the number of types of distractor. In the *Disjunction* condition there is only one kind of distractor, green horizontal lines. In the *Conjunction* condition there are two types: green vertical and white horizontal lines. It would have been desirable to equate the number of types of distractor in the two conditions, but the need to have two types of target in the *Disjunction* condition made it impossible to do so without introducing another colour or orientation into the display. Technical considerations precluded this tactic.[6] Nonetheless, results from an earlier study suggest that people are perfectly capable of subitizing targets in distractors when there are two types of distractor, as long as the distractors and targets differ on a simple feature, a property such as colour or orientation. For example, there was evidence of subitizing when subjects were required to count green items in a background of purple items and white items.

The target and distractor items could fall on any of 40 locations in a matrix. These locations are coded as positions 1-40 in Figure 4-5. The entire matrix occupied 8.5 by 10 cm, or 4.42 by 5.19 degrees visual angle when subjects were seated 110 cm from the display. In general, however, the display would occupy a somewhat smaller area than the full matrix. The minimum horizontal distance between items was 1.5 cm or 0.78 degrees visual angle whereas the minimal vertical distance between items was 1.8 cm or 0.94

---

[6] Apple high resolution graphics has two major problems. First, there is a severe aliasing problem that occurs whenever diagonal or curved lines are drawn. A diagonal line appears to be "dotted" rather than continuous as a result, and I didn't want to introduce line texture as yet another dimension in the discrimination. Second, the available colours were green, white, and purple. The purple and white are hard to distinguish for some reason: purples don't "pop out" easily from white, as shown in experiment 1. Consequently, I was left with vertical and horizontal as my two orientations, and green and white as my two colours.

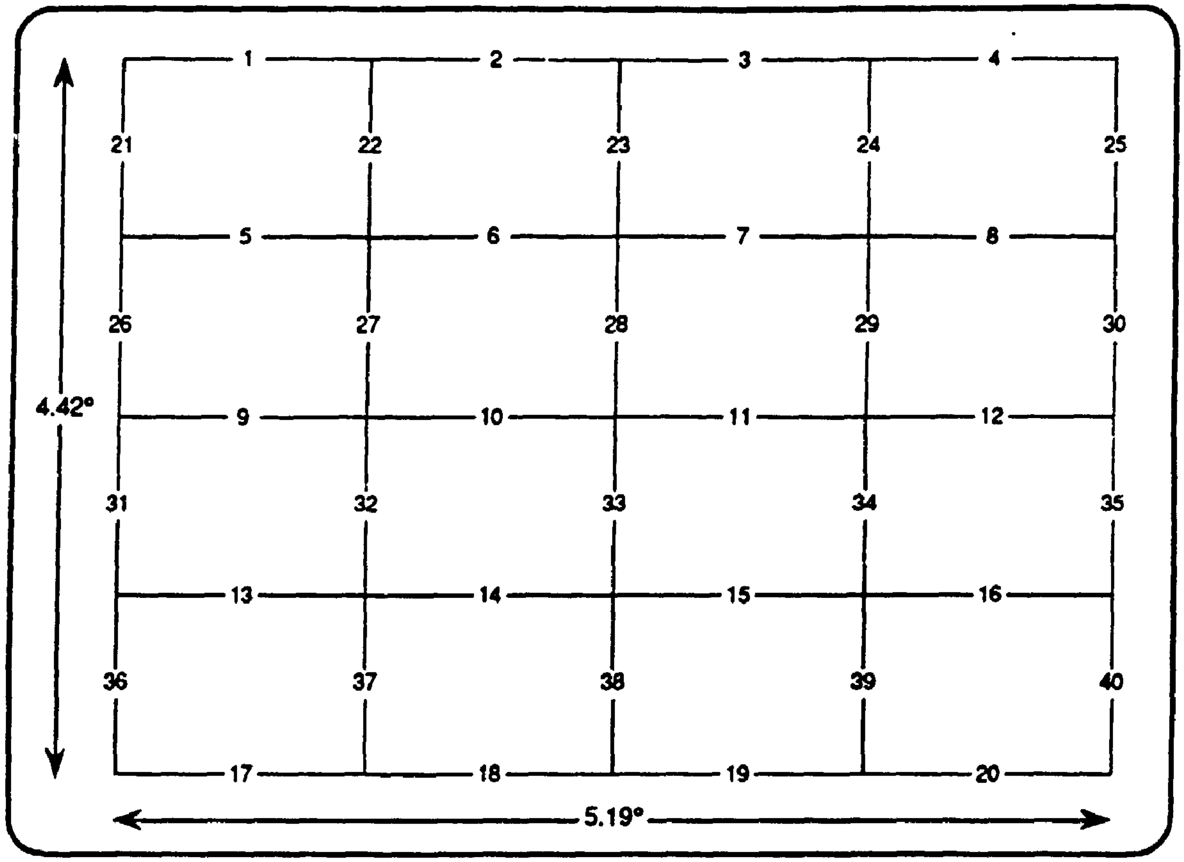

Figure 4-5: Item locations for experiment 5

degrees visual angle. Stimuli could fall within 1.2 cm, or 0.63 degrees, if they were diagonal neighbours, however. The positions of the target and distractor items were chosen randomly for each trial and subject.

Procedure

Subjects were seated in front of a computer terminal in a slightly darkened room. (Illuminance was 135 mW/m^2 as measured by a Techtronix J6502 photometer). Their task was to enumerate certain specified lines as fast as possible, with accuracy.

Each trial had four phases. First, the fixation display was shown for 512 msec. During this time the computer beeped to warn the subject the trial was about to start. The counting display appeared 256 msec later, and remained until the subject made a noise that set off the voice activated timer. The counting latency was considered the time between the onset of the counting display and the time when the voice activated timer registered a sound. Third, a coloured mask was projected for 512 msec. Finally the computer prompted the subject to type in the number that they had said, or an "M" for mistrial if something had gone wrong during the trial.

Subjects were given feedback about the accuracy of their response by the computer. If an error was made the computer beeped five times and reminded the subject what kind of items they were supposed to count. Mistrials and error trials were readministered at the end of the block. The appearance of these readministered trials was disguised, however, by changing the item locations so that the new locations were the mirror image of the old, rotated 90 degrees. Subjects were typically unaware that they had seen these disguised trials previously.

The experiment was run in three sessions, with the last session taking place within a week of the first. In order to minimize confusion, the *Disjunction* sessions were administered on consecutive days. Each session had 328 trials, 288 experimental trials and 40 catch trials. At the beginning of each session subjects did 36 practise trials.

Results

As in the previous studies, the latency data were analyzed in three ways: analysis of variance, trend analysis and regression were performed. Coverage of the various analyses will be quite selective, however. Only analyses with direct relevance for the hypotheses of the experiment will be discussed at any length. Descriptive statistics such as means and standard deviations are presented in Appendix G, as is an account of practise effects. Because the error rate was extremely low, 2% on average, and because some of the errors arose due to mistakes in typing rather than mistakes in counting, accuracy data will not be reported here. See Appendix G for details of the accuracy analyses.

Data pre-processing

The procedure for replacing outliers was the same as the one used in experiments 1 and 3, namely the latency farthest from the mean for each condition, number of distractors, and number of targets, for each subject was dropped from the analysis. There were no error trials to be considered because error trials were readministered until the subject got them right.

Analysis of Variance

Subjects were required to do two sessions of the *Disjunction* condition in order to ensure that there were enough observations that latencies to count vertical items could be compared to those to count white items. Only the first session *Disjunction* trials were analyzed for the comparison between *Disjunction* and *Conjunction* conditions, however. This was done to ensure that subjects were no more practised in the *Disjunction* condition than the *Conjunction* condition. Consequently, for the following analysis, the latencies for counting white and counting vertical trials in session 1 of the *Disjunction condition* are averaged and compared to the latencies to count white vertical items in the *Conjunction* condition. (The results are roughly the same whether only the first session, or both sessions are considered, however, as can be seen from Appendix G).

The average counting latencies for the two conditions are displayed in Figure 4-6. As can be seen from the graph, the number of target items clearly had an effect on latencies in both the *Disjunction* and *Conjunction* conditions, but there is evidence of a small end effects, as revealed by the slight decrease in the slope after 7. Also, it is apparent that the number of distractors also had an effect on latencies, but the effect was most pronounced in the *Conjunction* condition.

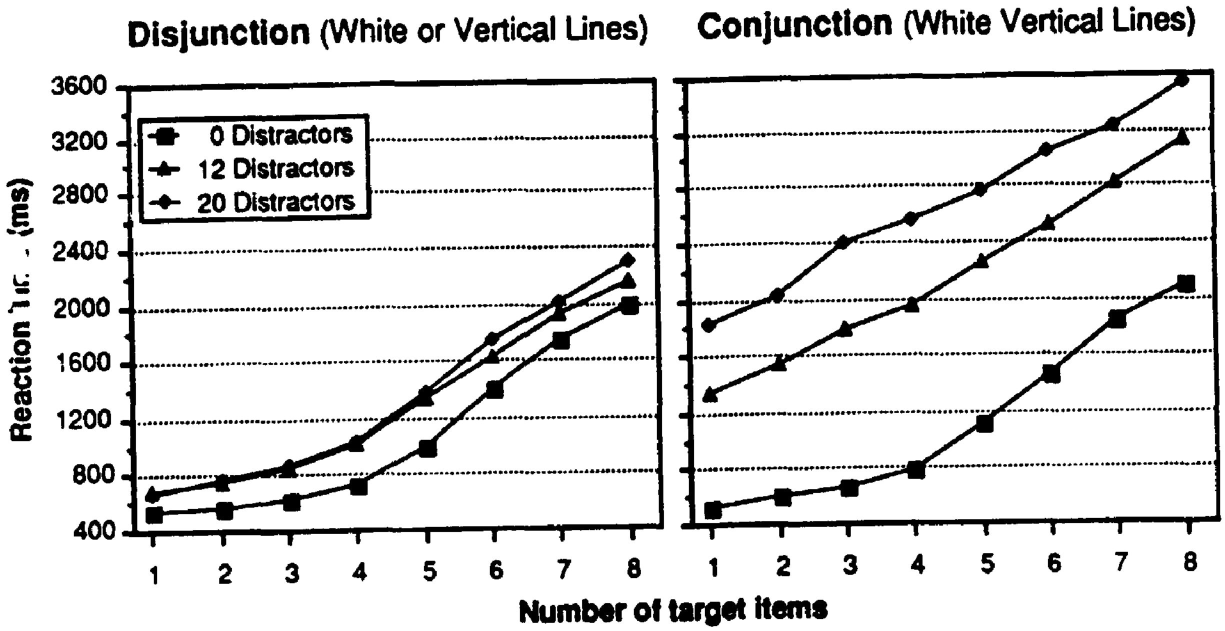

Figure 4-6: Average latencies to count coloured lines

Analysis of variance revealed all main effects and interactions to be significant. As usual, the number of target items had a significant effect; as the number of items to count increased, so did the counting latencies ($F(7,63)=162.4, p<.001$). Task had a significant effect ($F(1,9)=50.4, p<.001$). Subjects were 755 msec slower at counting white vertical lines than they were at counting white or vertical lines. Finally, the number of distractors also had a significant effect ($F(2,18)=87.9, p<.001$); overall, counting times increased as the number of distractor items increased from 1106 to 1734 to 2009 msec in the 0, 12 and 20 distractor conditions respectively.

Interactions between the factors complicate the picture, however. Task interacted with number such that number seemed to have a greater effect when subjects were required to count white vertical lines rather than white or vertical lines ($F(7,63)=2.30$, p

<.05); the difference between latencies for counting 1 and 8 items was 1516 msec on average in the *Disjunction* condition and 1691 msec in the *Conjunction* condition, which may reflect the larger terminal drop in the *Disjunction* condition. The number of distractors also had a much greater effect on latencies in the *Conjunction* than *Disjunction* condition ($F(2,18)=69.6$, $p < .001$). This difference manifests itself even in the latencies to count 1 item. In fact, the first 12 distractors added only 152 msec to the latencies to 1 in the *Disjunction* condition, or an average of 12.6 msec/distractor, and adding another 8 distractors *subtracted* approximately 3.3 msec/item from this value. Given this pattern of results, it may be that simply having distractors interferes with enumeration in the *Disjunction* condition, but the specific number of distractors makes little difference. In the *Conjunction* condition, the addition of the first 12 distractors added 809 msec to the latency to count 1, or approximately 67 msec/distractor whereas the remaining 8 distractors added another 491 msec or 61 msec/item. Consequently, it is apparent that the addition of distractors had not only a much stronger effect in the *Conjunction condition*, but also had a much more uniform effect; each additional distractor added approximately 65 msec to the latencies to count 1.[7] Finally, there was a three way interaction ($F(14,126)=2.4$, $p<.01$), a result of differential effects of the number of distractors on different conditions, depending on the number of targets. (Probably this has to do with the fact that the size of the end effect varies with condition and number of distractors).

Latencies for session 1 and session 2 trials were averaged in order to compare performance when subjects were counting white items as opposed to vertical items in the *Disjunction* condition. (See Figure 4-7).[8] Latencies to count white items and count

---

[7] It is interesting to compare these values to the slopes in the pilot study. Each additional distractor added approximately 3 msec to positive slopes when there were 13-21 items in the Disjunction condition as opposed to approximately 20 msec/item in the Conjunction condition. As in experiment 3, the number of distractors had a greater effect in the counting study than in the search pilot.

[8] Outliers were dropped by dimension in this analysis.

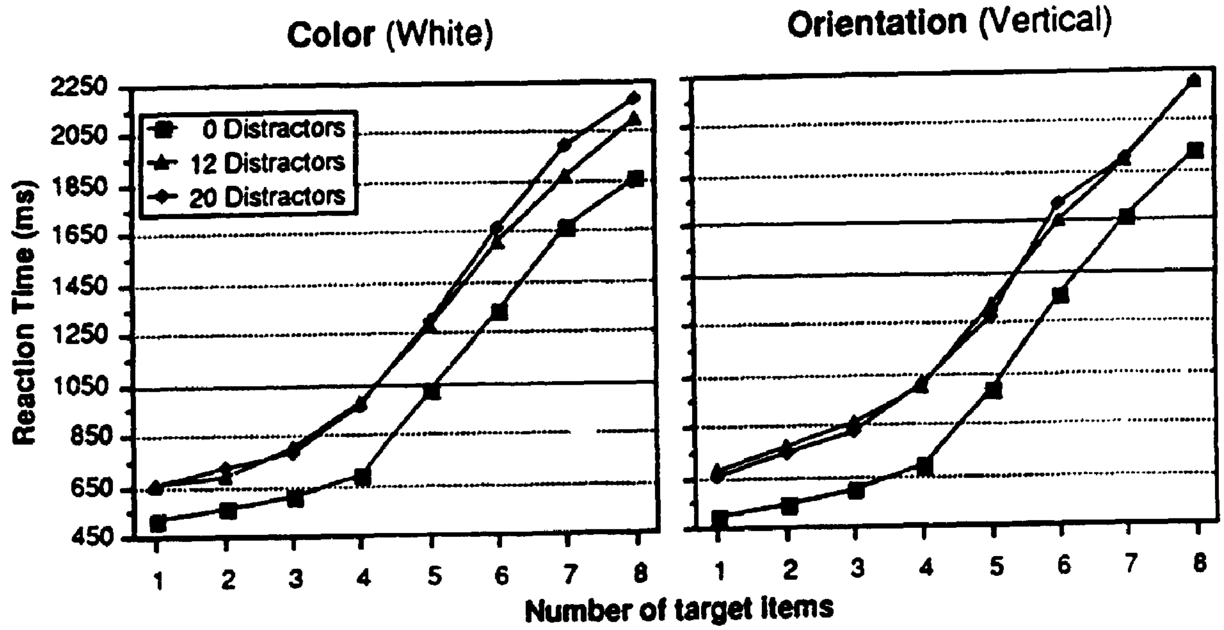

Figure 4-7: Average latencies to count white vs vertical lines

vertical items were not significantly different ($F(1,9)=1.7, p > .1$). Moreover, the particular dimension (white vs. vertical) did not interact with either number of targets ($F(7,63)=1.7$, $p > .1$) or number of distractractors ($F(2,18)=1.6, p > .1$). Nor did dimension enter into a three way interaction ($F(14,126)=0.7, p > .1$). Contary to expectation, the type of stimuli to be enumerated made little difference in the *Disjunction* condition. Subjects were as adept at enumerating white horizontal in green horizontal lines as they were at enumerating green vertical in green horizontal lines.

Trend analysis

In order to determine if subitizing occured, trend analysis was performed. The goal was to discover if there were any deviations from linearity in the latency functions. These deviations are the trademark of the shift from the subitizing process to the counting process. In order to discover if there were any deviations from linearity, latencies for the

1-7 range were analyzed for each condition and number of distractors. (Latencies to count 8 items were not included in the analysis, because of evidence of a terminal drop in latency. In three sessions subjects had ample opportunity to infer that there was rarely more than 8 items to be counted, even with occasional catch trials to 9. This would explain why the latencies to count 8 items were not as high as would be expected). Trend analysis only registers changes in slope, however. In order to find out if the change in slope came about from an increase in slope, rather than a decrease, graphs of the latencies were also consulted.

Trend analysis was performed at two levels. First, between subjects analysis was done. In this type of analysis each case was one subject's average reaction time for a particular condition, and number of targets and distractors. Second, within subjects trend analysis was done. In this type of analysis each subject's dataset was analyzed independently. A case in this type of analysis would be a counting latency for one trial for one subject. Within subject analysis was performed in order to ensure that averaging across subjects did not obscure deviations from linearity produced by the shift from the subitizing to the counting process.

First, between subjects analysis will be discussed. In all conditions, and with every number of distractors, linear trends were evident when latency was plotted as a function of the number of targets. As predicted, however, there were significant deviations from linearity when the task was counting white or vertical lines, as is evident from Figure 4-7. Whether there were 0, 12 or 20 distractors, there was always a deviation from linearity similar to those produced by the change from subitizing to counting in the *Disjunction condition* ($F(5,63) = 7.7$, $p < .001$; $F(5,63) = 2.7$, $p < .05$; $F(5,63) = 3.4$, $p < .01$ respectively). The deviation became significant at 5 in the 0 distractor trials ($F(3,45) = 4.1$, $p < .05$) suggesting that people in general subitize up to 4 items when there were no distractors. The deviation from linearity did not become significant until 6 in the

12 and 20 distractor trials $F(4,54)=3.0$, $p < .05$ and $F(4,54)=4.2$, $p < .01$ respectively), probably an artifact of the greater within cell variability in these conditions. (See Appendix H). In contrast, in the *Conjunction* condition, when the task was counting white vertical lines, subjects only showed deviations from linearity when there were no distractors ($F(5,63)= 5.8$, $p < .001$). As in the 0 distractor *Disjunction condition*, subjects once again seem to subitize to 4 ($F(3,45)=3.7$, $p < .05$). When there were distractors there were no significant deviations from linearity in the reaction time functions in the *Conjunction condition* ($F(5,63)=0.08$, $p > .1$ and $F(5,63)=0.05$, $p > .1$ for the 12 and 20 distractor trials respectively).

The within subject trend analysis revealed a similar pattern of results. The individual data are admittedly noisy, based on a mere 11 reaction times per cell, but nonetheless, a consistent story emerges. (See Table 4-3). Graphs of latencies for each subject are available in Appendix G. Subjects seem capable of subitizing regardless of the condition whenever there are no distractor items. Once there are distractors, however, there are a significantly greater proportion of subjects with deviations from linearity in the *Disjunction* conditions than the *Conjunction* conditions, 9/10 as compared with 0/10 for the 12 distractor case ($\chi^2(1)=16.3, p<.01$), and 9/10 as opposed to 1/10 in the 20 distractor case ($\chi^2(1)=12.8, p<.01$).

An estimate of how high people can subitize can be obtained by observing the point at which the the deviation from linearity emerges. On average, in this study people seem to subitize to 4 or 5, as mentioned previously. When trend analyses of individual datasets are performed, subitizing ranges vary between 2 and 6, however, as can be seen from Table 4-3. There were greater variations between individuals in this study than in any of the others yet performed.

Table 4-3

Trend analysis of individual datasets for coloured lines study

NUMBER OF SUBJECTS SHOWING DEVIATIONS FROM LINEARITY

	0 distractors	12 distractors	20 distractors
Disjunction	10/10	9/10	9/10
Conjunction	10/10	0/10	0/10

NUMBER OF SUBJECTS WITH DEVIATIONS FROM LINEARITY

	0 distractors	12 distractors	20 distractors
DISJUNCTION CONDITION: Counting White OR Vertical Lines			
Total	(N=10)	(N= 9)	(N= 9)
# subitizing to 2	1	1	-
# subitizing to 3	2	2	2
# subitizing to 4	4	3	2
# subitizing to 5	3	2	5
# subitizing to 6	-	1	-
CONJUNCTION CONDITION: Counting White Vertical Lines			
Total	(N=10)	(N= 0)	(N= 1)
# subitizing to 2	1	-	-
# subitizing to 3	1	-	1
# subitizing to 4	6	-	-
# subitizing to 5	2	-	-
# subitizing to 6	-	-	-

Regression analysis

As in the previous studies, the subitizing range was considered 1-3 for purposes of regression in order to avoid inflating the subitizing slope too much with trials in which subitizing did not occur. Similarly, as recommended by Siegler (1987), data from the one subject who did not show evidence of subitizing was dropped from the regression analysis to avoid inflating the subitizing slopes. In addition, the counting range was considered 5-7 in order to avoid deflating the counting slope with the terminal drop, or the decrease in slope after 7 due to the subjects' knowledge of the maximal number of items in the display. See Table 4-4. The first thing to notice is that the data are noisier than that of experiment 3. Generally, the fits of the regression lines are not as good. This may in part reflect greater variability between subjects in this experiment. (This difference was also manifest in the raw latencies, and differences in how high people subitized). In the *Disjunction* condition the subitizing slope gradually increased with the addition of distractors, from 49 to 65 to 92 msec/item, though there were no significant differences between the slopes of the three conditions. In the *Conjunction* condition, there was only evidence of subitizing when there were no distractors; the slope was 60 msec/item, more than the 49 msec/item observed with the same number of distractors in the *Disjunction condition*, and also more than the slopes for subitizing white items (49.2 msec/item) or vertical items (49.5 msec/item) with no distractors. Once again, however, all the subitizing slopes fell within each other's 95% confidence interval. If *Conjunction* condition trials with distractors are considered, the picture is quite different, however. The slope in the 1-3 range is 209 msec/item when there are 12 distractors and 273 msec/item with 20 distractors.

Naturally, in the conditions in which subitizing was evident, the slope in the 5-7 range significantly exceeded the slope in the 1-3 range. The slopes in the 5-7 range did not differ significantly from one another, however. This is not surprising considering the size of the 95% confidence intervals. Notice, however, that when *Disjunction* and

Table 4-4

Regression analysis for coloured lines study

Slopes are in milliseconds per item.

	Slope	95% C.I.	R
SUBITIZING RANGE (1-3)			
DISJUNCTION CONDITION: Counting White OR Vertical Lines			
0 distractors	49	14 - 83	.50
12 distractors	65	22 - 109	.53
20 distractors	92	54 - 130	.71
CONJUNCTION CONDITION: Counting White Vertical Lines			
0 distractors	60	30 - 90	.64
12 distractors	209	70 - 349	.53
20 distractors	273	32 - 514	.42
COUNTING RANGE (5-7)			
DISJUNCTION CONDITION: Counting White OR Vertical Lines			
0 distractors	366	233 - 499	.75
12 distractors	283	164 - 401	.70
20 distractors	326	185 - 467	.69
CONJUNCTION CONDITION: Counting White Vertical Lines			
0 distractors	361	212 - 510	.71
12 distractors	266	-4 - 537	.38
(1-8)	256	206 - 307	.75
20 distractors	225	-115 - 565	.26
(1-8)	238	177 - 300	.66

Conjunction conditions are compared, the slopes in the 5-7 range are almost identical when there are no distractors. Once there are distractors, however, the slope in the 5-7 range is greater in the *Disjunction* condition than *Conjunction* condition, 283 and 326 msec/item as opposed to 266 and 225 msec/item in the 12 and 20 distractor conditions respectively.

Discussion

As predicted, there was only evidence of subitizing in a field of distractors when the target and distractor items differed in terms of a feature, or a property that promotes "pop out" in search (i.e., colour or orientation). When targets and distractors differed only in the combination of features, and moreover, the number of distractors was adequate to produce the typically 2:1 ratio of negative to positive slopes, subitizing was not evident. From this I would like to conclude that subitizing in a field of distractors is only possible when preattentive information distinguishes target from distractor. This finding is consistent with the idea that subitizing exploits a limited capacity preattentive mechanism for item individuation such as the FINST mechanism, although there is no specific support for the FINST hypothesis. Counting slopes are less affected by the differences between target and distractor, because attentional scrutiny is always necessary for the counting process proper.

The results of this study are consistent with the previous study in that both show that under conditions where attentional processing is required to distinguish targets from distractors, subitizing is not evident. Consequently, even when there is no particular need for high resolution processing subitizing disappears with the need for spatial attention. (There is no reason that distinguishing a white vertical line from horizontal and green vertical lines should require more high resolution information than distinguishing green vertical from green horizontal lines). There are a number of other effects in common. Most important, the interaction between condition and the number of distractors is apparent in both studies. Consequently, although increasing the number of distractors increases latencies to count 1 in both conditions, the increase is greatest when attentional scrutiny is required. For example, in experiment 4 each addition X distractor added 36 msec to the time to count one O whereas each additional Q distractor added 250-300 msec. Similarly, in experiment 5 each additional distractor added 12 msec/item for up to 12 items in the *Disjunction condition* but 65 msec/item in the *Conjunction condition*. Therefore, in both experiments 4 and 5 the cost of feature integration dwarfs the cost of filtering; each distractor added more to the processing time if attention was required to distinguish the distractor from the target items to be counted. Nonetheless, the filtering costs are a little higher than those typical in search studies. (In fact, in the associated pilot studies in Appendices E and G the slopes were only 8 msec/distractor and 3 msec/distractor respectively). This difference may reflect differences in the demands of search and enumeration tasks (cf., Francolini and Egeth, 1980).

When there are differences between the experiments 4 and 5, they are primarily in the magnitude of the effects. The number of distractors had a much greater effect on the latencies in experiment 4, when subjects were counting O's. This difference was evident both in latencies to count when the target and distractor differed by a property that promoted "pop out" in search, and when the target and distractor differed by a property that required attentional scruntiny. Consider first the latencies for trials in which the targets were designed to "pop out" of the distractors. In the *OX* condition each additional X added around 36 msec to the time to count an O. In the *Disjunction* condition of this experiment, each additional distractor added approximately 12 msec with up to 12 distractors, and after that the addition of another 8 items *subtracted* 3 msec/item. In conditions in which attentional scruntiny was required, the *OQ* in experiment 4, and the *Conjunction* condition in the present experiment, there was also a consistent pattern. In the *OQ* condition, each additional distractor added 250-300 msec to the latencies to count 1, whereas in the *Conjunction* condition, each distractor only added around 70 msec on average. Even if subjects only checked half of the distractors in the *Conjunction* condition, and thus the slope were truly 140 msec/item there is a great difference in the magnitude of the distractor effect betweem the two experiments. In general, the distractor effects were three or four times as strong in Experiment 4 than Experiment 5, regardless of condition. This difference may reflect differences in stimuli (letters vs coloured lines), differences in the number of distractors (2 and 4 vs 12 and 20), differences in the subjects, or any combination of these. At this point it is hard to tell.

In conclusion, subjects seem incapable of subitizing targets in a field of distractors when the targets and distractors differ only in the combination of features, whereas subitizing seems possible when subjects are enumerating targets that differ from the background on the basis of a feature.

Chapter Five
General Discussion

Although none of these studies is definitive, all converge on a common conclusion. Whenever there is no need for attentional processing, either because the property that distinguishes targets from distractors is a pop out feature, or because low level grouping processes deliver contour clusters which each correspond to an item, subitizing is evident from latency data. Specifically, the slope in the 1-3 range is significantly different from that of the 5-7 range, and moreover, the slope in the 1-3 range is low. Whenever attention is required either to compute a spatial relation, to resolve the object as a whole, or to combine parts or dimensions into a unified object representation, there is little evidence of subitizing from latency data. The slope in the 1-3 range is not significantly different from the slope in the 5-7 range, which suggests that the same process is being used for both ranges. Furthermore, the slope is high; it is the counting process that is being employed for small and large numbers of items.

These findings are consistent with the idea subitizing relies on preattentive information. In order for this preattentive information to be useful there must be some way to individuate the feature clusters derived from low level analysis, however. Pylyshyn's FINST mechanism provides a way to accomplish this task, and moreover, it is independently motivated on functional grounds; there is a need for a mechanism to individuate feature clusters in such a way that the attentional focus could be directed towards them, although the properties and retinal coordinates of the cluster change. These findings are also consistent with the idea that the counting process requires spatial attention. This is why counting and subitizing slopes are similar in cases where spatial attention is required to distinguish target items from distractors. The same process is being employed in the 1-3 and 5-7 ranges, and that process is very time consuming, as is the counting process. Consequently, the difference between subitizing and counting is that the latter requires spatial attention, but the former can carry on without it, because all that is necessary for subitizing is checking the assigned reference tokens or FINSTs (Pylyshyn, 1989).

Consequently, at this point it is possible to re-address the questions set at the outset of the paper. Why are there two enumeration processes? Why can't we subitize any number of items? According to the FINST hypothesis, the reason is that we have a limited capacity mechanism for individuating feature clusters. Although this mechanism permits parallel tracking of a small number of items (cf., Pylyshyn and Storm, 1988), the mechanism is of limited capacity. There are only a small number of spatial reference tokens or FINSTs, consequently, when the number of reference tokens is exceeded a different process is required, the counting process proper. Although these experiments can offer no direct support for the FINST hypothesis, no other theory predicts a relationship between attentional phenomena, spatial relations, and enumeration.[1]

The data speak against a number of alternative explanations for the subitizing phenomena. Why are there two enumeration processes? Why can't we subitize any number of items? One idea is that we can't subitize any number of items because of working memory limitations. We can only hold so many pieces of information in working memory at once. Once there are more items in the display than slots in working memory, a different process is required. Though not stated in these terms, this idea emerged early in the subitizing research with the philosopher Sir William Hamilton (1860), experienced a renaissance in 1956 with Miller, (who finally rejected it), and then

---

[1]Ironically, the subitizing range was originally referred to as the span of attention (Jevons, 1871). At the time "attention" referred to consciousness.

found its way into early work of Klahr (1973b).[2] The enumeration process admittedly requires memory both for storing addition subtotals and keeping track of the marked (already counted) locations. A memory based account does not explain why perceptual factors influence the subitizing slopes, or why subitizing is not apparent with some types of visual stimuli. Why is it more time consuming to enumerate concentric items than items that are side by side? Why is it more difficult to enumerate purple items than green items? Similarly, a memory based account does not explain why subitizing is possible in the presence of some types of distractors and not others. There is no reason that adding Q distractors would affect latencies to count O's any more than adding X distractors.

A second explanation of the subitizing phenomena hinges on the relationship between number and item density, or more accurately spatial frequency. In the world, number is intimately related to density. When there is a large number of items in the visual field, on average the items are more densely packed than when there are few. Why can't we subitize any number of items? Atkinson, Campbell and Francis (1976) suggest that there are special neural units responsive to number when there are small numbers of items. They cite an article by Glazer, Ivanoff and Tscherbach (1973) who found units in the visual cortex of the cat that preferentially responded to black and white bars of a

---

[2] The working memory explanation of subitizing has appeal in that it can encompass both spatial enumeration (e.g., counting dots on a screen) and temporal enumeration (i.e., counting events, tones or flashes of light). Although both tasks involve producing numeric responses, different factors appear to be operative in spatial and temporal enumeration. For example, there is evidence that the subitizing range, as indicated by the range of perfect enumeration accuracy, is smaller for temporal counting than spatial counting. When using Morse code, people often mistake the signal for I (..) with the signal for S (...) and D (-..) for B (-...). This tendency is resistant to corrective training (Taubman, 1950a, cf., 1950b). Thus subjects only seem capable of subitizing 2 identical tokens with perfect accuracy in temporal enumeration, whereas with spatial enumeration subjects typically appear capable of subitizing up to 4. Further, there is evidence that the heterogeneity of items has different effects for temporal and spatial enumeration. Thus, for example, for spatial enumeration the heterogeneity of items has little effect within the subitizing range (Folk, Egeth, and Kwak, 1988; Frick, 1987), and only improves the accuracy of response when there are a large number of items and these heterogeneous items are clustered into homogeneous groups (Beckwith and Restle, 1966). In contrast, for temporal counting, heterogeneity of the items is very important. Thus, when there are several presentations of the same word in rapid visual serial presentation a phenomenon called repetition blindness occurs (Kanwisher, 1987); subjects are less sensitive to the second presentation of the word, and are often unaware that the word has been presented twice, even if several different words intervene between the first and second presentation.

particular spatial frequency and number of cycles. These units were responsive only to low spatial frequencies. From this they concluded that the reason that we can only subitize small numbers is that there are no such units for high spatial frequencies. The problem with this account is that is seems to be aimed at the wrong level of abstraction; we enumerate objects, or more precisely grouping of edges rather than the edge points themselves. Further, it is difficult to extend the explanation from linear arrays of equispaced, equisized dots to randomly arranged O's in displays of X's let alone gamboling holsteins. In addition, it is difficult to explain why subjects were perfectly capable of subitizing when required to count items of a particular colour, but not when required to count connected items though the density of the items was the same in the two cases.

A third explanation of the subitizing phenomena has to do with the relationship between number and geometry. It is a fact of geometry that 1 dot forms a point, 2 dots can be connected by a line, and 3 dots most often fall into a triangular configuration. According to Mandler and Shebo (1982) subjects use the "canonical pattern" of items in the display to determine cardinality when there are small numbers of items. When there are large numbers another process must be used. Why can't we subitize any number of items? Pattern stops being be a reliable cue for number past 3. Although this account has common sense appeal because it seems that we can use pattern to avoid counting items in some situations, there are a number of problems with this explanation. First, it leaves too many questions unanswered. What "canonical" (i.e., rotation, translation and scale invariant) cue differentiates any 3 itemed display from any 4, 5, or 6 item display? It is hard to think of a cue other than number that could distinguish every 3 item display from every possible 4 item display, given the wide variety of items and item configurations.[3] Why is there a subitizing slope at all given that complexity, at least as measured in terms of the number of parts, does not predict recognition latencies (Biederman, 1988)? Does it really take longer to recognize a triangle than a line, a line than a point? In addition, this account does not fit very well with the data. Although geometry stops being a reliable cue to number at 3, many counting studies, including several presented here, show people subitizing to 4 on average. Further, there is no reason that subitizing should be possible with one sort of distractor and not another if simple geometric cues are being used. Finally, subjects can easily and accurately enumerate parallel (linear) arrangements of lines and corners though the geometric cues are all wrong in these cases. (Linear arrangements of items should prompt the answer "two"). For these reasons it seems unlikely that people use geometric arrangement of the stimuli to determine cardinality when there are small numbers of items.

Directions for future research

In this thesis an attempt was made to look at enumeration and spatial attention using a variety of different counting tasks. Consequently, no one task was studied in depth; there are many questions left to answer. First, there is a need to find out why there were different subitizing slopes for different coloured items in the first experiment. Second, there is a need to learn more about the situations in which subitizing clusters of contours is possible. In particular, there is a need to distinguish between the effects of having items one inside another as opposed to concentric, as in experiment 2. Third, there is the relationship between the number of distractors, search and the subitizing process to investigate. The latency to count one item and the subitizing slope both gradually increase with the number of distractors, even if target items "pop out". But what does this mean? In reaction time studies in general it is difficult to distinguish the effects of

---

[3] Subjects were no slower at subitizing 3 items in a "non-canonical" linear configuration than in a "canonical" triangular configuration (Trick, 1987). In fact, subjects were slightly faster at linear configurations. Even Mandler and Shebo (1982) found little benefit of "canonical" pattern in the 1-3 range, even when the same pattern was shown again and again.

perceptual factors from those of decisional factors, or factors relating to how much evidence a subject demands before they make a response. If subjects were inclined to become more cautious when there were a large number of items in the display this might cause an increase in the latencies to count an item in a reaction time study. Moreover, given that eye movements are more likely to occur when there are large numbers of items simply because the items will tend to cover a larger visual angle, and given that eye movements inflate the subitizing slope (Klahr, 1973a), it is possible that the subitizing slope increases with the number of distractors simply because of eye movements. There were good pragmatic reasons for using reaction time as a dependent variable and for not controlling eye movements. Nonetheless, because of these practises it was difficult to interpret changes in the subitizing slope. For this reason these studies need to be replicated using a masking paradigm (e.g., Sagi and Julesz, 1984).

For purposes of the FINST hypothesis there is also a need to discover when the ability to track items in parallel breaks down. If the FINST mechanism is indeed limited capacity then the capacity limitation should be apparent in reduced performance in the multiple target tracking task as the number of targets exceeds the number of FINSTs. From this point, it would also be possible to forge a link between the subitizing research, FINST theory and the research on motion. Ullman's Structure from Motion theorem (1979) *assumes* the ability to track 4 items over 3 frames, for example. In addition, however, there is the question of whether entire areas can be FINSTed, or if only edges or edge clusters can be tokened. Similarly, there is the question of whether different sensory modalities share the same spatial reference tokens, and the question of how spatial information from different sensory modalities (hearing and touch and smell as well as vision) is brought together into a unified description of the world. There are many details to be worked out before a complete understanding of the indexing process can be accomplished.

Conclusion

The goal of this research was to work out a model of subitizing and counting that could be incorporated into a general theory of visual perception (Marr, 1982, Ullman, 1984) and spatial attention (Treisman, 1988). In this paper I have tried to show that subitizing, the rapid apprehension of number when there are fewer than 4 or 5 items, makes use of a preattentive mechanism that individuates selected feature clusters for attentional access (the FINST mechanism, Pylyshyn, 1989). This research contributes to the research on subitizing and counting in two ways. First, I have tried to show that subjects are capable of subitizing target items in a background of distractor items under certain conditions, a finding predicted by none of the current theories of subitizing and counting. Second, I have tried to show that the unit of analysis for subitizing is the contour cluster rather than the edge point or the object-as-whole, also unpredicted by the current theories of subitizing and counting. This research contributes to the research on attention in suggesting that we may have a preattentive mechanism for individuating a small number of items in the visual array, and we can selectively individuate a subset of items on the basis of their features. It is hoped that this information will contribute to a better understanding of how information from the enviroment is synthesized into a representation of the world suitable for the purposes of visual motor coordination and object recognition.

Appendix A
Experiment 1: Counting connected items

A.1. Analysis by session

Average reaction time in milliseconds for each session

Outliers are dropped on the basis of Session. Standard deviations are in parentheses.

COUNTING COLOURED ITEMS

	4 link	5 link	6 link
Session 1			
1	852 (260)	817 (166)	838 (204)
2	905 (207)	909 (216)	881 (149)
3	1019 (250)	1063 (214)	1083 (240)
4	1207 (299)	1274 (226)	1270 (272)
5	1649 (345)	1545 (196)	1570 (304)
6	1823 (290)	1839 (374)	1890 (359)
7	2285 (498)	2165 (323)	2306 (361)
8	2346 (338)	2440 (391)	2357 (323)
Session 2			
1	746 (177)	767 (189)	748 (207)
2	836 (170)	847 (227)	834 (223)
3	921 (233)	918 (207)	1045 (205)
4	1162 (278)	1108 (257)	1157 (238)
5	1411 (263)	1449 (280)	1390 (245)
6	1755 (209)	1636 (255)	1760 (385)
7	1975 (355)	2008 (356)	2067 (420)
8	2132 (396)	2189 (532)	2162 (473)
Average			
1	799	792	793
2	871	878	858
3	970	991	1064
4	1185	1191	1214
5	1530	1497	1480
6	1789	1738	1825
7	2130	2087	2187
8	2239	2314	2260

COUNTING CONNECTED ITEMS

	4 link	5 link	6 link
Session 1			
1	2136 (358)	2437 (487)	2569 (461)
2	2337 (357)	2700 (437)	2824 (443)
3	2474 (291)	2756 (415)	3015 (393)
4	2664 (305)	3067 (393)	3274 (399)
5	2781 (253)	3118 (388)	3479 (525)
6	3099 (371)	3512 (381)	3689 (383)
7	3342 (405)	3643 (456)	4007 (457)
8	3458 (306)	3918 (372)	4125 (444)
Session 2			
1	1924 (321)	2176 (463)	2277 (607)
2	2169 (423)	2478 (428)	2629 (509)
3	2265 (286)	2622 (494)	2787 (546)
4	2486 (446)	2778 (396)	3087 (546)
5	2584 (416)	2967 (526)	3327 (554)
6	2887 (309)	3174 (462)	3467 (547)
7	2979 (361)	3372 (548)	3730 (635)
8	3154 (435)	3483 (481)	3745 (575)
Average			
1	2030	2307	2423
2	2253	2589	2727
3	2370	2689	2901
4	2575	2923	3181
5	2683	3043	3403
6	2993	3343	3578
7	3161	3508	3869
8	3306	3701	3935

Analysis of variance: Full with practise effects

Condition $F(1,7)=168.6$, $p < .001$
Session $F(1,7)=38.6$, $p < .001$
Condition X Session $F(1,7)=2.7$, $p > .1$
Link length $F(2,14)=96.6$, $p < .001$
Condition X Link length $F(2,14)=62.4$, $p < .001$
Session X Link length $F(2,14)=0.2$, $p > .1$
Condition X Session X Link length $F(2,14)=0.2$, $p > .1$
Number $F(7,49)=502.5$, $p < .001$
Condition X Number $F(7,49)=7.8$, $p < .001$
Session X Number $F(7,49)=3.9$, $p < .005$
Condition X Session X Number $F(7,49)=0.8$, $p > .1$
Link length X Number $F(14,98)= 2.6$, $p < .005$
Condition X Link length X Number $F(14,98)=1.6$, $p = .092$

Session X Link length X Number F (14,98)=0.9, p > .1
Number X Dimension F (7,49)=0.3, p < .1
Session X Link length X Number X Condition F (14,98)=0.3, p > .1

Overall analysis adding together Sessions

Condition F (1,7)=223.5, p < .001
Link length F (2,14)=89.3, p < .001
Condition X Link length F (2,14)=66.7, p < .001
Number F (14,98)=171.3, p < .001
Condition X Number F (7,49)=1.3, p > .1
Link length X Number F (14,98)=2.9, p = .001
Condition X Link length X Number F (14,98)=4.4, p < .001

Appendix A (continued)

Regression analyses of practise effects in slope

Reaction time in milliseconds

COLOUR CONDITION

	Overall	Block 1	Block 2
1-3	106.7	109.5	103.9
5-7	316.0	332.0	300.1

CONNECTED CONDITION

	Overall	Block 1	Block 2
1-7	204.6	209.1	200.2

Appendix A (continued)

A.2. Analysis by dimension

Average reaction time in milliseconds for each dimension

Outliers are dropped on the basis of colour or contour orientation depending on the condition. Standard deviations in parentheses.

COUNTING COLOURED ITEMS

	4 link	5 link	6 link
Green items			
1	702 (171)	727 (184)	727 (234)
2	738 (169)	753 (181)	791 (173)
3	851 (193)	866 (201)	902 (197)
4	1054 (265)	1113 (248)	1074 (215)
5	1390 (337)	1332 (240)	1337 (257)
6	1683 (177)	1547 (261)	1671 (312)
7	1987 (355)	1928 (358)	2006 (376)
8	2109 (338)	2183 (277)	2161 (356)
Purple items			
1	876 (227)	851 (192)	857 (201)
2	1024 (220)	1021 (229)	906 (176)
3	1057 (244)	1151 (192)	1224 (208)
4	1317 (272)	1264 (209)	1350 (259)
5	1719 (230)	1727 (295)	1630 (258)
6	1920 (351)	1924 (364)	1958 (348)
7	2358 (550)	2248 (340)	2302 (348)
8	2478 (378)	2369 (438)	2373 (386)
Average			
1	789	789	792
2	881	887	849
3	954	1009	1063
4	1186	1189	1212
5	1555	1530	1484
6	1802	1736	1815
7	2173	2088	2154
8	2294	2276	2267

Analysis of variance: Colour analysis

Link length F (2,14)=0.9, p > .1
Colour F (1,7)=63.9, p < .001
Link length X colour F (2,14)=0.6 , p > .1
Number F (7,49)=344.6, p < .001
Link length X Number (14,98)=1.3, p > .1
Colour X Number F (7,49)=3.5, p < .005
Link length X Colour X Number F (14,98)=1.2, p > .1

COUNTING CONNECTED ITEMS

	4 link	5 link	6 link
Horizontal orientation			
1	2137 (331)	2354 (439)	2502 (481)
2	2327 (304)	2573 (308)	2799 (438)
3	2340 (282)	2716 (431)	3031 (439)
4	2645 (368)	2947 (414)	3212 (496)
5	2770 (377)	3035 (439)	3358 (408)
6	2963 (294)	3417 (442)	3677 (444)
7	3134 (317)	3645 (486)	3822 (388)
8	3264 (425)	3709 (337)	4029 (424)
Vertical orientation			
1	1948 (359)	2265 (484)	2358 (720)
2	2148 (396)	2660 (555)	2625 (498)
3	2377 (393)	2609 (464)	2789 (482)
4	2519 (408)	2876 (406)	3230 (564)
5	2675 (448)	3034 (482)	3353 (530)
6	2957 (281)	3228 (401)	3528 (470)
7	3166 (448)	3350 (574)	3915 (571)
8	3276 (366)	3575 (499)	3751 (463)
Average			
1	2043	2310	2430
2	2238	2617	2712
3	2359	2663	2910
4	2582	2912	3221
5	2723	3035	3356
6	2960	3323	3603
7	3150	3498	3869
8	3270	3642	3890

Analysis of variance: Connected analysis

Link length $F(2,14)=99.2$, $p < .001$
Chain orientation $F(1,7)=2.0$, $p > .1$
Link length X Chain orientation $F(2,14)= 0.7$, $p > .1$
Number $F(7,49)=215.1$, $p < .001$
Link length X Number $F(14,98)=3.3$, $p < .001$
Chain orientation X Number $F(7,49)= 0.8$, $p > .1$
Link length X Chain orientation X Number $F(14,98)= 3.4$, $p < .001$

Horizontal vs Vertical Chain slopes

	Slope
CONNECTED CONDITION (1-7)	
Overall	202.6
Horizontal	197.8
Vertical	207.4

Appendix A (continued)

A.3. Error analyzed by session and dimension

Percent error

Standard deviations are in parentheses. Note that when an error is made, the trial is redone, thus permitting another error to be made on that trial. Also, errors may result from mistakes in typing as well as mistakes in counting.

	4 link	5 link	6 link
COUNTING COLOURED ITEMS			
Session 1			
GREEN items			
1	0 (0)	6.25 (11.57)	0 (0)
2	0 (0)	0 (0)	3.13 (8.84)
3	3.13 (8.84)	0 (0)	0 (0)
4	0 (0)	3.13 (8.84)	0 (0)
5	3.13 (8.84)	3.13 (8.84)	0 (0)
6	6.25 (11.57)	0 (0)	3.13 (8.84)
7	0 (0)	6.25 (11.57)	9.38 (12.94)
8	3.13 (8.84)	0 (0)	12.50 (13.36)
PURPLE items			
1	6.25 (11.57)	0 (0)	3.13 (8.84)
2	0 (0)	0 (0)	3.13 (8.84)
3	0 (0)	6.25 (11.57)	3.13 (8.84)
4	9.38 (18.60)	6.25 (11.57)	6.25 (11.57)
5	3.13 (8.84)	3.13 (8.84)	9.38 (18.60)
6	6.25 (11.57)	12.50 (26.73)	6.25 (17.68)
7	9.38 (18.60)	9.38 (18.60)	15.63 (22.90)
8	3.13 (8.84)	6.25 (11.57)	9.38 (18.60)
Session 2			
GREEN items			
1	6.25 (11.57)	0 (0)	6.25 (11.57)
2	0 (0)	0 (0)	3.13 (8.84)
3	0 (0)	3.13 (8.84)	3.13 (8.84)
4	3.13 (8.84)	0 (0)	0 (0)
5	0 (0)	0 (0)	0 (0)
6	0 (0)	9.38 (18.60)	0 (0)
7	6.25 (11.57)	3.13 (8.84)	6.25 (11.57)

8	3.13 (8.84)	3.13 (8.84)	12.50 (13.36)
PURPLE items			
1	0 (0)	0 (0)	3.13 (8.84)
2	12.50 (18.90)	3.13 (8.84)	0 (0)
3	3.13 (8.84)	6.25 (17.68)	3.13 (8.84)
4	9.38 (18.60)	9.38 (12.94)	6.25 (17.68)
5	6.25 (17.68)	6.25 (11.57)	9.38 (18.60)
6	0 (0)	12.50 (18.90)	6.25 (11.57)
7	12.50 (18.90)	9.38 (12.94)	12.50 (23.15)
8	12.50 (18.90)	9.38 (18.60)	6.25 (17.68)

Analysis of variance: Colour trials

Session F (1,7)=0.1, p > .1
Link length F (2,14)=0.3, p > .1
Session X Link Length F (2,14)=0.2, p > .1
Number F (7,49)=3.2, p < .01
Session X Number F (7,49)=0.1, p > .1
Link length X Number F (14,98)=0.9, p > .1
Session X Link Length X Number F (14,98)=0.7, p > .1
Colour F (1,7)=9.4, p < .05
Session X Colour F (1,7)= 0, p > .1
Link length X Colour F (2,14)=0.4, p > .1
Session X Link length X Colour F (2,14)= 0.4, p > .1
Number X Colour F (7,49)= 1.5, p > .1
Session X Number X Colour F (7,49) = 0.6, p > .1
Link length X Number X Colour F (14,98)= 1.1, p > .1
Session X Link length X Number X Colour F (14,98)= 0.6, p > .1

Appendix A (continued)

Percent error

Standard deviations are in parentheses Note that when an error is made, the trial is redone, thus permitting another error to be made on that trial. Also, errors may result from mistakes in typing as well as mistakes in counting.

COUNTING CONNECTED ITEMS

	4 link	5 link	6 link
Session 1			
HORIZONTAL chained items			
1	0 (0)	0 (0)	0 (0)
2	6.25 (17.68)	3.13 (8.84)	0 (0)
3	3.13 (8.84)	0 (0)	0 (0)
4	6.25 (11.57)	3.13 (8.84)	9.38 (18.60)
5	0 (0)	3.13 (8.84)	6.25 (11.57)
6	3.13 (8.84)	6.25 (11.57)	3.13 (8.84)
7	0 (0)	6.25 (11.57)	9.38 (12.94)
8	12.50 (13.36)	9.38 (12.94)	12.50 (18.90)
VERTICAL chained items			
1	0 (0)	0 (0)	0 (0)
2	0 (0)	6.25 (11.57)	6.25 (11.57)
3	3.13 (8.84)	0 (0)	3.13 (8.84)
4	6.25 (11.57)	6.25 (11.57)	0 (0)
5	3.13 (8.84)	3.13 (8.84)	0 (0)
6	0 (0)	9.38 (12.94)	12.50 (18.90)
7	6.25 (11.57)	0 (0)	3.13 (8.84)
8	3.13 (8.84)	6.25 (11.57)	12.50 (13.36)
Session 2			
HORIZONTAL chained items			
1	0 (0)	0 (0)	0 (0)
2	0 (0)	9.38 (18.60)	0 (0)
3	3.13 (8.84)	0 (0)	0 (0)
4	3.13 (8.84)	6.25 (17.68)	0 (0)
5	0 (0)	0 (0)	0 (0)
6	3.13 (8.84)	12.50 (18.90)	6.25 (11.57)
7	3.13 (8.84)	3.13 (8.84)	0 (0)
8	6.25 (11.57)	3.13 (8.84)	6.25 (17.68)

VERTICAL chained items

1	3.13 (8.84)	0 (0)	0 (0)
2	0 (0)	0 (0)	0 (0)
3	0 (0)	0 (0)	6.25 (11.57)
4	3.13 (8.84)	3.13 (8.84)	0 (0)
5	6.25 (11.57)	0 (0)	0 (0)
6	0 (0)	3.13 (8.84)	9.38 (12.94)
7	3.13 (8.84)	3 13 (8.84)	6.25 (11.57)
8	9.38 (12.94)	6.25 (11.57)	6.25 (11.57)

Analysis of variance: Connected trials

Session $F(1,7)=2.4, p > .1$
Link length $F(2,14)=0.4, p > .1$
Session X Link Length $F(2,14)=0.6, p > .1$
Number $F(7,49)= 6.6, p < .001$
Session X Number $F(7,49)=0.5, p > .1$
Link length X Number $F(14,98)=1.7, p = .06$
Session X Link Length X Number $F(14,98)=0.4, p > .1$
Orientation $F(1,7) =0.1, p < .05$
Session X Orientation $F(1,7)=0.1, p > .1$
Link length X Orientation $F(2,14)=0.6, p > .1$
Session X Link length X Orientation $F(2,14)=2.1, p > .1$
Number X Orientation $F(7,49)=0.3, p > .1$
Session X Number X Orientation $F(7,49) =1.2, p > .1$
Link length X Number X Orientation $F(14,98)=0.9, p > .1$
Session X Link length X Number X Orientation $F(14,98)=1.2, p > .1$

Full error analysis with practise effects

Session $F(1,7)=2.4, p > .1$
Link length $F(2,14)=0.4, p > .1$
Session X Link length $F(2,14)=0.6, p > .1$
Number $F(7,49)=6.6, p < .001$
Session X Number $F(7,49)=0.5, p > .1$
Link length X Number $F(14,98)= 1.7, p = .06$
Session X Link length X Number $F(14,98)= 0.4, p > .1$
Dimension $F(1,7)= 0.1, p > .1$
Session X Dimension $F(1,7)= 0.1, p > .1$
Link length X Dimension $F(2,14)= 0.6, p > .1$
Session X Link length X Dimension $F(2,14)=2.1, p > .1$
Number X DimensionF $(7,49)=0.3, p > .1$
Session X Number X Dimension $F(7,49) = 1.2, p > .1$
Link length X Number X Dimension $F(14,98)= 1.0, p > .1$
Session X Link length X Number X Dimension $F(14,98)= 1.2, p > .1$

Overall analysis adding together sessions
Condition $F(1,7)=1.5, p > .1$

Link length F (2,14)=0.6, p > .1
Condition X Link length F (2,14)=0.14, P > .1
Number F (14,98)=5.9, p < .001
Condition X Number F (7,49)=1.5, p > .1
Link length X Number F (14,98)=1.8, p < .05
Condition X Link length X Number F (14,98)= 0.6, p > .1

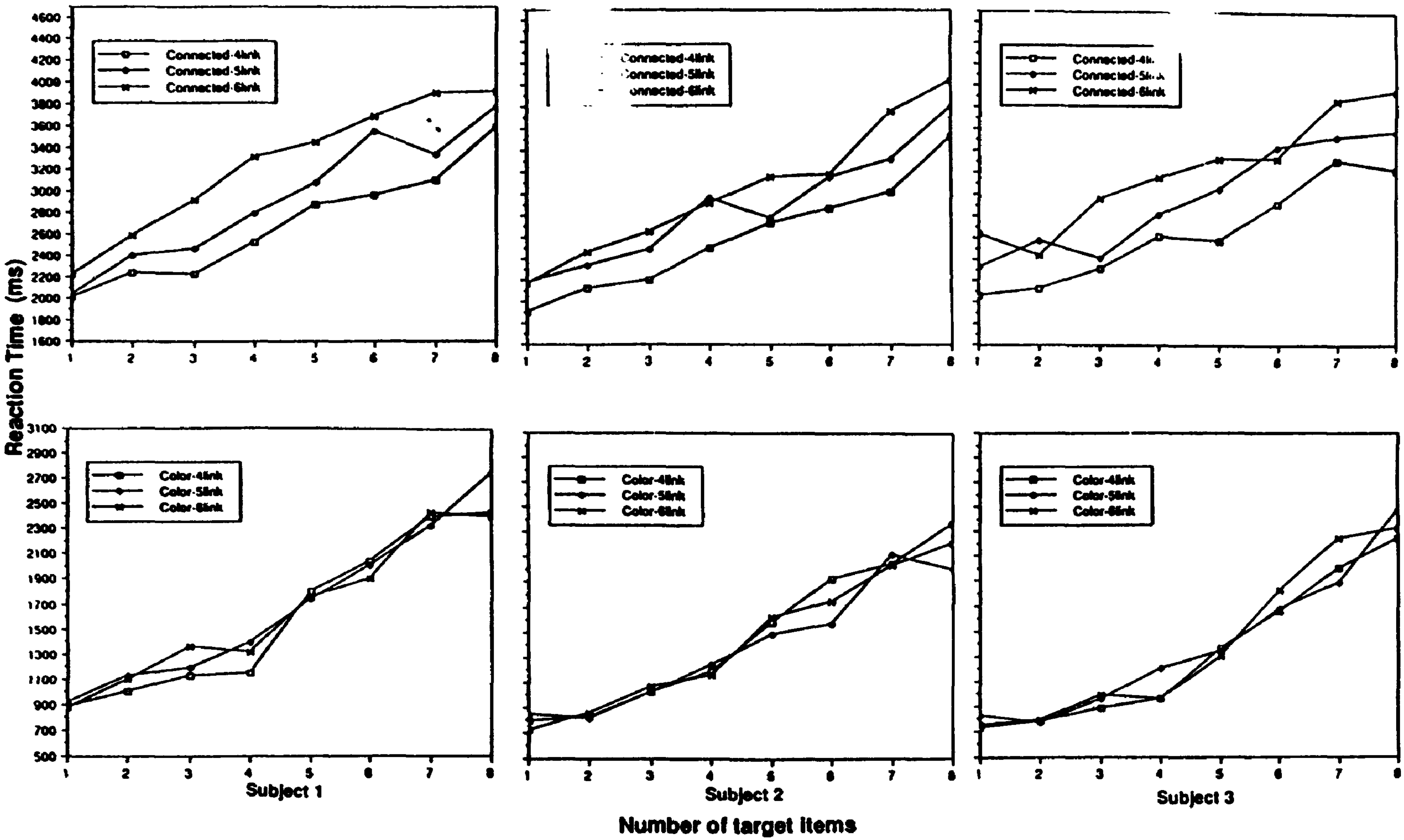

Figure A. Average response latencies for counting connected blocks (top) or colored blocks (bottom) for each subject.

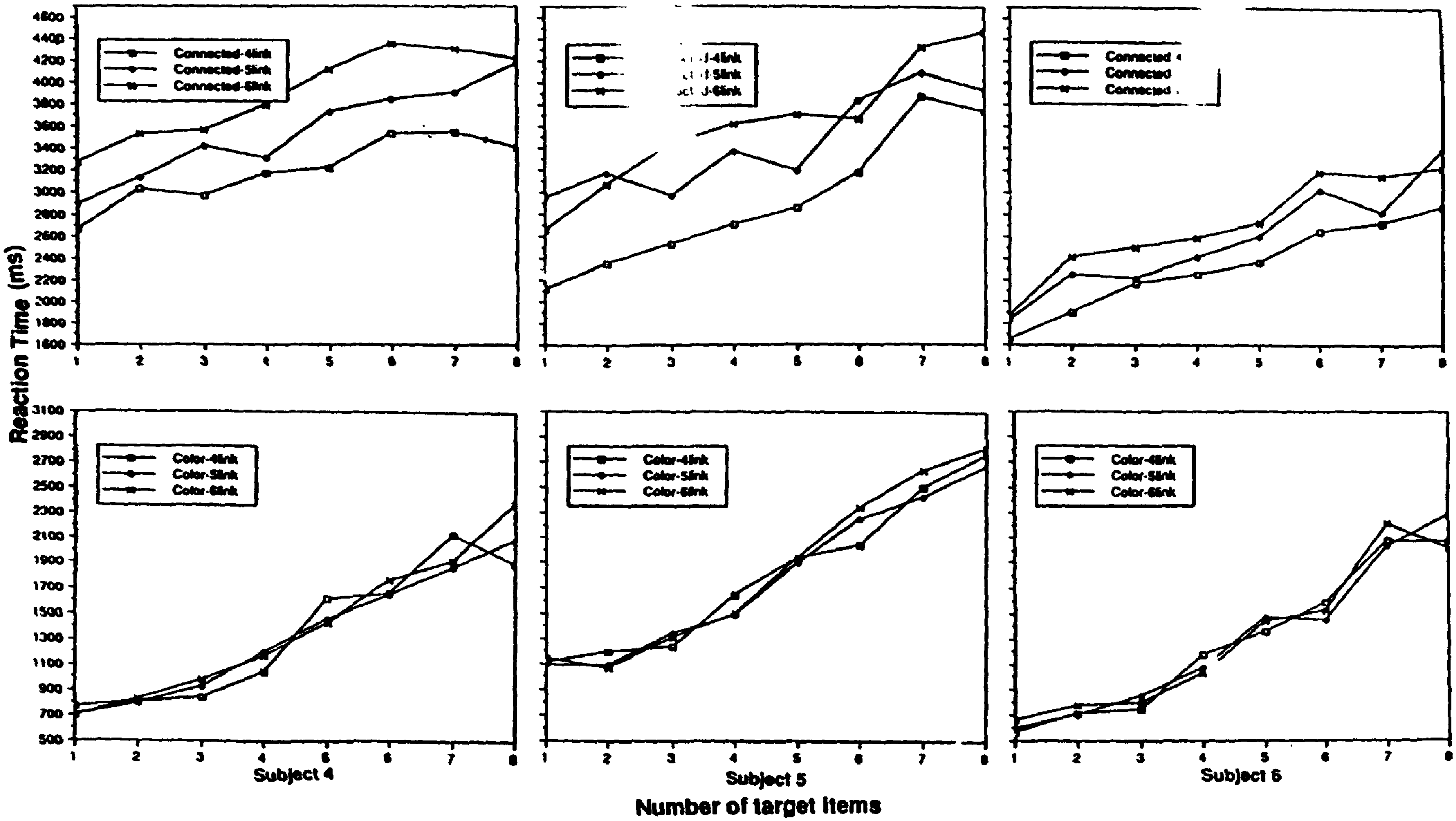

Figure A, continued.

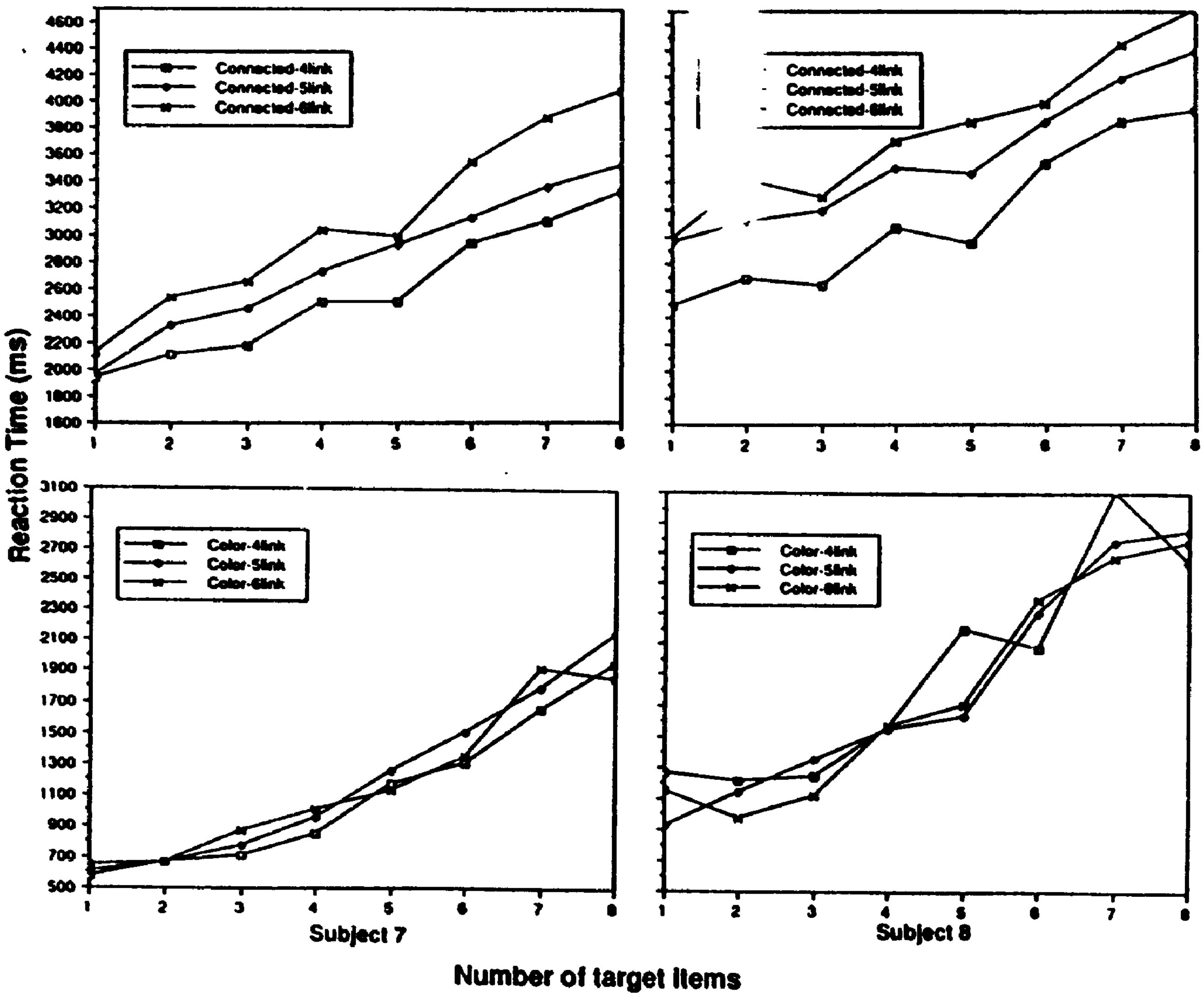

Figure A, continued.

Appendix B
Special analysis of connected trials

The probability that all targets and distractors fall in a smaller area than covered by the contour varies with number of target items and link length, in this experiment. Recall that the position of target items was chosen randomly to fall within a certain link length. There was no requirement that targets be located near the end of the contour, however. Thus, it was possible that for a given 6 link trial, all the targets might fall within the 4 link area. Naturally, this occurence would be more probable if there were a small number of targets. The position of the target makes little difference, however, if there are distractor items to check beyond the last target; attention would presumably be required to determine that each distractor was not, in fact, connected to the targets, and this would involve tracing the chain until a break was found. However, the probability of distractors "beyond" the last target also varies with link length. Recall that the number of possible distractor locations varied inversely with link length, because locations on the unconnected segment of the chain could serve distractor locations also. For 6 link trials there were 42 possible locations for distractors, whereas for five there were 48. These six extra distractor positions fall at points on the contour beyond the first gap. Consequently, the probability that all six distractors fall within four segments in a 6 link trial is greater than for the same situation in a 5 link trial.

An example might clarify this point. Say that the task was counting one item on a horizontal 6 link horizontal contour. Refer back to Figure 2-2. By counting the number of possible target positions in a 6 link contour (42), and then finding out the number of target positions that fall within 4 links (30), it is evident that there is a 30/42 probability that the target will fall in the first 4 links or .714. Similarly, the number of distractor items, or items on vertical contours is 42. The number of distractors that fall within the first 4 links is 28. Thus, the probability that all six distractors lie within a 4 link area is

28/42 * 27/41 * 26/40 * 25/39 * 24/38 * 23/37 = .0718. Together the probability that the target and all distractors fall inside a 4 link area for a 6 link contour is around .05. In contrast, for a 5 link trial, the proability that the target falls within the first 4 links is 30/36 or .833; this probability is greater than for the 6 link trials because there are fewer possible target locations. There are 48 possible distractor locations, 42 on vertical contours plus 6 on the unconnected horizontal contour. The probability that all six items fall within the same area is 28/48 * 27/47 * 26/46 * 25/45 * 24/44 * 23/43 = .032 approximately. Thus, although the probability that the target is in a smaller area than the contour distance is greater in the 5 link than 6 link trials, the joint probability that targets and distractors all fall within this area is smaller, around .03. In fact, as can be seen from the calculations below, if you consider all the ways that targets and distractors could fall short of the full contour length, there are great differences between the 4, 5 and 6 link trials for one target displays. When subjects are counting to 1, there is a 33% chance that items will not occupy the full link area for 6 link trials, as opposed to 3% for 5 link, and effectively 0% for 4 link trials.

PROBABILITY ANALYSIS

Counting one item. Six distractors

6 LINK curves: 7 scan segments if the entire chain is scanned.

a. Probability that subjects will only need to scan two segments

Target in 2 scan segments	Distractors all in 1 scan segment
12/42	7/42 * 6/41 * 5/40 * 4/39 * 3/38 * 2/37

b. Probability that subjects will only need to scan three segments

Target in 3 scan segments	Distractors all in 2 scan segments
18/42	14/42 * 13/41 * 12/40 * 11/39 * 10/38 * 9/37

c. Probability that subjects will only need to scan four segments

Target in 4 scan segments	Distractors all in 3 scan segments
24/42	21/42 * 20/41 * 19/40 * 18/39 * 17/38 * 16/37

d. Probability that subjects will only need to scan five segments

Target in 5 scan segments   Distractors all in 4 scan segments

30/42 28/42 * 27/41 * 26/40 * 25/39 * 24/38 * 23/37

e. Probability that subjects will only scan need to scan six segments

Target in 6 scan segments Distractors all in 5 scan segments
36/42 35/42 * 34/41 * 33/40 * 32/39 * 31/38 * 30/37

Summed probablity .33

5 LINK curves: 6 scan segments if the entire chain is scanned.

a. Probability that subjects will only need to scan two segments

Target in 2 scan segments Distractors all in 1 scan segment
12/36 7/48 * 6/47 * 5/46 * 4/45 * 3/44 * 2/43

b. Probability that subjects will only need to scan three segments

Target in 3 scan segments Distractors all in 2 scan segments
18/36 14/48 * 13/47 * 12'46 * 11/45 * 10/44 * 9/43

c. Probability that subjects will only neea to scan four segments

Target in 4 scan segments Distractors all in 3 scan segments
24/36 21/48 * 20/47 * 19/46 * 18/45 * 17/44 * 16/43

d. Probability that subjects will only need to scan five segments

Target in 5 scan segments Distractors all in 4 scan segments
30/36 28/48 * 27/47 * 26/46 * 25/45 * 24/44 * 23/43

Summed probablity .03

4 LINK curves: 5 scan segments if the entire chain is scanned.

a. Probability that subjects will only need to scan two segments

Target in 2 scan segments Distractors all in 1 scan segment
12/30 7/52 * 6/51 * 5/50 * 4/49 * 3/48 * 2/47

b. Probability that subjects will only need to scan three segments

Target in 3 scan segments Distractors all in 2 scan segments
18/30 14/52 * 13/51 * 12/50 * 11/49 * 10/48 * 9/47

c. Probability that subjects will only need to scan four segments

Target in 4 scan segments Distractors all in 3 scan segments
24/30 21/52 * 20/51 * 19/50 * 18/49 * 17/48 * 16/47

24/30 21/52 * 20/51 * 19/50 * 18/49 * 17/48 * 16/47

Summed probablity .00

Consequently, there is a greater probability that targets and distractors will not occupy the full contour area in the 6 link trials than the 5 or 4 link trials. In these 6 link trials subjects should respond faster than if they had to scan the full 6 link contour. This is because boundary tracing is a time consuming task. If it were possible to avoid boundary tracing part of the contour then it should not take as long to make a response.

In order to test this hypothesis, an analysis of variance was performed putting two distance measures into competition in the *Connected condition*. The first was link length, or the length of contours before a gap. The second was segment area. Segment area was the number of full line segments that had to be scanned in order to pass through all the targets and distractors. See Figure B-1. Notice that for a 5 link horizontal contour there are six horizontal segments. The segment length for a 5 link contour in which targets occupied the full contour area would be six. Similarly, for a 6 link horizontal contour there are seven horizontal segments; a normal 6 link trial would have targets and distractors distributed around all seven segments. If the targets or distractors were all concentrated in the first six horizontal segments, the trial would be a deemed a 6 segment trial, however. Segment area was calculated in a similar way for vertical segment trials, although vertical line segments were counted instead of horizontal segments.

Response latencies were compared between five link trials (i.e., 6 segment) in which the full contour length had to be traversed, six link trials in which subjects could stop short at the same distance as five link trials (i.e., 6 segments) and six link trials in which the full contour distance had to be traversed (i.e., 7 segments). Thus, there were three factors, segment distance, number (1 and 2), and session. Session was analyzed in order to determine if there were any practise effects, i.e., to determine if subjects had to "learn" to short cut contour tracing. If they did, session should interact with segment length in such a way that there would be a larger difference between 6 and 7 segment six link trials in session 2 than session 1.

As can be seen from Table B-1, segment distance did indeed have a significant effect on counting latencies ($F(2,14)=5.8$, $p<.05$). Newman-Keuls test of means revealed that for each number and block there were no significant differences in the means for five link (6 segment) and six link (6 segment) trials ($p>.05$). In almost every case, there were significant differences between the 6 and 7 segment trials, however ($p < .05$). The exception was for trials in session 2 in which subjects were counting 2 targets; the difference between six link-6 segment and six link-7 segment trials fell slightly short of significance. Number and block both had significant effects on latencies in this analysis ($F(1,7)=28.9, p < .005$ and $F(1,7)=8.1$, $p < .05$ respectively) but neither factor interacted with segment length. Thus, even before practise, subjects seem to respond faster when the boundary tracing process could be cut short because targets and distractors did not occupy the entire contour distance. Given that the segment area rather than the contour length per se, seemed to best predict reaction time in these cases, the latencies to count 1 or 2 items probably underestimate the time required to scan a 6 link contour. Once there are more than two target items, the probability that all targets and distractors fall short of the full link length gradually decreases. Consequently, the reaction times for 4 or more targets should be a better reflection of the time required to boundary trace a 6 link contour while counting.

Therefore, to summarize, in the *Connected* condition subjects responded significantly faster when the targets and distractors did not occupy the full length in 6 link contours. Thus, subjects were on average 165 msec faster at counting 1 item when all the targets and distractors were located within the first 6 segments of the 7 segment contour. This result can be taken as evidence that the boundary tracing process stopped

Table B-1

Average response latencies in *connected* condition trials as a function of link length and segment distance

Reaction time in milliseconds. Standard deviations are in parentheses.

	1 item	2 items
Block 1		
5 link-6 segment	2452 (474)	2728 (455)
6 link-6 segment	2440 (705)	2745 (449)
6 link-7 segment	2605 (441)	2874 (429)
Block 2		
5 link-6 segment	2177 (452)	2454 (452)
6 link-6 segment	2228 (647)	2544 (456)
6 link-7 segment	2392 (638)	2659 (558)

short before the end of the contour on these trials. These results can also be taken as support for the FINST hypothesis. Presumably, in the *Connected* condition FINSTs could be of no use in ascertaining the cardinality of the display because FINSTs could not be selectively assigned to target items. (Targets and distractors differed by a spatial relation (connectedness) that required spatial attention). This preattentive mechanism might still play a role in response latencies in that boundary tracing could be stopped short once there were no further FINSTED (target OR distractor) items ahead, however. This finding suggests that there is information about the existence of feature clusters yet to be accessed, before the attentional focus goes to the location in question, there is a sense of "some" ahead though the number name has not been accessed.

If this interpretation were correct, it might even help explain end effects, the effect that produced the drop in slope at 8 in this study. End effects are one of the most persistant and troublesome artifacts in the counting literature. They occur whenever subjects have knowledge of the maximal number of items *n*, and as a result begin to count to *n* almost as fast, or even faster than they count *n-1* items (Klahr and Wallace, 1976).[1] End effects are a great nuisance because latencies for *n* items must consequently be dropped from regression or trend analysis; counting slopes would be smaller than they should be, and quadratic or cubic trends might emerge because of the terminal drop alone if latencies for *n* item trials were included. These end effects persist even when annoying error messages are given to discourage guessing, and even when there are catch trials to *n + 1*, as in this study.

What causes end effects? Say that the maximal number of items in the display is 8. Subjects may be counting and arrive at 6, and still have a sense of other unmarked items ahead, though they have not moved the attentional focus to where the points are, in order

[1] Sometimes these effects even extend to *n-1* so that subjects are counting to *n-1* almost as fast as they can count *n-2*.

to find out how many. In this case subjects may simply respond "8", knowing that there are seldom more than 8 items in the display and that there are "some" yet ahead to be inspected. This sense of "some" precedes knowledge of number, because the number name corresponding to "some" has not yet been accessed. Thus preattentive knowledge of the existence of unmarked (but presumably indexed) items is being used in conjunction with knowledge of the end of range in order to short cut the counting process. This would explain why the time to count *n* items is still as long or slightly longer than the time to count *n-1* items. If subjects were simply guessing, while using a density heuristic, they should be more inaccurate than they are, and much faster. For example, in this experiment item density *per se* should not be a very reliable cue to number, given that there are an unpredictable number of distractor items. If subjects were using simple density they should err more often than they do. As can be seen from Appendix A, the error rates at 8 for the two conditions in which the end effects are strongest, are on average only 9.4% and 10.2% (6 link trials in the *Connected* and *Colour* conditions respectively). Moreover, density information should emerge preattentively; differences in point density in random dot patterns produce effortless texture segregation. Subjects should be able to respond very rapidly to preattentive information, within a second at least. In this study, even with a substantial drop in latency at 8, the latencies are still at least 2 seconds.

In summary, it appears that subjects don't always scan the full contour in order to count connected items. If all the targets and distractors run out before the end of the contour, boundary tracing stops short, and consequently subjects respond faster than they would if they were to scan the full contour distance.

Appendix C
Search pilot for Experiment 1: Comparison of green and purple targets

This pilot study was performed in order to ensure that green items pop out of purple and white items, and purple items pop out of white and green items, as assumed in the rationale for experiment 1. In this experiment subjects were required to search for items of a particular colour, and make one response in a specified target item was in the display and another if it was not. In the *Green* condition subjects were required to search for green items in purple and white items. In the *Purple* condition subjects were required to search for purple items in green and white items. Display size was manipulated. There was between 1 and 21 items in each display. The prediction was that display size would not have a significant effect on response latencies regardless of the colour of the items. This result follows from research that suggests that easily discriminable colours "pop out" from each other in search (Treisman and Gelade, 1980). Second, it was predicted that although subjects would in general require longer to indicate that the specified targets were *not* in the display, the response would not interact with display size, as shown in other search studies when pop out occurs (Treisman and Gelade,1980).

Method

Subjects

Five subjects participated in the study. for payments of $10. Two were female and the remaining three were male. All subjects were graduate students at the University of Western Ontario, and veterans of either counting or search studies. There were two 20 minute sessions in the experiment. Each subject participated in every condition.

Apparatus and Stimuli

In this experiment an Apple II+ computer was used to present the displays and record the data. Subjects were required to indicate whether a target item was in the display by pushing keys on the computer terminal. The "0" key had an silver "Y" pasted over it, and was designated the "Yes" key whereas the "1" key had an silver "N" pasted over it, and was designated the "No" key.

The items were 0.5 cm vertical lines that were either green, white or purple. (The background was black). These lines were located on a 84 point grid identical to the one employed in Experiment 1. When subjects were seated 110 cm from the display, the grid occupied 6.74 by 5.7 degrees, and the minimum horizontal distance between lines on the grid was 1.12 degrees, whereas the minimal vertical distance was 0.95 degrees. The number of items in a display varied between 1 and 21; there was either 1, 13, 17, 21 items in the display. Half the time there was one target item in amongst the distractors and half the time there was no target. Targets and distractors varied according to condition. In the green condition the target items were green lines and the distractors were purple or white lines. Half the distractors were purple and the other half white. In the purple condition the targets were purple and the distractors green or white. In this condition half the distractors were green and the other half white.

A coloured random dot mask preceded and followed the display. In the centre of the mask was a 0 5 cm black square which served as the fixation point.

Procedure

The subjects' task was to indicate whether a specified target item was in the display by pressing keys on a computer terminal. There were two conditions. In the Green condition subjects had to push the yes key if was a green item in the display whereas in the Purple condition subjects pushed the yes key if there was a purple item in the display. Condition was blocked. The order in which the green and purple sessions was counterbalanced.

Trials had two phases. First there was a fixation display. Subjects were shown a coloured random dot mask and were instructed to fixate on a central black square. This fixation display remained for 416 msec. The computer beeped to warn the subjects that the trial was about to start during this time, 256 msec before the onset of the search display. Second, the search display came on. The green, white and purple lines remained on the screen until subjects made a response by hitting a computer key. If an error was made, or subjects hit a key that was neither the "yes" or "no" keys, the computer beeped twice to indicate an error. Trials in which subjects hit a key other than the yes or no keys were re-administered at the end of the block, with changes in the positions of the items

Each session began with 12 practise trials. Each session took 20 minutes, and involved 240 trials. The first session was performed within a day of the last.

Results and Discussion

Error trial latencies were dropped from the analysis, as were latencies beyond two standard deviations of the mean. (This process was carried out for each subject, condition, response, and display size). A fixed factors analysis of variance was then performed on the averaged reaction times.

First an analysis of variance was performed on the entire range of display sizes. Colour had a significant effect on reaction times ($F(1,4)=33.8, p = .005$); subjects were on average approximately 50 msec slower at deciding if there was a purple item in green and white distractors than deciding if there were a green item in purple and white distractors. (See Figure C). Subjects were faster at responding when targets were in the display when the targets were not in the display ($F(1,4)=103.3$, $p = .001$). Display size also had a significant effect on latencies ($F(3,12)=8.0$, $p < .005$), with subjects being faster to respond when there were fewer items in the display. However, the display size effect did not interact with response as would be expected if serial self terminating search was

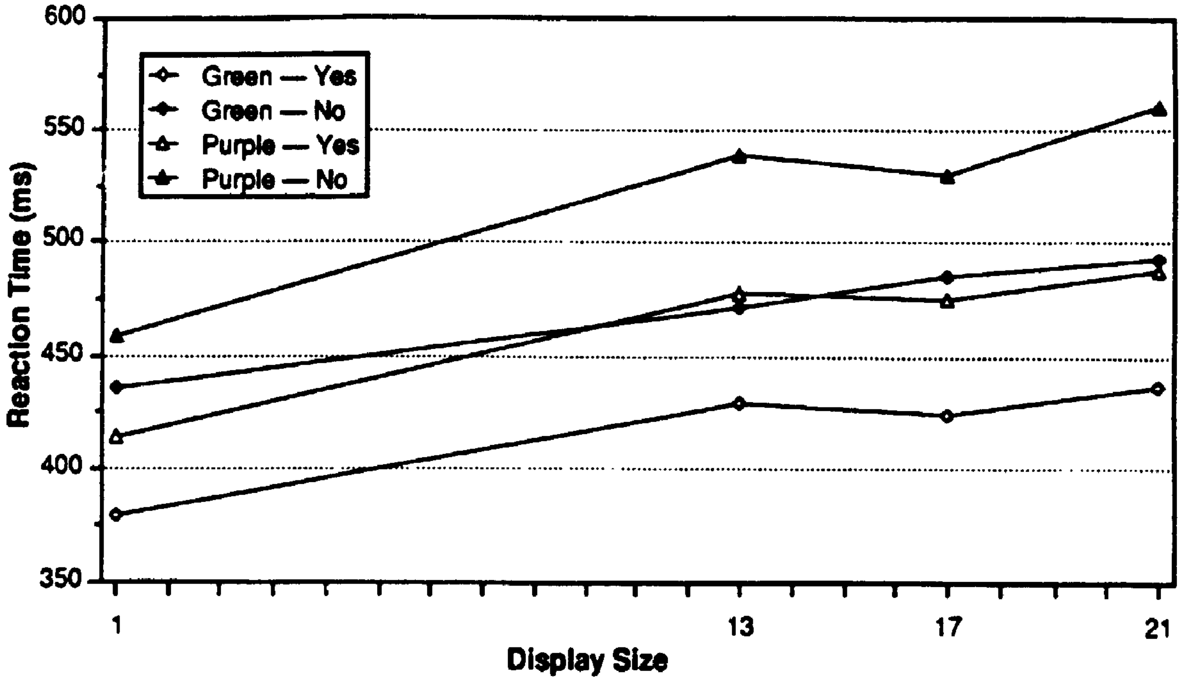

Figure C-1: Search latencies for green vs purple items

taking place ($F(3,12)=0.4$, $p > .1$). Moreover, display size did not interact with colour (F [illegible] was there evidence of a three way interaction between colour, display size and response ($F(3,12)=0.7$, $p > .1$).

By analyzing the trials in which there were several distractors in the display separately from those in which there was only 1 item, a slightly different picture emerges, however. The display size effect disappears once the 1 item trials are removed from the analysis ($F(3,12)=1.5$, $p > .2$). Moreover, the difference between the Green and Purple conditions shrinks to a marginally significant 30 msec effect when there is 1 item ($F(1,4)=5.8$, $p = .073$) as compared to a robust 55 msec effect when there is 13-21 items in the display ($F(1,4)=29.8$, $p < .01$). Thus, although subjects are slower to respond to even 1 item in the *Purple condition*, this difference is slightly, though not significantly, inflated once there are distractors.

Trend analysis was performed on the latencies, taking reaction time as a function display size. Although weak linear trends emerged, no significant deviations from linearity were apparent. Regression was performed on the latencies to determine the slope of the linear functions. The slope of the latency functions in the *Green condition* was 2.8 msec/item for positive responses and 2.9 msec/item for negative responses. In the *Purple condition* the slope of the function for the positive trials was 3.8 msec/item and 4.9 msec/item for the negative trials. Each of these slopes fell within each other's 95% confidence interval, and moreover, none of these slopes differed significantly from 0.

Because display size had little effect on reaction times for either the green or purple items, and because there was no interaction between response type and display size, green and purple were judged adequately discriminable to produce pop out in search.

Appendix D
Experiments 2 and 3

D.1. Experiment 2: Counting concentric items

Standard deviations are in parentheses.

REACTION TIME IN MILLISECONDS

	Same size	Different size	Concentric
1	472 (42)	478 (44)	486 (36)
2	549 (67)	557 (57)	612 (83)
3	585 (57)	615 (73)	890 (184)
4	709 (107)	827 (108)	1309 (200)
5	1147 (260)	1306 (219)	1788 (193)
6	1527 (268)	1542 (250)	2150 (276)
7	1881 (346)	2010 (298)	2502 (353)
8	2049 (423)	2236 (381)	2786 (633)

PERCENT ERROR

	Same size	Different size	Concentric
1	0 (0)	0 (0)	0 (0)
2	0 (0)	0 (0)	0 (0)
3	0 (0)	0 (0)	5 (17.3)
4	.96 (3.5)	0 (0)	.96 (3.5)
5	2 (4.7)	2 (4.7)	2 (6.7)
6	3 (5.4)	.96 (3.5)	2 (4.7)
7	4 (7.8)	3 (7.4)	5 (9.5)
8	4 (7.8)	3 (5.4)	9 (11.6)

Error analysis

Overall, the error rate was very low in this study, probably because subjects were encouraged to be as accurate as possible. Nonetheless, number had a significant effect on accuracy ($F(7,77)=3.7$, $p<.001$). Subjects made no errors when counting small numbers of items, up to 3 in the *Same size* and 4 in *Different size* condition. In the *Concentric* condition performance was errorless only to 2 items. In the *Concentric* condition subjects made more errors within the subitizing range than in the other conditions, because of the high error rate at 3. Given that the counting process is more error prone than subitizing this result suggests that subjects are counting even when there are only a small number of concentric rectangles. This interaction between condition and number was not statistically significant. As with the latency data, 2 in the *Concentric* condition is an embarassment. The error rate at 2 is much lower than expected, perhaps because subjects were very familiar with configurations of 2 concentric rectangles (e.g., picture frames, door frames), and thus tended not to err.

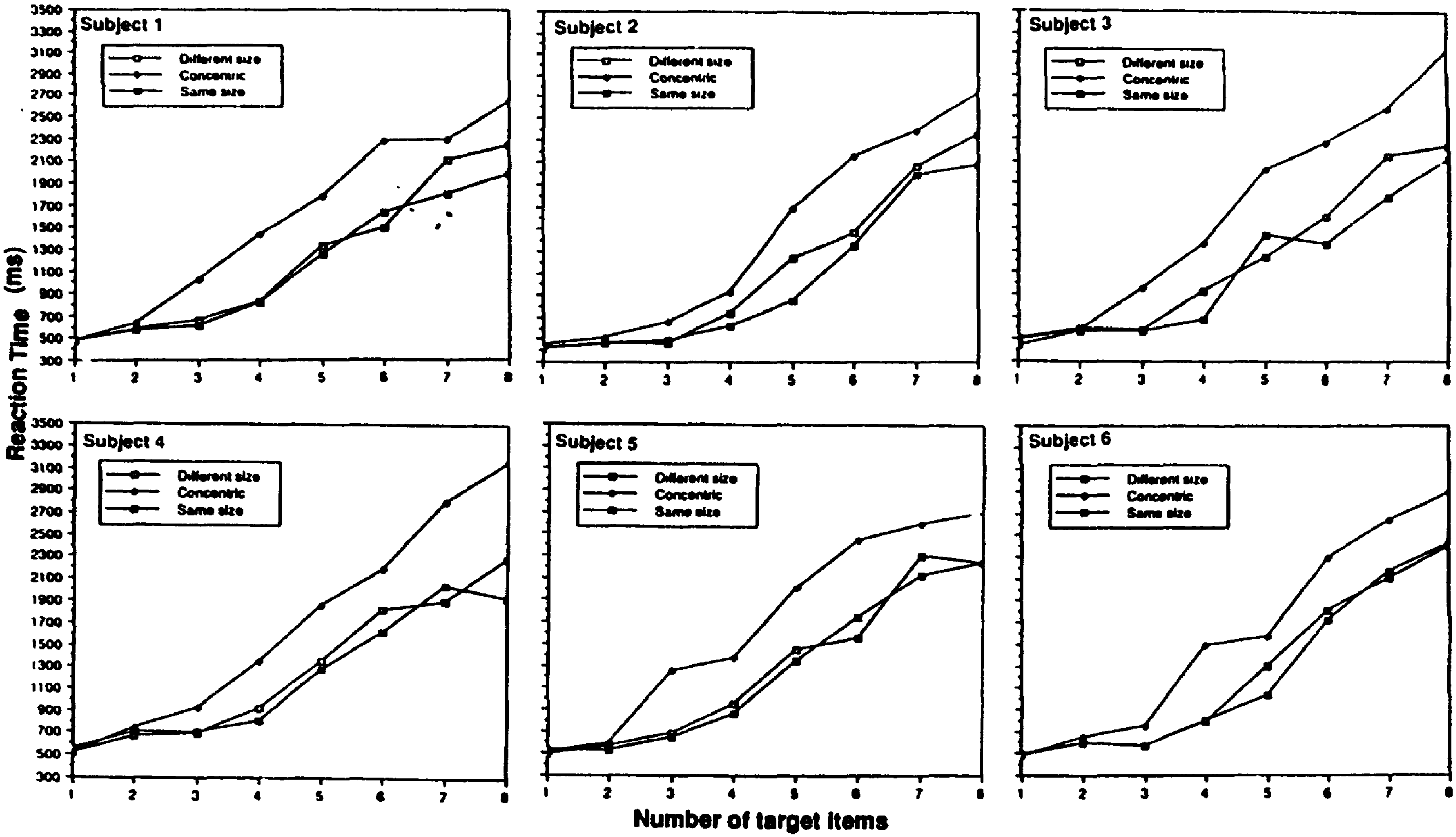

Figure D.1. Average response latencies to count rectangels of the Same size, Different sizes, or presented Concentrically for each subject.

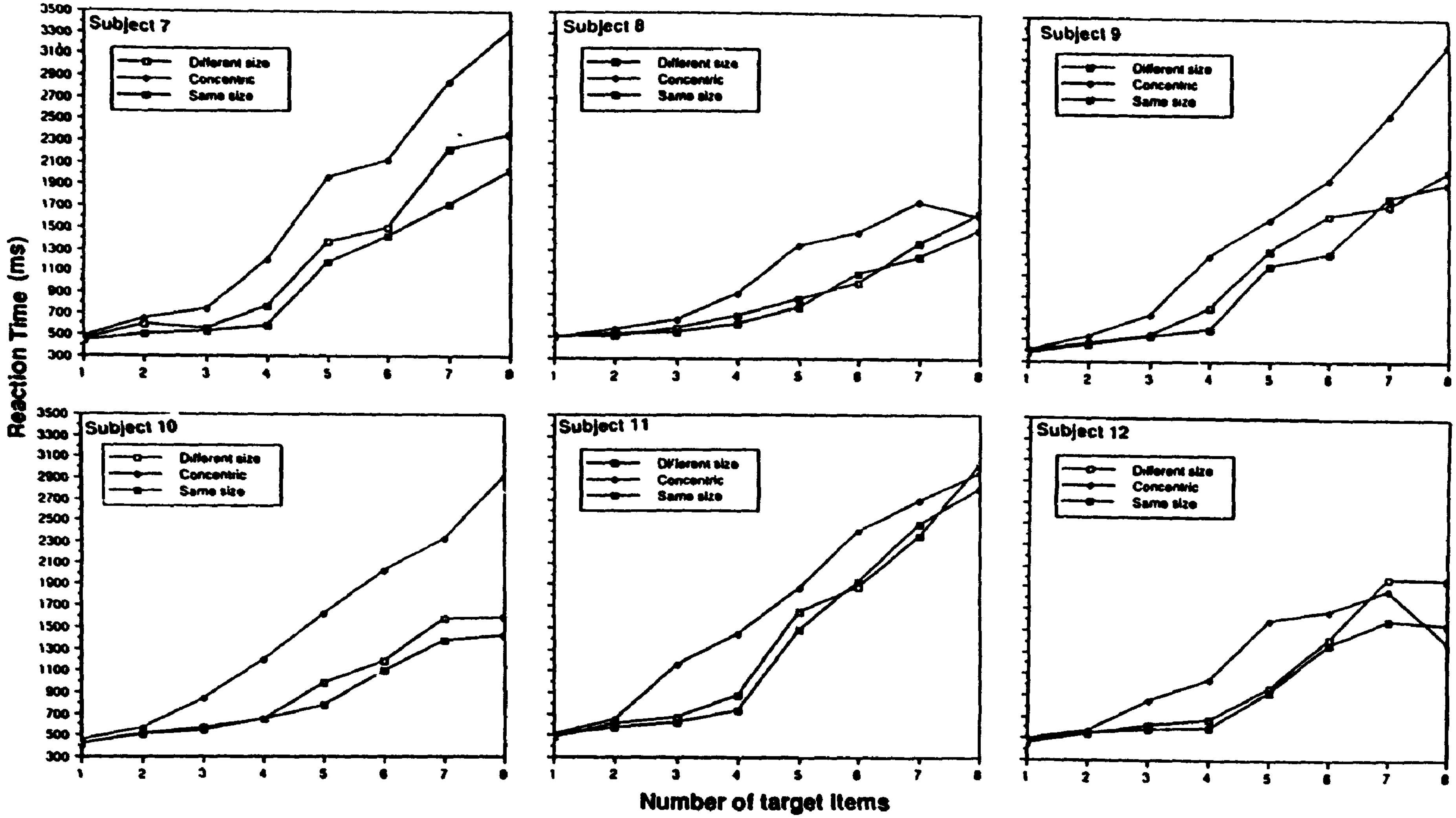

Figure D.1., continued.

Appendix D (continued)

D.2. Experiment 3: Counting parallel lines and parallel corners

Parallel lines and corners study: Reaction time in milliseconds

Standard deviations are in parentheses

REACTION TIME IN MILLISECONDS

	Parallel lines	Parallel corners
1	448 (53)	460 (53)
2	522 (67)	519 (69)
3	555 (54)	573 (61)
4	701 (96)	711 (115)
5	979 (118)	980 (157)
6	1369 (168)	1318 (171)
7	1715 (188)	1646 (265)
8	1975 (295)	1869 (297)

PERCENT ERROR

	Parallel lines	Parallel corners
1	0 (0)	0 (0)
2	0 (0)	0 (0)
3	0 (0)	0 (0)
4	.96 (3.5)	0 (0)
5	2 (4.7)	2 (4.7)
6	3 (5.4)	.96 (3.5)
7	4 (7.8)	3 (7.4)
8	4 (7.8)	3 (5.4)

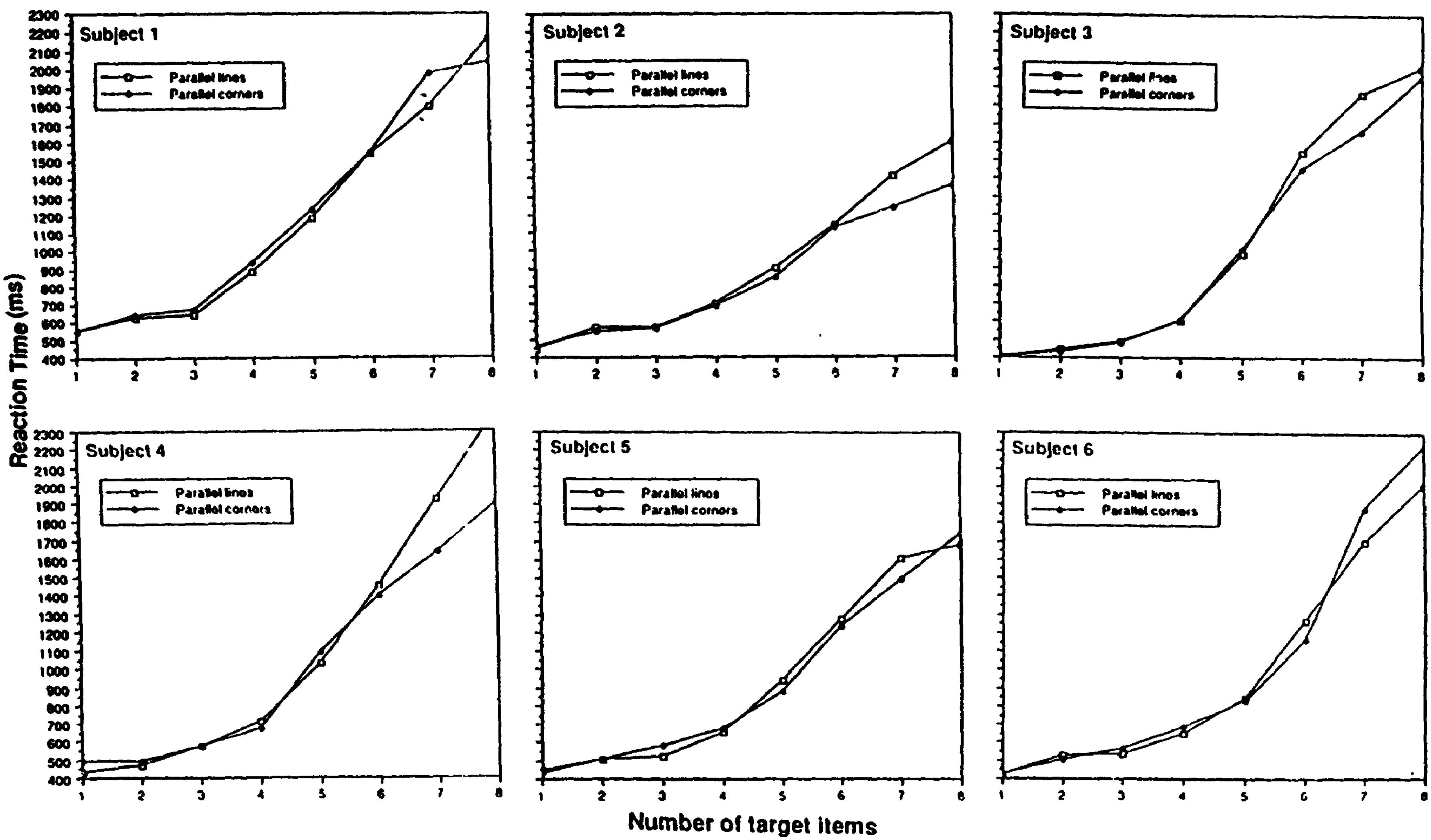

Figure D.2. Average response latencies for counting parallel lines or corners for each subject.

D.3. Recent experiments on the ability to enumerate concentric items

Two recent studies that the ability to subitize concentric items might vary with the size of the items. The first was run with undergraduates and the second with experienced counters, graduate students. In both studies there was evidence of subitizing of concentric items when the number of possible item sizes was increased to 18, and the maximal ring size was increased from 7 to 9 degrees visual angle. Thus, when the items are more spread out on average, subitizing is evident for concentric items. Though the slope in the 1-3 range for concentric items was twice the slope for parallel lines and corners in the same study, there was a clear break between the slopes for 1-3 and 5-7 items.

If subjects are capable of subitizing large concentric items, why did Saltzman and Garner fail to find evidence of subitizing with 14 degree concentric circles, however? Perhaps the variable of importance isn't size as much as the spacing. In Saltzman and Garner's study the items were equispaced. When there are a small number of possible item sizes, as in experiment 2, there is a greater probablity of items being equispaced. On the other hand, with the same number of item sizes there is evidence of subitizing in experiment 3. Moreover, there is clear evidence of subitizing when subjects are required to enumerate equispaced items in other studies (e.g., Frick, 1987).

It is puzzling. At this point it is not clear why sometimes there is evidence of subitizing of concentric items, and sometimes there is not. Nonetheless, another recent study shows that if subjects are required to count whole (4 sided) concentric boxes in a field of two or four 3 sided boxes, the ability to subitize is completely lost. In this case subjects could no longer focus on one side of the display and ignore the other because the number of sides or corners was not an accurate reflection of the number of items. Even the strategy of moving the attentional focus through the display and counting edge crossings or corners would not work, given this task; subjects *have* to consider each object as a whole before they enumerate it. The task proved exceedingly difficult, as judged by the counting latencies; the slopes were constant in the 1-3 and 5-7 range, and moreover very large, 400-500 msec/item even with only 2 distractors. Further, the error rate was astonishing. Subjects made errors on 40% of the trials in which they were required to enumerate 6 concentric rectangles when there were 4 three sided distractor items, for example. If subjects were required to count non-concentric 4 sided boxes and there are 2 or 4 three sided boxes as distractors, subitizing was once more apparent Moreover, the error rate was low, 5.5% on average when there were 4 distractors.

Appendix E
Search pilot for Experiment 4: Finding O's in X's and O's in Q's

The goal of this exercise was to determine if the pop out phenomena could be generated using letters from the Apple character set. Specifically, in both conditions the task was to indicate whether there were any O's in the display. In one condition the distractor letters were X's and in another they were Q's, however. Treisman has shown that O's will pop out of X's; subjects were able to indicate the presence of an O in a display of X's in a time roughly independent of the number of items in the display (Treisman and Gelade, 1980). In contrast, O's do not pop out of Q's; the time required for subjects to indicate that there is an O in the display increases with the display size (Treisman, 1985). In Treisman's study, however, the position of the stem on the Q's varied, and consequently there is some question of whether there will be evidence of serial search using the Apple character set given that there is a particular location to check. According to Ullman (1984) if there is one position to check for a particular feature, a template scheme might be adequate for the purpose and thus pop out should occur.[1]

[1]Efforts had been made to create larger stimuli using high resolution graphics on the Apple, and in which the position of the stem on the Q in particular could be varied, but typically, these artificial letters produced pop out regardless of the condition. Thus, even in situations in which pop out was not supposed to occur, i.e., when subjects were searching for O's in in two or four Q's. Consequently, in order to ensure one condition in which pop out did not occur, I resorted to using the characters from the Apple character set, in which I could obtain at least one case in which pop out did not occur. Unfortunately, I was not able to replicate Treisman's asymmetry; there is evidence of serial search when subjects were looking for Q's among O's as well as O's among Q's

Method

Subjects

Three subjects volunteered to participate in the study. The subjects were graduate students and research assistants at the University of Western of Ontario One was female. One of the three had never been in a search study before.

Materials and Stimuli

An Apple II+ computer was used to display the stimuli and record the data. Subjects responded by either pushing the "0" key on the terminal, that had an orange "Y" pasted above it, standing for Yes, or by pushing the "1" key on the terminal that had an orange "N" pasted above it, standing for No.

Stimuli were identical to those used in experiment 2. The letters used in the experiment were from the standard Apple character set, and each occupied 0.69 by 0.40 cm or a 0.36 by 0.21 degrees visual angle when subjects were seated 110 cm from the video screen. The location of the stimuli was chosen from a 49 point matrix that covered a 5.97 by 4.2 degree area of visual angle. The minimal distance between letterswas 0.73 horizontal and 0.36 vertical degrees.

Subjects were required to either to indicate whether there was an O in the display. There were also a number of distractor letters, in one condition the distractors were X's and in another they were Q's. There were four display sizes, there was either 1, 3, 5 or 9 letters in the display. Half of the time there was a letter O in the display and half the time there was not.

Before the letter display was presented, and after the subject had made a response a black and white random dot mask was shown. In the centre of the mask was a 0.5 cm black square that served as the fixation point.

Procedure

The subjects' task in this experiment was to press the key designated as the "Yes" key if there was a letter O in the display, and to push the key designated as "No" if there were no O's in the display. There were two conditions. In one condition the distractor letters were X's whereas in the other distractors were Q's. Condition was blocked so that subjects did a session in one condition immediately followed by a session in the other. The order that these sessions were presented was counterbalanced, so that half of the subjects search for O's in X's first, and the others searched for O's in Q's. Before each condition subjects were given 12 practise trials to get accustomed to their experimental task

Each trial had two phases. First, in the fixation phase, for a period of 512 msec, subjects were shown the random dot mask, and required to focus on the fixation point. During this time, and 256 msec before the display was presented the computer beeped to warn the subject that the display was imminent. Second, the letter display was presented. The display remained on until subjects made a response by pushing either the keys designated as the Yes and No keys. Reaction time was measured from the onset of the letter display until the time when subjects made their key press response. If subjects were incorrect or pushed a key that was neither the yes or no key the computer beeped twice at them to indicate that they had made a mistake. Trials in which subjects had pushed neither the yes or no key were re-administered at the end of the block, with changes in the spatial arrangement in the letters. In particular, re-administered trials were mirror reflections of the item locations, rotated 90 degrees.

The entire experiment took approximately 30 minutes, involving 160 trials in each condition plus 12 practise trials.

Results and Discussion

Only reaction times from correct responses were analyzed. For each subject, latencies outside two standard deviations of their respective mean for each condition, display size and response were dropped from the analysis, also.

A fixed factors analysis of variance was performed on latencies. As predicted, condition had a significant effect on reaction time, with subjects requiring more time to search for O's when the distractors were Q's than when the distractors were X's ($F(1,2)=3009.7, P<.001$). Further, the mean time for negative responses was longer than for positive ($F(1,2)=81.8, p<.01$). Finally, overall subjects took longer to respond when there were larger numbers of items in the display ($F(3,6)=687.3, p<.001$). There was also a number of interactions between factors. First, the difference betweeen positive and negative trial latencies was greater when subjects were searching for O's in Q's than O's in X's ($F(1,2)=77.6, p<.01$). Further, the number of distractors had a much greater effect on latencies when subjects were searching among Q's than X's ($F(3,6)=885.6, p<.001$). See Figure E. Finally, there was a three way interaction between condition, response and the number of distractors ($F(3,6)=21.5, p<.002$).

Trend analysis was performed on the latencies, and in all cases no significant deviations from linearity occured, though there were clear linear effects. Regression was performed on the latencies in order to determine the slopes for the search functions. When subjects were required to search for O's in X's the slopes for no response and yes response trials were virtually identical (7.5 msec/item as opposed to 7.9 msec/item respectively). In contrast, when subjects were required to search for O's among Q's the slopes for negative trials was almost twice that for positive trials (208 msec/item as opposed to 117 msec/item), a difference that is significant.

Consequently, there was clear evidence of pop out when subjects were searching

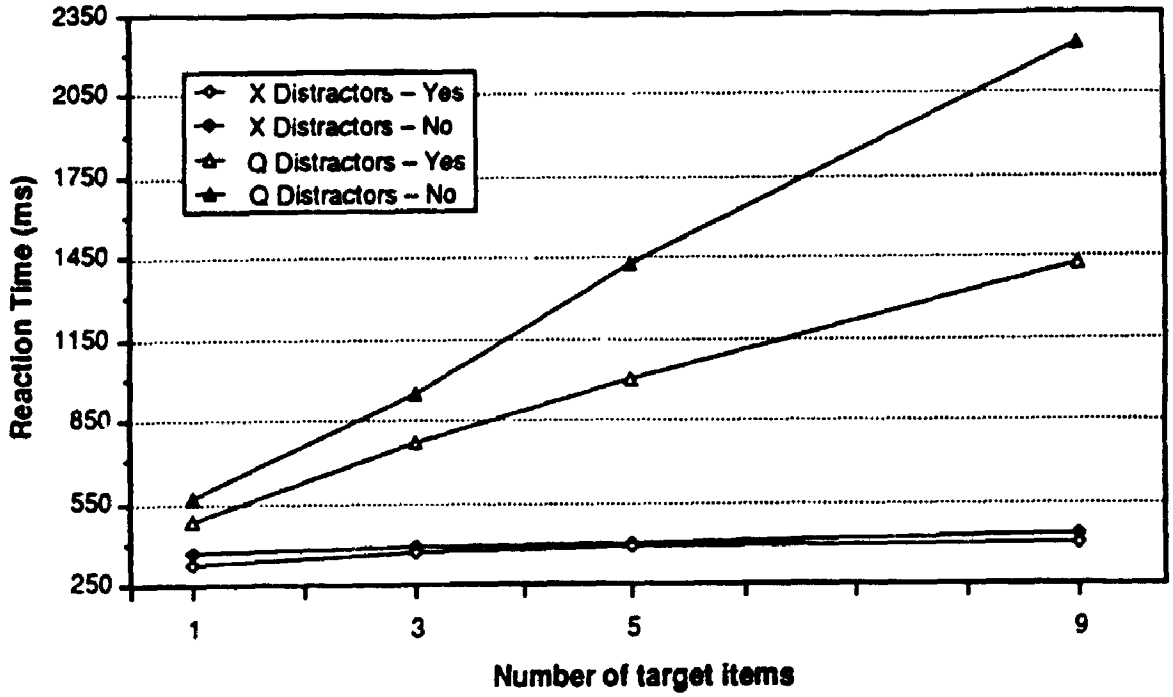

Figure E-1: Search latencies for O's in distractor letters

for O's in a background of X's; the slope for both positive and negative trials was a mere 8 msec/item. In contrast, searching for an O amid Q's, even amid Q's in which the stem always was located on the bottom of the item, produced evidence of serial search. There was a substantial slope, in excess of 100 msec/item, and moreover, the slope for negative trials was twice that of positive trials, 208 msec/item as opposed to 117 msec/item. Consequently, the stimuli were judged adequate for use in experiment 2.

Appendix F
Experiment 4: Counting O's study

F.1. Counting latency by condition and number of distractors

Average reaction time in milliseconds

Standard deviations are in parentheses.

POP OUT CONDITION. Counting O's in a background of X's

	0 X's	2 X's	4 X's
1	453 (52)	526 (46)	596 (74)
2	502 (49)	616 (68)	719 (101)
3	542 (60)	716 (101)	847 (147)
4	654 (107)	888 (165)	1041 (178)
5	915 (237)	1184 (212)	1293 (172)
6	1274 (210)	1384 (161)	1607 (213)
7	1467 (171)	1708 (207)	1865 (183)
8	1668 (220)	1852 (266)	2071 (320)

ATTENTIVE SEARCH CONDITION: Counting O's in a background of Q's

	0 Q's	2 Q's	4 Q's
1	528 (68)	1041 (121)	1645 (108)
2	682 (47)	1418 (119)	1968 (157)
3	785 (74)	1687 (110)	2273 (151)
4	996 (89)	2011 (155)	2639 (175)
5	1456 (140)	2285 (165)	2914 (191)
6	1729 (179)	2560 (238)	3258 (235)
7	1966 (219)	2876 (207)	3535 (201)
8	2324 (213)	3106 (245)	3734 (294)

Appendix F (continued)

F.2. Error as a function of condition and number of distractors

Percent error

Standard deviations are in parentheses.

POP OUT CONDITION: Counting O's in a background of X's

	0 X's	2 X's	4 X's
1	0 (0)	0 (0)	.7 (2.4)
2	0 (0)	0 (0)	0 (0)
3	0 (0)	0 (0)	0 (0)
4	0 (0)	7 (2.4)	0 (0)
5	0 (0)	.7 (2.4)	1.4 (3.3)
6	0 (0)	2.8 (5.4)	2.8 (4.1)
7	2.1 (5.2)	2.8 (6.5)	4.8 (7.5)
8	2.1 (3.8)	4.2 (1.1)	4.8 (8.3)

ATTENTIVE SEARCH CONDITION: Counting O's in a background of Q's

	0 Q's	2 Q's	4 Q's
1	.7 (2.4)	.7 (2.4)	0 (0)
2	.7 (2.4)	2.1 (3.8)	1.4 (3.3)
3	0 (0)	1.4 (3.3)	.7 (2.4)
4	.7 (2.4)	4.2 (7.5)	4.2 (5.6)
5	0 (0)	4.2 (5.6)	3.5 (4.3)
6	1.4 (3.3)	5.6 (5.4)	2.8 (4.1)
7	2.1 (3.8)	3.5 (5.6)	4.8 (6.6)
8	3.5 (5.6)	4.2 (4.3)	5.6 (8.2)

Error analysis

Subjects made very few errors, only 1.83% on average. Nonetheless, condition, number of items, and number of distractors all had effects on the error rate. Subjects tended to make more errors when counting in the *OQ* than the *OX* condition ($F(1,11)=8.0, p<.05$). In fact, subjects were approximately twice as likely to err in the *OQ* condition; the average error rate was 2.42% as opposed to 1.24%. Also, the probability of error increased with the number of O's ($F(7,77)=6.3$, $p<.001$). Distractors also affected error rate ($F(2,22)=6.3$, $p<.01$). Accuracy was best when there were no distractors in the display. With no distractors in the *OX* condition the average error rate is 0.53% when there is no distractors whereas in the *OQ* condition the average error rate is 1.14%. (1)

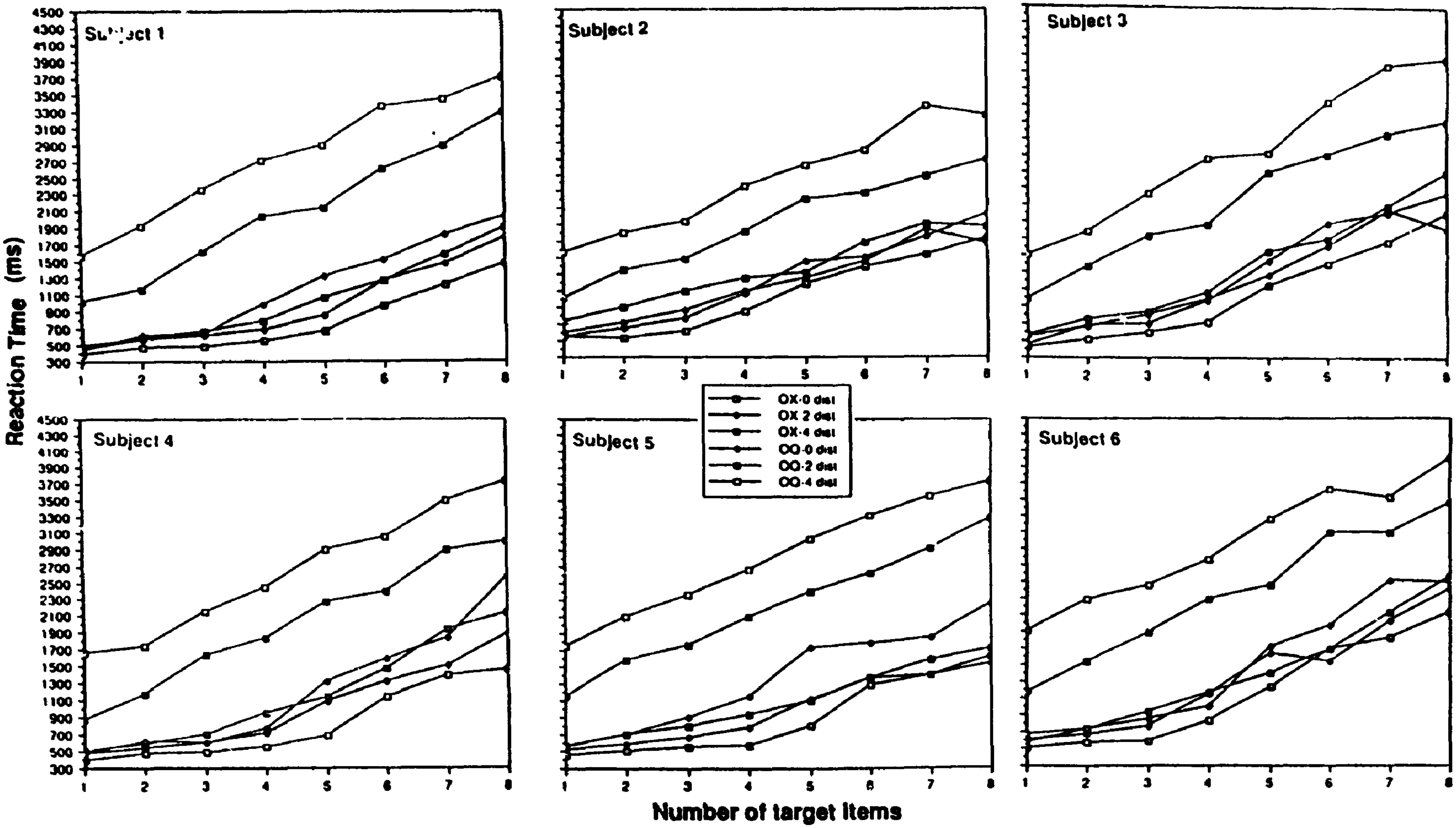

Figure F. Average response latencies for counting "O"s in a background of varying nu.nber of distractor letters for each subject.

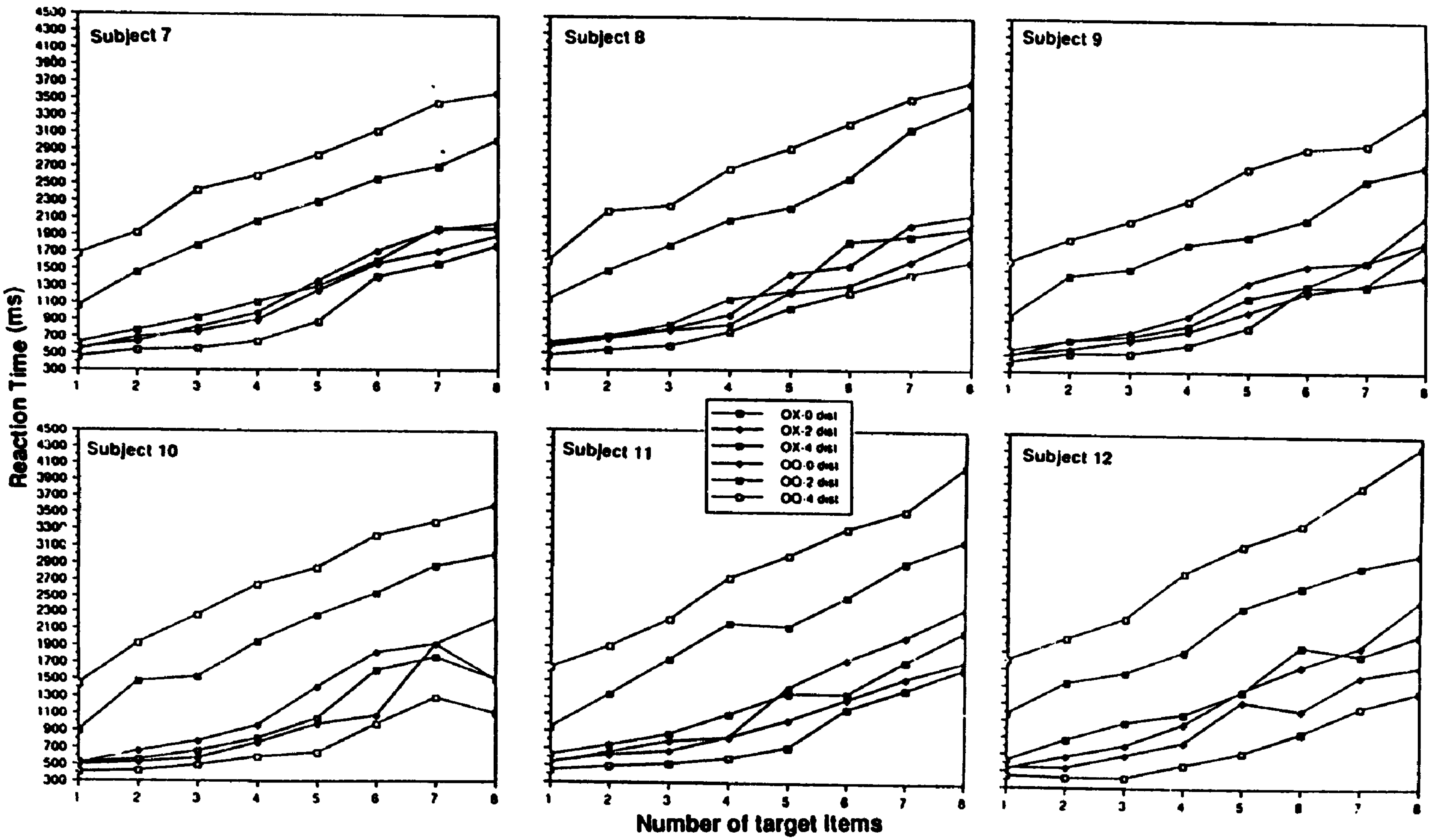

Figure F, continued.

Appendix G
Search pilot for Experiment 5: Finding white vertical vs. white or vertical lines

The goal of this experiment was to find out how many distractors are needed before the typical 2:1 ratio of negative to positive slopes for *Conjunction search* are apparent. Pashler (1987) found that this pattern of results does not emerge until there are more than 8 items in a display. Before that point Feature and Conjunction search are quite similar in that the ratio of negative to positive slopes is approximately 1:1.

Method

Design

In this experiment subjects were required to indicate whether or not certain designated target items were in the display by pushing "Yes" or "No" keys. There were three factors. The first was task. In the *Disjunction* condition subjects were required to push the "Yes" key if there was a vertical item or white item in the display, and the "No" key otherwise. (Vertical and white items never appeared in the same display). In the *Conjunction* condition subjects were required to push the "Yes" key if there were a vertical white item in the display and "No" otherwise. The second factor was response. In positive trials the specified target was in the display and thus the correct response would be "Yes" whereas in negative trials the target was not in the display and thus the correct response would be "No". The third factor was display size, and had to do with the number of items in the display. There were either 1, 3, 5, 9, 13, or 21 items in the display. These particular display sizes were chosen so that positive trials would be precise analogues of trials in which subjects had to count 1 item in a field of 0, 2, 4, 8, 12 and 20 distractors. Condition was blocked in this experiment, so that subjects performed the *Disjunction* task one day and the *Conjunction* task another. The order in which the two sessions were performed was counterbalanced.

The dependent variable was reaction time, the time required for subjects to make a response.

Subjects

Five subjects were persuaded to participate for payment in *Hershey's* chocoloate kisses. Three subjects were female and the remainder were male. All were graduate students or research assistants at the University of Western Ontario. Two of the five subjects had been in search experiments before. The experiment involved two sessions. Each subject participated in every condition of the experiment.

Materials and Stimuli

An Apple II+ computer was used to present the displays and record the data. Subjects were required to indicate whether a target item was in the display by pushing keys on the computer terminal. The "0" key had an orange "Y" pasted over it and was designated the "Yes" key whereas the "1" key had an orange "N" pasted over it and was designated the "No" key.

The stimuli used in this study were identical to the ones used in experiment 4. The items were 0.5 cm vertical or horizontal lines that were either green or white. Each line occupied approximately 0.26 degrees visual angle when the subject was seated 110 cm from the display. These lines could appear in any of 40 positions on a 40 point grid. The minimal horizontal distance between items was 1.5 cm or 0.78 degrees visual angle. The minimal vertical distance between points was 1.8 cm or 0.94 degrees visual angle, and the minimal diagonal distance was 1.2 cm or 0.625 degrees. The total area that the items could appear in was 8.5 by 10 cm, or 4.42 by 5.19 degrees visual angle.

The number of items in a display varied between 1 and 21; there was either 1, 3, 5, 9, 13, or 21 items in the display. The composition of the display varied according to condition. In the *Disjunction* condition the target items were either green vertical lines or white horizontal lines. The distractors were always green horizontal lines. In the

Conjunction condition the target items were always white vertical lines and the distractors were green vertical and white horizontal lines. (There were approximately equal numbers of each type of distractor). In either condition, in half the trials a target item was present in the display, and in half the target was absent.

The fixation display was a coloured random dot mask with a 0.5 cm black square in the centre of the screen. This fixation point occupied approximately 0.26 degrees visual angle when subjects were seated 110 cm from the display.

Procedure

The subjects' task was to indicate whether a specified target item was in the display by pressing keys designated as "Yes" and "No" buttons. Subjects were encouraged to respond as rapidly as they could, with accuracy.

Each trial had two phases. First, the fixation display appeared. Subjects were instructed to fixate on a central block in a coloured random dot mask. The fixation display remained on for 416 msec. During this time, and 256 msec before the experimental display appeared, the computer beeped to warn the subjects that the trial was imminent. The experimental display of vertical and horzintal lines then appeared. This display remained on until subjects made a response by hitting a computer key. If subjects made a mistake or hit a key that was neither the "Yes" or "No" button, the computer beeped twice to indicate the error. Trials in which subjects missed the "Yes" or "No" button were re-administered at the end of the block. These readministered trials were disguised by changing the position of the items in the display. Items in readministered trials were in locations that were the mirror image of the original locations, and rotated by 90 degrees.

The session involved 240 experimental trials and began with 12 practise trials. At most a session would take 40 minutes.

aggravating the problem by increasing differences in the proportion of missing cases per cell, latency analyses were performed on raw data for within cell analysis.[2]

Analysis of variance

As can be seen from Figure 4-3, both the number of O's and the number of

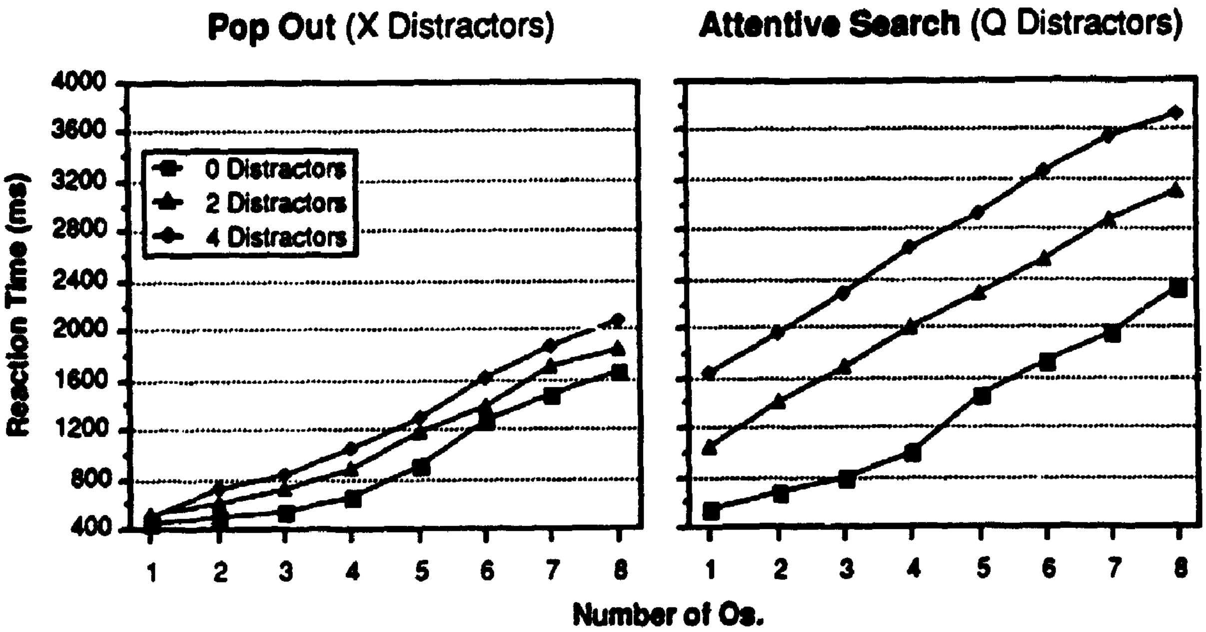

Figure 4-3: Average latency to count O's in distractor letters

distractors affected reaction time. In fact, fixed factors analysis of variance revealed all main effects and interactions to be significant. Overall, counting took longer in the session in which subjects were required to count O's in the background of Q's ($F(1,11)=618.7, p<.001$); subjects were 960 msec slower at counting in the *OQ* condition than the *OX* condition. Reaction times increased as the number of O's increased ($F(7,77)=768.5, p<.001$). These increases were greater in the *OQ* session than the *OX* session ($F(7,77)=57.0, p<.001$), however. The difference between latencies for counting 1

[2]These analyses were also performed on averages in which the latencies beyond two standard deviations of the mean for each subject, condition, and number were dropped (Trick and Pylyshyn, 1989). The results are not qualitatively different.

Results and Discussion

Latencies for trials in which there were errors were dropped from the analysis, as were trials with reaction times that were outside two standard deviations for each condition, display size, and response for each subject. From the remaining latencies an average was calculated for each subject, and this average latency was entered into a fixed factors analysis of variance.

All main effects and interactions were significant. See Figure G. As predicted,

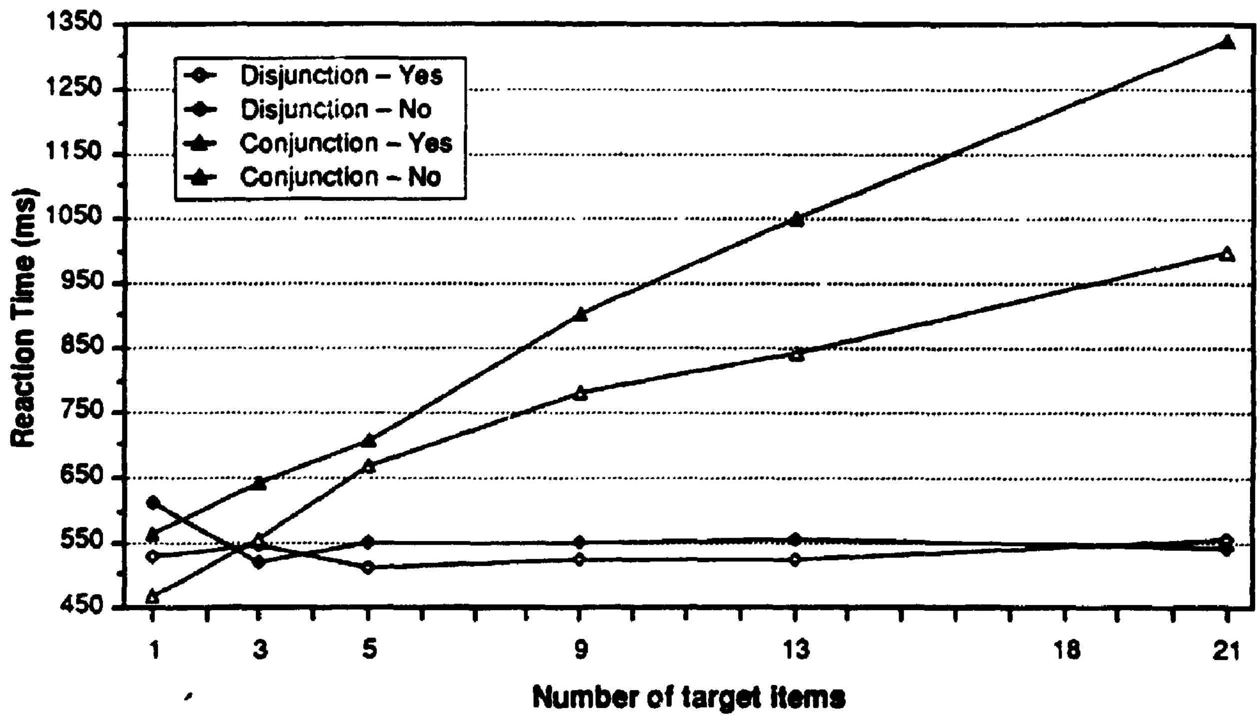

Figure G-1: Search latencies for coloured lines

condition had a significant effect on response latencies. Overall, it took subjects 248 msec longer to respond in the *Conjunction* condition than in the *Disjunction* condition ($F(1,4)=86.4, P=.001$). Moreover, subjects were fastest in positive trials ($F(1,4)=92.0, p=.001$); "Yes" responses were made 85 msec faster than "No" responses. Finally, the number of items in the display also had an effect on latencies ($F(5,20)=77.0, p<.001$), though there was also a strong interaction between condition and display size ($F(5,20)= 172.1, p < .001$), such that the number of distractors had the greatest effect on latencies in the *Conjunction* condition. The difference between positive and negative responses was much greater when subjects were searching for white vertical items than when they were searching for white or vertical items ($F(1,4)=19.2, p<.05$), 147 msec as opposed to 23. Finally, there was a three way interaction between condition, response and number of distractors ($F(5,20)=21.0, p<.001$).

Given that condition interacts with every other factor, latencies for *Disjunction* trials were analyzed separately from *Conjunction* trials. Display size had no significant effect on latencies when subjects were searching for white or vertical items ($F(5,20)=1.5, p>.1$). Subjects were actually slower to respond when there was 1 item in the display than when there were 21, probably because it was necessary to consider the colour or orientation of the item in the former case, rather than simply responding if there was an "odd man out". Response type had a marginal effect, with negative trials taking slightly longer than positive ($F(1,4)=4.8$, $.05 < p < .1$). Finally, there was an interaction between display size and response type ($F(5,2)=3.5, p < .05$), probably because "No" responses were made slightly faster than "Yes" responses when there were 3 or 21 distractors.

For the *Conjunction* condition, both display size and response had significant effects on reaction times. The larger the number of items in the display, the longer it took subjects to decide whether there was a target ($F(5,2)=140.8$, $p < .001$). In fact, on average it required 645 msec longer to make the decision when there were 21 in the display than when there was 1 in the display. "No" responses were made significantly slower than "Yes" responses ($F(1,4)=47.5, p < .005$), by a factor of 147 msec. As expected, there was also a significant interaction between display size and response ($F(5,20)=9.7, p < .001$); this interaction is typical of results from Treisman's *Conjunction* search also. The interaction between response and number of distractors did not emerge until there were

13 distractors, however ($F(2,8)=1.5$, and 1.6 $p > .2$ when there was 1-5 or 1-9 distractors as opposed to $F(2,8)=17.8, p=.001$ for 1-13). Whenever there were 9 or more distractors a significant interaction was apparent, however ($F(1,4)=8.5, p<.05$ when there were 13-21 distractors). The late emergence of the interaction between display size and response is reminiscent of findings from Pashler (1987).

Trend analysis was performed on the latencies. In the *Disjunction* condition there were no linear trends in latency as as would be expected when the slope is near 0 ($F(1,24)< 1$, $p > .1$ both positive and negative trials). Neither were there any deviations from linearity ($F(4,24)< 1$), $p > .1$) for both positive and negative trials. Nonetheless, the slopes will be broken down by range for purposes of comparison with the *Conjunction* trials. When there were fewer than 13 items, the slope for positive trials was -1.5 msec/item whereas the slope for negative trials was -4.5 msec/item. When there were 13 or more items, The slope for positive trials was 3.4 msec/item whereas the slope for negative trials that was -1.8 msec/item.

In contrast, there were strong linear trends in the data in the *Conjunction* condition ($F(1,24)=108.2, p < .001$ and $F(1,24)=262.9$, $p < .001$ for the positive and negative trials respectively). Moreover, there were significant quadratic components in the latencies for the "Yes" responses in the *Conjunction* condition ($F(1,24)=4.4, p < .05$). Regression was performed in order to determine the slopes of the functions. When there were 9 or fewer items in the display, the slopes for "Yes" and "No" functions were almost identical, 39.3 msec/item and 42.2 msec/item for for positive trials and negative trials respectively., or approximately 1.1/1. When there were 13 or more items in the display the slope for negative trials was almost twice that of positive, 34.3 msec/item as opposed to 19.6 msec/item, or 1.8/1. For both positive and negative trials, the slope drops with an increase in display size after 9, but the drop is greatest for the positive trials.

As a result of this experiment 12 and 20 distractors were used in experiment 4.

because they best approximate the standard pattern of negative to positive slopes in Conjunction search reported in many search studies, although the optimal 2:1 pattern was never achieved.

Appendix H
Experiment 5: Counting coloured lines

H.1. Session analysis: Average counting latency as a function of session

Standard deviations in parentheses.[1]

	0 distractors	12 distractors	20 distractors
DISJUNCTION CONDITION: Counting White OR Vertical Lines			
Session 1			
1	540 (118)	692 (101)	665 (67)
2	570 (93)	746 (105)	768 (86)
3	621 (100)	855 (162)	861 (122)
4	723 (132)	1020 (180)	1024 (209)
5	999 (215)	1346 (295)	1382 (321)
6	1403 (285)	1632 (292)	1741 (274)
7	1751 (361)	1936 (379)	2010 (361)
8	1992 (356)	2156 (412)	2299 (488)
Session 2			
1	482 (51)	647 (61)	651 (119)
2	537 (84)	711 (98)	705 (99)
3	595 (87)	807 (163)	750 (103)
4	674 (94)	952 (195)	942 (204)
5	969 (271)	1239 (225)	1207 (282)
6	1301 (321)	1601 (325)	1633 (376)
7	1576 (330)	1826 (402)	1897 (437)
8	1780 (425)	2190 (484)	2106 (413)
CONJUNCTION CONDITION: Counting White Vertical Lines			
1	528 (53)	1338 (257)	1828 (411)
2	600 (80)	1532 (370)	2039 (527)
3	659 (88)	1795 (442)	2393 (603)
4	795 (154)	1958 (496)	2563 (696)
5	1110 (278)	2259 (612)	2760 (596)
6	1465 (265)	2521 (574)	3040 (813)
7	1851 (398)	2821 (625)	3229 (665)
8	2095 (378)	3139 (782)	3533 (827)

[1]For *Disjunction* condition trials, latencies to count white items and latencies to count vertical items are averaged in this analysis. Outliers are dropped on the basis of session. Thus, for each subject, condition, number and session, the reaction time farthest from the mean is dropped in this analysis.

Appendix H (continued)

Analysis of variance was performed, comparing session 1 and session 2 latencies in the *Disjunction* condition. Subjects were significantly faster at enumerating white or vertical lines in the second session than the first ($F(1,9)=25.1, p < .001$), requiring 1157 msec as opposed to 1239 msec on average. As usual, the number of targets had an effect on how rapidly the items could be counted ($F(7,63)=115.6$, $p < .001$); subjects required 613 msec to count 1 item as opposed to 2087 to count 8, on average, for example. The number of distractors also had an effect ($F(2,18)=79.2$, $p < .001$), with subjects requiring 1032 msec to count with no distractors, and approximately 250 msec more to count when there were distractors (1272 and 1290 msec for 12 and 20 distractors respectively) on average. Session did not interact with either the number of targets ($F(7,63)=1.6, p > .1$) or the number of distractors ($F(2,18)1.9$, $p > .1$), so there is little evidence that practise moderated either of these effects.

The situation is complicated by interactions, however. There was an interaction between the number of targets and distractors, however ($F(14,126)=3.4$, $p < .001$), with the number of distractors having less influence on the latencies to count 1 target than those to count 8. For example, the addition of 12 distractors added 158 msec to the latencies to count 1 and 287 msec to the latencies to count 8 items. There was, however a marginal three way interaction ($F(14,126)=1.6$, $p=.073$), such that interaction between the number of distractors and targets was much stronger in the second session. Thus, for session 1, the addition of 12 distractors added 152 msec to the latencies to count 1, and the additional 8 subtracted 26.4 msec, whereas for latencies to count 8, the addition of 12 distractors added 164 msec and the addition of 8 more added another 143 msec. In session 2, however, latencies to count 1 were 165 msec slower when 12 distractors were added, and an additional 8 distractors added only 4 msec whereas for latencies to count 8 items 12 distractors added 409 msec and the addition of an additional 8 distractors *subtracted* 84 msec from the counting latencies.

Appendix H (continued)

Regression: Subitizing and counting slopes in the *Disjunction* condition by session

All subjects were included in these analyses, regardless of whether or not they showed evidence of subitizing with distractors in the *Disjunction* condition.

	Slope	95% C.I.	R
SUBITIZING RANGE (1-3)			
Overall			
0 distractors	48	20 - 77	.40
12 distractors	81	43 - 118	.49
20 distractors	74	41 - 107	.62
Session 1			
0 distractors	40	-6 - 87	.32
12 distractors	81	24 - 138	.49
20 distractors	98	55 - 141	.66
Session 2			
0 distractors	56	22 - 90	.54
12 distractors	80	28 - 132	.51
20 distractors	50	1 - 98	.37
COUNTING RANGE (5-7)			
Overall			
0 distractors	340	246 - 434	.69
12 distractors	294	194 - 394	.61
20 distractors	329	221 - 438	.62
Session 1			
0 distractors	376	244 - 508	.74
12 distractors	295	149 - 441	.62
20 distractors	314	169 - 458	.64
Session 2			
0 distractors	304	165 - 442	.65
12 distractors	293	146 - 441	.61
20 distractors	344	177 - 512	.62

Appendix H (continued)

Individual trend analysis for *Disjunction* condition, analyzed by session

NUMBER OF SUBJECTS WITH DEVIATIONS FROM LINEARITY IN LATENCIES

	0 distractors	12 distractors	20 distractors
Overall	10/10	10/10	10/10
Session 1	10/10	9/10	9/10
Session 2	10/10	10/10	10/10

NO. OF SUBJECTS WITH DEVIATIONS FROM LINEARITY AT EACH NUMBER

	0 distractors	12 distractors	20 distractors
Overall			
Total	(N=10)	(N=10)	(N=10)
# subitizing to 2	2	1	1
# subitizing to 3	1	3	5
# subitizing to 4	5	3	2
# subitizing to 5	2	3	2
# subitizing to 6	-	-	-
Session 1			
Total	(N=10)	(N= 9)	(N= 9)
# subitizing to 2	1	1	-
# subitizing to 3	2	2	2
# subitizing to 4	4	3	2
# subitizing to 5	3	2	5
# subitizing to 6	-	1	-
Session 2			
Total	(N=10)	(N=10)	(N=10)
# subitizing to 2	3	-	2
# subitizing to 3	-	3	4
# subitizing to 4	5	4	1
# subitizing to 5	2	2	3
# subitizing to 6	-	1	-

Appendix H (continued)

H.2. Dimension analysis: Average counting latency as a function of dimension

Standard deviations are in parentheses.[2]

	0 distractors	12 distractors	20 distractors
DISJUNCTION CONDITION: Counting White Lines			
1	518 (96)	659 (70)	652 (86)
2	559 (84)	695 (109)	721 (116)
3	607 (88)	805 (176)	783 (114)
4	691 (98)	979 (210)	964 (228)
5	1026 (305)	1275 (265)	1304 (267)
6	1331 (331)	1609 (310)	1665 (356)
7	1663 (325)	1875 (404)	1997 (415)
8	1861 (381)	2098 (400)	2179 (440)
DISJUNCTION CONDITION: Counting Vertical Lines			
1	500 (61)	681 (86)	652 (85)
2	544 (82)	772 (96)	753 (71)
3	602 (88)	864 (171)	824 (125)
4	692 (111)	1001 (183)	1015 (200)
5	989 (238)	1318 (262)	1280 (307)
6	1356 (277)	1650 (338)	1724 (309)
7	1661 (330)	1899 (346)	1902 (375)
8	1929 (382)	2208 (500)	2202 (427)
CONJUNCTION CONDITION: Counting White Vertical Lines			
1	528 (53)	1338 (257)	1828 (411)
2	600 (80)	1532 (370)	2039 (527)
3	659 (88)	1795 (442)	2393 (603)
4	795 (154)	1958 (496)	2563 (696)
5	1110 (278)	2259 (612)	2760 (596)
6	1465 (265)	2521 (574)	3040 (813)
7	1851 (398)	2821 (625)	3229 (665)
8	2095 (378)	3139 (782)	3533 (827)

[2]Session 1 and session 2 trials in the *Disjunction* condition are averaged in this analysis. Outliers are dropped on the basis of dimension. Thus, for every subject, condition, number and dimension (i.e., white vs vertical) the reaction time farthest from the mean was dropped from analysis.

Appendix H (continued)

Regression: Subitizing and counting slopes in the *Disjunction* condition by dimension

All subjects were included in these analyses, regardless of whether or not they showed evidence of subitizing with distractors in the *Disjunction* condition.

	Slope	95% C.I.	R
SUBITIZING RANGE (1-3)			
Overall			
0 distractors	48	22 - 74	.44
12 distractors	82	43 - 122	.48
20 distractors	76	44 - 107	.53
Counting white lines			
0 distractors	45	4 - 85	.40
12 distractors	73	16 - 130	.44
20 distractors	66	18 - 114	.47
Counting vertical lines			
0 distractors	51	16 - 86	.49
12 distractors	92	36 - 147	.54
20 distractors	86	42 - 129	.61
COUNTING RANGE (5-7)			
Overall			
0 distractors	327	234 - 420	.68
12 distractors	295	196 - 395	.61
20 distractors	329	223 - 434	.63
Counting white lines			
0 distractors	318	174 - 462	.65
12 distractors	299	155 - 449	.61
20 distractors	347	189 - 505	.65
Counting vertical lines			
0 distractors	336	208 - 464	.71
12 distractors	290	147 - 434	.62
20 distractors	311	159 - 463	.62

Appendix H (continued)

Individual trend analysis for *Disjunction* condition, analyzed by dimension

NUMBER OF SUBJECTS WITH DEVIATIONS FROM LINEARITY IN LATENCIES

	0 distractors	12 distractors	20 distractors
Overall	10/10	10/10	10/10
White	10/10	10/10	10/10
Vertical	10/10	9/10	9/10

NO. OF SUBJECTS WITH DEVIATIONS FROM LINEARITY AT EACH NUMBER

	0 distractors	12 distractors	20 distractors
Overall			
Total	(N=10)	(N=10)	(N=10)
# subitizing to 2	2	1	1
# subitizing to 3	1	3	4
# subitizing to 4	5	3	2
# subitizing to 5	2	3	3
# subitizing to 6	-	-	-
Counting white lines			
Total	(N=10)	(N=10)	(N=10)
# subitizing to 2	2	3	-
# subitizing to 3	1	3	4
# subitizing to 4	3	1	3
# subitizing to 5	3	3	3
# subitizing to 6	1	-	-
Counting vertical lines			
Total	(N=10)	(N=9)	(N=9)
# subitizing to 2	2	-	2
# subitizing to 3	1	3	3
# subitizing to 4	5	2	1
# subitizing to 5	2	3	3
# subitizing to 6	-	1	-

Appendix H (continued)

H.3. Error as a function of session and condition

Percent error for each session

Standard deviations in parentheses. Note that when an error is made, the trial is redone, thus permitting another error to be made on that trial. Also errors arise due to mistakes in typing as well as mistakes in counting.

	0 distractors	12 distractors	20 distractors
DISJUNCTION CONDITION: Counting White OR Vertical Lines			
Session 1			
1	0 (0)	0 (0)	0 (0)
2	0 (0)	0 (0)	0 (0)
3	0 (0)	0 (0)	.83 (2.67)
4	0 (0)	0 (0)	0 (0)
5	.83 (2.67)	1.67 (3.50)	0 (0)
6	3.33 (8.08)	2.50 (5.58)	3.33 (4.30)
7	1.67 (5.25)	5.00 (8.92)	2.50 (5.58)
8	3.33 (4.30)	5.00 (5.82)	4.17 (4.42)
Session 2			
1	0 (0)	0 (0)	0 (0)
2	0 (0)	.83 (2.67)	0 (0)
3	.83 (2.67)	0 (0)	0 (0)
4	.83 (2.67)	1.67 (3.50)	1.67 (3.50)
5	.83 (2.67)	0 (0)	.83 (2.67)
6	0 (0)	2.50 (7.92)	7.50 (13.92)
7	2.50 (5.58)	2.50 (5.58)	1.67 (3.50)
8	.83 (2.67)	1.67 (3.50)	2.50 (4.00)
CONJUNCTION CONDITION: Counting White Vertical Lines			
1	0 (0)	.83 (2.67)	4.17 (7.08)
2	0 (0)	1.67 (3.50)	3.33 (4.30)
3	0 (0)	2.50 (7.92)	9.17 (11.42)
4	0 (0)	2.50 (4.00)	5.83 (6.83)
5	0 (0)	3.33 (5.83)	5.83 (10.42)
6	0 (0)	3.33 (5.83)	6.70 (7.67)
7	1.67 (3.50)	3.33 (5.83)	7.50 (8.25)
8	3.33 (5.83)	7.50 (6.17)	7.50 (10.75)

Appendix H (continued)

Analysis of variance was performed comparing the session 1 *Disjunction* condition error rate with the *Conjunction* condition error rate. Subjects made significantly more errors in the *Conjunction* condition than the *Disjunction* condition, 3.36% as opposed to 1.42% on average ($F(1,9)=7.59, p<.05$). Also, the number of targets had a significant effect, with subjects making more errors the larger the number of targets that they had to count ($F(7,63)=7.29, p<.001$), varying from 0.83% at 1 to 5.14% at 8. In addition, subjects made more errors when the number of distractors was larger ($F(2,18)=5.65, p<.05$). On average subjects made 0.92% errors when there were no distractors, 2.46% when there were 12, and 3.79% when there were 20. This effect was stronger in the *Conjunction* condition than the *Disjunction* condition ($F(2,18)=7.59, p<.005$ for the interaction); in the *Conjunction* condition the error rate rose from 0.67% to 3.17% to 6.25% as the number of distractors increased whereas in the *Disjunction* condition the error rate was 1.17% with no distractors, 1.75% with 12 distractors and 1.3% when there were 20 distractors. All other effects were not significant, with calculated F values less than 1.

Recall, however, that subjects performed two sessions in the *Disjunction* condition. If the error rate between the two sessions is compared there are no significant effects, except for the effect of number ($F(7,63)=3.91, p<.001$). Thus, in the *Disjunction* condition subjects typically made more mistakes when they were required to count higher numbers of targets. There was little evidence that the subjects' accuracy improved with practise ($F(1,9)= 0.23, p > .1$). The number of distractors also had little effect on the accuracy in *Disjunction* condition ($F(2,18)=0.8, p < .1$); subjects were almost as accurate when there were 20 green horizontal lines to ignore as when there was none, 1.54% errors as opposed to 0.96%. None of the remaining effects are approach significance either. (1)

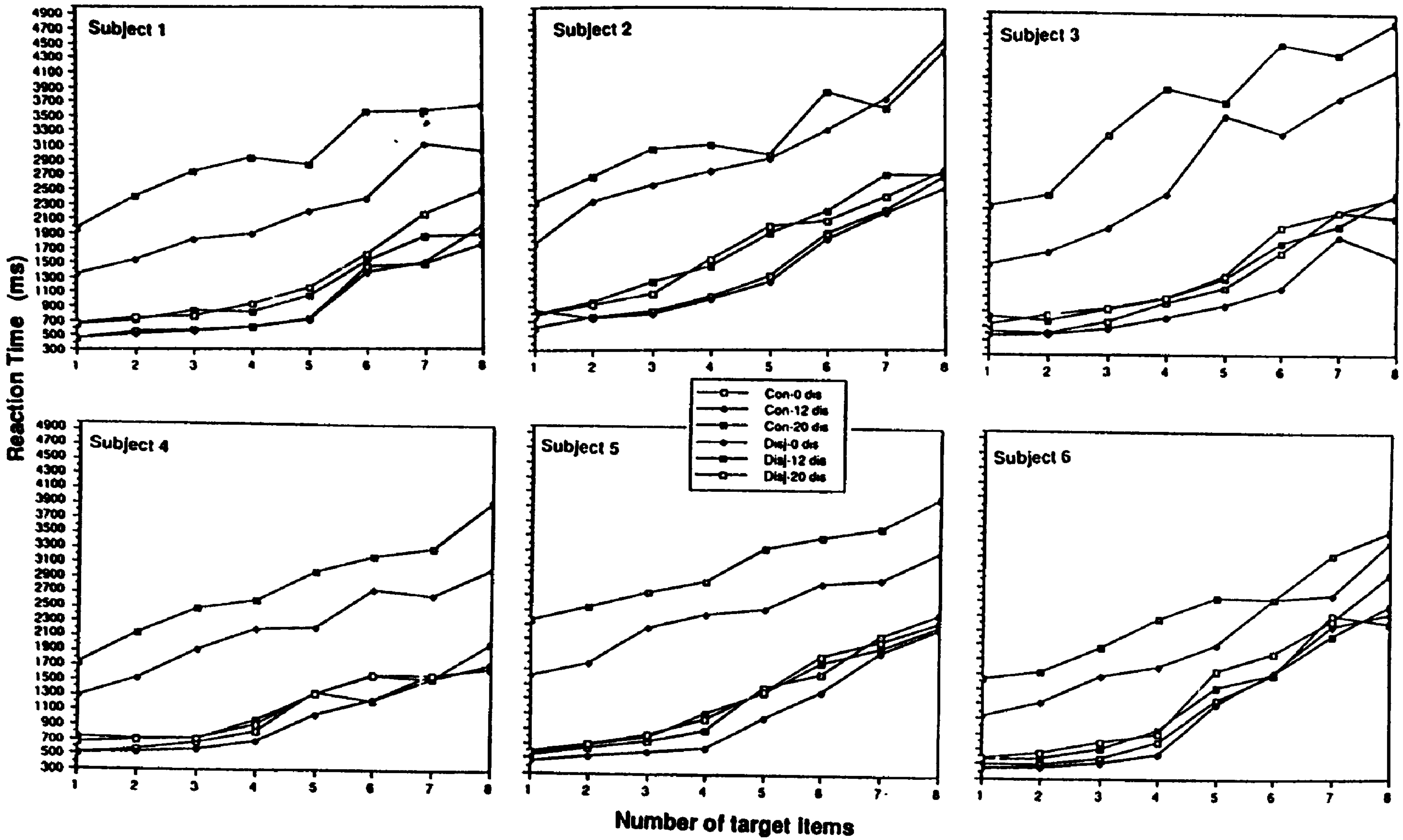

Figure H. Average response latencies for counting White lines, Vertical lines (Disjunction) or White Vertical lines (Conjunction) in a background varying in number of distractors.

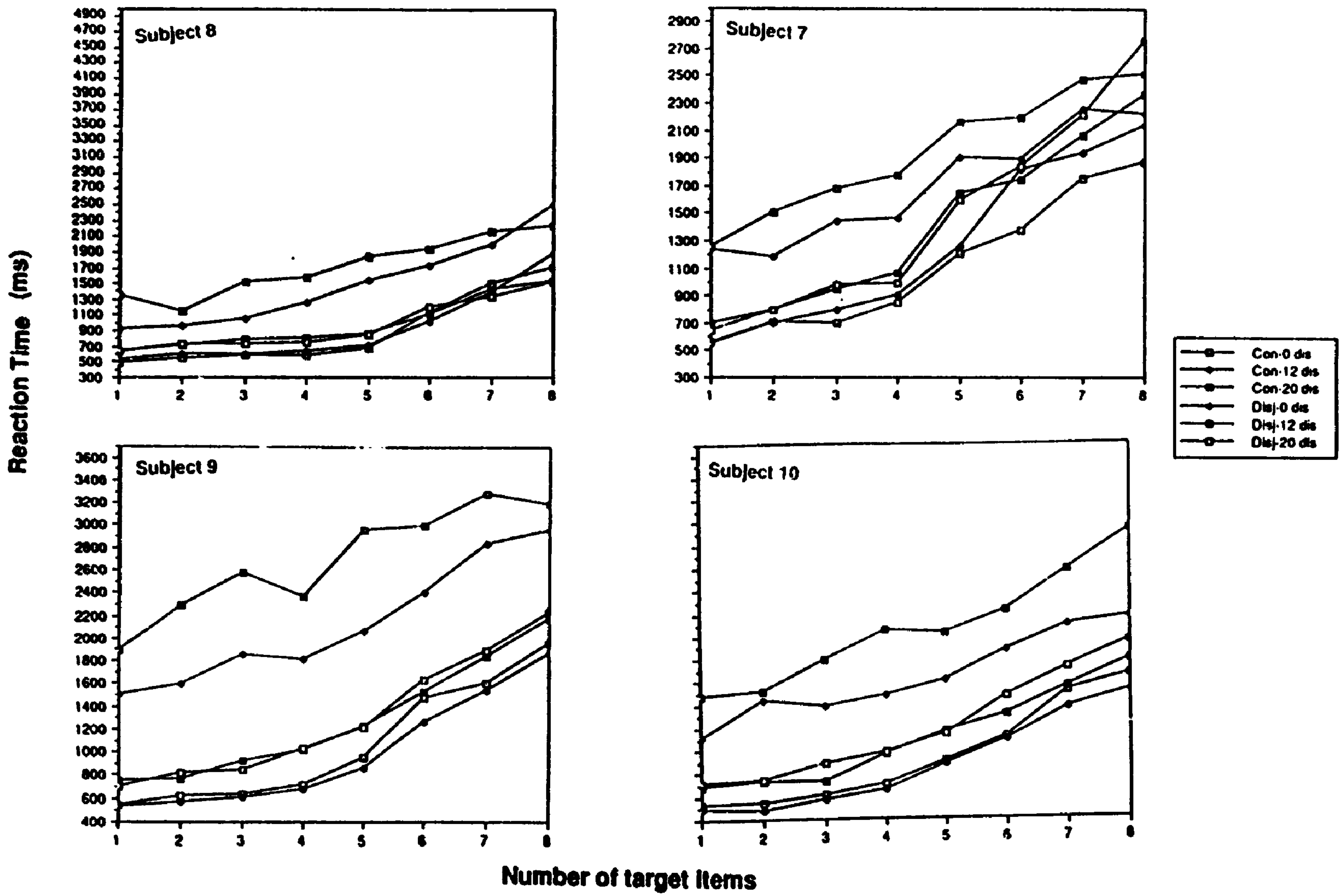

Figure H, continued.

Bibliography

Allport, A. (1975). Critical notice: The state of cognitive psychology. Quarterly Journal of Experimental Psychology, 27, 141 - 152.

Akin, O. & Chase, W. (1978). Quantification of three-dimensional structures. Journal of Experimental Psychology: Human perception and performance, 4 (3), 397 - 410.

Aoki, T. (1977). On the counting process of patterned dots. Tohoku Psychologica Folia, 36, 15 - 22.

Atkinson, J., Campbell, F. & Francis, M. (1976a). The magic number 4 $\pm$ 0: A new look at visual numerosity judgements. Perception, 5, 327-334.

Bashinski,H. & Bacharach, V. (1980). Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. Perception and Psychophysics, 28, 241-248.

Beck, J. (1982). Textural segmentation. In J. Beck (ed.), Organization and representation in perception. Hillsdale, NJ: Erlbaum.

Beck, J., Prazdny, K., and Rosenfled, A. (1983). A theory of textural segmentation. Human and Machine Vision, 1-38.

Beckwith, M. & Restle, F. (1966). Process of enumeration. Psychological Review, 73, 437-444.

Biederman, I. (1988). Aspects and extensions of a theory of human image understanding. In Z. Pylyshyn (ed.), Computational processes in human vision: An interdisciplinary perspective. Norwood, NJ: Ablex publishing corporation.

Briand, K. & Klein, R. (1987). Is Posner's "Beam" the same as Treisman's "Glue"?: On the relation between visual orienting and feature integration theory. Journal of Experimental Psychology: Human Perception and Performance, 13 (2), 228-241.

Burkell, J. (1986). Perception of isoluminant colour changes. Unpublished Masters thesis.

Chi, M., & Klahr, D. (1975). Span and rate of apprehension in children and adults. Journal of Experimental Child Psychology, 19, 434-439.

Egeth, H., Virzi, R., and Garbart, H. (1984). Searching for conjunctively defined targets. Journal of Experimental Psychology: Human Perception and Performance, 10, 32 - 39.

Ericksen, C. & Hoffman, J. (1972). Temporal and spatial characteristics of selective encoding from visual displays. Perception and Psychophysics, 12 (2b), 201-204.

Ericksen, C. & Hoffman, J. (1973). The extent of processing of noise elements during selective encoding of visual displays. Perception and Psychophysics, 14 (1), 155-160.

Ericksen, C. & Schultz, D. (1977). Retinal locus and acuity in visual information processing. Bulletin of the Psychonomic Society, 9, 81-84.

Ericksen, C. & St. James, J. (1986). Visual attention within and around the field of focal attention: A Zoom lens model. Perception and Psychophysics, 40 (4), 225 - 240.

Ericksen, C. & Yeh, Y. (1985). Allocation of attention in the visual field. Journal of Experimental Psychology: Human Perception and Performance, 11, 583-597.

Folk, C., Egeth, H., & Kwak, H. (1988). Subitizing: Direct apprehension or serial processing? Perception and Psychophysics, 44 (4), 313 - 320.

Fodor, J. (1983). Modularity of Mind. Cambridge, Mass.: Bradford Books.

Fodor, J. & Pylyshyn, Z. (1981). How direct is visual perception: Some reflections on Gibson's "ecological" approach. Cognition, 9, 139 - 196.

Francolini, C. & Egeth, H. (1979). Perceptual selectivity is task dependent: The pop-out effect poops out. Perception and Psychophysics, 25 (2), 99 - 110.

Francolini, C. & Egeth, H. (1980). On the nonautomaticity of "automatic" activation: Evidence of selective seeing. Perception and Psychophysics, 27 (4), 331-342.

Frick, R. (1987). The homogeneity effect in counting. Perception and Psychophysics, 41 (1), 8 - 16.

Hamilton, W. (1860). Lecture XIV, Consciousness: Attention in general. Lectures on metaphysics. (Vol. 1). Boston: Gould and Lincoln.

Howe, E. & Jung, K. (1987). Judgment of numerosity: Effect of symmetry and goodness of dot patterns. Acta Psychologica, 64 (1), 3 - 12.

James, W. (1890). Attention. The principles of psychology. (Vol. 1). New York: Dover Publications.

Jevons, W. (1871). The power of numerical discrimination. Nature, 3, 281-282.

Jolicoeur, P. (1988). Curve tracing operations and the perception of spatial relations. In Z. Pylyshyn (ed.). Computational processes in human vision: An interdisciplinary approach. (pp. 133 - 168). Norwood, NJ: Ablex Publishing.

Jolicoer, P., Ullman, S. , and Mackay, L. (1986). Curve tracing operations and the perception of spatial relations. Memory and Cognition, 14, 129-140.

Jonides, J. (1980). Towards a model of the mind's eye. Canadian Journal of Psychology, 34, 103-112.

Jonides, J. and Yanstis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. Perception and Psychophysics, 43 (4), 346 - 354.

Julesz, B. (1980). Spatial nonlinearities in the instantaneous perception of textures with identical power spectra. Philosophical Transactions of the Royal Society, London, Series B, 290, 91-94.

Julesz, B. (1984). Towards an axiomatic theory of preattentive vision. In G. Edelman, W. Gall, & W. Cowan (eds). Dynamic aspects of neocortical function. Toronto: John Wiley & sons.

Kahneman, D. & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman and D. Davies (eds.) Varieties of attention. (pp 29-61). Orlando, FL: Academic.

Kanwisher, N. (1987). Repetition blindness: Type recognition without token individuation. Cognition, 27, 117-143.

Kaufman, E. Lord, M., & Reese, T. (1949). The discrimination of visual number. American Journal of Psychology, 62, 498 - 525.

Klahr, D. (1973a). Quantification processes. In W.G. Chase (Ed.) Visual information processing. New York: Academic Press. pp. 3-34.

Klahr, D. (1973b). A production system for counting, subitizing and adding. In W.G. Chase (Ed.) Visual information processing. New York: Academic Press. pp. 527-546.

Klahr, D. & Wallace, J. (1976) Cognitive development: An information processing view. Hillsdale, NJ: Lawrence Erlbaum Assoc.

Laberge, D. (1983). Spatial extent of attention to letters and words. Journal of Experimental Psychology: Human Perception and Performance, 9, 371-379.

Laberge, D. & Brown, V. (1986). Variations in size of the visual field in which targets are presented: An attentional range effect. Perception and Psychophysics, 40 (3), 188 - 200.

Liss, P. & Reeves, A. (1983). Interuption of dot processing by a backward mask. Perception, 12, 513-529.

Logie, R. & Baddeley, A. (1987). Cognitive processes in counting. Journal of Experimental Psychology, 13 (2), 310 - 326.

Lorinstein, B. & Haber, R. (1975). Perceived numerosity: An information processing analysis. Canadian Journal of Psychology, 29 (3), 224-236.

Mandler, G. & Shebo, B. (1982). Subitizing: An analysis of its component processes. Journal of Experimental Psychology: General, 111 (1), 1-22.

Marr, D. (1982). Vision. San Francisco: W.H. Freeman and Company.

Miller, G. (1956). The magical number 7 ± 2: Some limits on our capacity to store information. Psychological Review, 63, 81-97.

Milligan, G., Wong, D., & Thompson, P. (1987). Robustness properties of nonorthogonal analysis of variance. Psychological Bulletin, 101 (3), 464-470.

Minsky, M. & Papert, S. (1969). Perceptrons: An introduction to computational geometry. Cambridge, Massuchusetts: M.I.T. Press.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. Cognitive Psychology, 9, 353-383.

Oyama, T., Kikuchi, T., and Ichihara, S. (1981). Span of attention, backward masking and reaction time. Perception and Psychophysics, 29 (2), 106-112.

Parkman, J., and Groen, G. (1971). Temporal aspects of simple addition and comparison. Journal of Experimental Psychology, 89 (2), 335 - 342.

Pashler, H. (1987). Detecting conjunctions of colour and form: Reassessing the serial search hypothesis. Perception and Psychophysics, -' (3), 191-201.

Posner, M. (1978). Chronometric explorations of mind. Hillsdale, NJ: Erlbaum.

Posner, M., Nissen, M. & Ogden, W. (1978). Attended and unattended processing modes: the role of set for spatial location. In H. Pick & I. Saltzman (Eds.) Modes of per-eiving and processing information. Hillsdale, NJ: Lawrence Erlbaum.

Posner, M., Snyder, C. & Davidson, B. (1980). Attention and detection of signals. Journal of Experimental Psychology: General, 109, 160-174.

Potter, M., and Levy, E. (1968). Spatial enumeration without counting. Child development, 39, 265-272.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. Cognition, 32, 65 - 97.

Pylyshyn, Z. & Storm, R. (1988). Tracking multiple independent targets: Evidence for both serial and parallel stages. Spatial vision, 3 (3), 179 - 197.

Rock, I. (1983). The logic of perception Cambridge, Mass.: Bradford Books.

Sagi, D. & Julesz, J. (1984). Detection vs discrimination of visual orientation. Perception, 619-628.

Saltzman, I. & Garner, W. (1948). Reaction time as a measure of the span of attention. Journal of Psychology, 25, 227-241.

Shaw, M. (1984). Division of attention among spatial locations: A fundamental difference between detection of letters and detection of luminance increments. In H. Bouma and D. Bouwhuis (Eds.) Attention and performance X: Control of language processes (pp 109 -121). Hillsdale, NJ: Erlbaum.

Shrager, J., Klahr, D., & Chase, W. (1982). Segmentation and quantification of random dot patterns. Paper presented to the 23 annual meeting of the Psychonomics Society.

Siegler, R. (1987). The perils of averaging data over strategies: An example from children's addition. Journal of Experimental Psychology: General, 116, 250-264.

Simons, D. & Langheinrich, D. (1982). What is magic about the magical number four? Psychological Research, 4, 283-294.

Starkey, P., & Cooper, R. (1980). Perception of number by human infants. Science, 210, 1033-1035.

Svenson, O., & Sjoberg, K. (1983). Speeds of subitizing and counting processes in different age groups. Journal of Genetic Psychology, 142, 203-211.

Taubman, R. (1950a). Studies in judged number: I. The judgment of auditory number. Journal of General Psychology, 43, 167-194.

Taubman, R. (1950b). Studies in judged number: II. The judgment of visual number. Journal of General Psychology, 43, 195-219.

Treisman, A. (1985). Preattentive processing in vision. Computer vision, graphics, and image processing, 31, 156-177.

Treisman, A. (1988). Features and objects: The Fourteenth Bartlett Memorial Lecture. Quarterly Journal of Experimental Psychology, 40a (2), 201-237.

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., Kahneman, D. & Burkell, J. (1983). Perceptual objects and the cost of filtering. Perception and Psychophysics, 33, 527-532.

Treisman, A. & Paterson, R. (1984). Emergent features, attention and object perception. Journal of Experimental Psychology: Human Perception and Performance, 10, 12-31.

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 diagnoistic for the preattentive processing of separable features. Journal of Experimental Psyhcology: General, 114, 285-309.

Treisman, A., Sykes, M. & Gelade, G. (1977). Selective attention and stimulus integration. In S. Domic (Ed.) Attention and Performance, V. Hillsdale, N.J.:Erlbaum.

Trick, L. (1987). Subitizing and canonical pattern. Paper presented at annual meeting of the Canadian Psychological Association. Vancouver, B.C.

Trick, L. & Pylyshyn, Z. (1988). When subitizing fails: The importance of preattentive item indexing for subitizing. Cogmem, 35. Centre for Cognitive Science. University of Western Ontario.

Tsal, Y. (1983). Movements of attention across the visual field. Journal of Experimental Psychology: Human Perception and Performance, 9, 523-530.

Ullman, S. (1979). The interpretation of visual motion. Cambridge, Mass: MIT press.

Ullman, S. (1984). Visual routines. Cognition, 18, 97-159.

Van Oeffelen, M. & Vos, P. (1982). Configurational effects on the enumeration of dots: Counting by groups. Memory and Cognition, 10 (4), 396-404.

Warren, H. (1897). The reaction time of counting. Psychological Review, 4, 569 - 591.

Wolfe, J., Cave, K., and Franzel, S. (1987). A modified Feature Integration model for visual search. Unpublished manuscript. Department of Brain and Cognitive Science. Submitted.

Woodworth, R. & Schlosberg, H. (1954). Experimental Psychology, (pp. 96-100). New York: Holt, Rinehart and Winston.

Wright, R. (1989). The perception of inside/outside spatial relations. Unpublished doctoral dissertation.

Zeki, S. (1978). Functional specialization in the visual cortex of the rhesus monkey. Nature, 274, 423-428.

Trick, L. and Pylyshyn, Z. (1989). Preattentive indexing and visual counting: FINSTs and the enumeration of concentric items. Published in the Proceedings from the Eleventh Annual Conference of the Cognitive Science Society, August, 1989, Ann Arbor.

Trick, L. (1989) Parallel and serial processing in visual analysis. Poster presented at the Canadian Institute for Advance Research graduate student conference, June, 1989, London, Ontario.

Trick, L. and Pylyshyn, Z. (1989). Preattentive item indexing and subitizing. Poster presented at the annual meeting of the Canadian Psychology Association, June, 1989, Halifax.

Trick, L. and Pylyshyn, Z. (1988). Spatial attention and subitizing: An investigation of the FINST hypothesis. Published in the Proceedings from the Tenth Annual Conference of the Cognitive Science Society, August, 1988, Montreal.

Trick, L. and Pylyshyn, Z. (1988). When subitizing fails: The importance of preattentive item indexing on subitizing. Cogmem, 35, University of Western Ontario, Centre for Cognitive Science.

Trick, L. (1988).Spatial attention and counting. Poster presented at the annual meeting of the Canadian Psychology Association, June, 1988, Montreal.

Trick, L. (1987).Subitizing and counting: The effect of item configuration. Paper presented at the annual meeting of the Canadian Psychology Association, June, 1987, Vancouver.

Trick, L. (1987). Three studies about subitizing and counting. Poster presented at the Canadian Institute for Advanced Research graduate student conference, May, 1987, York University, Toronto.