PREATTENTIVE INDEXING AND VISUAL COUNTING: FINSTS AND THE ENUMERATION OF CONCENTRIC ITEMS

LANA TRICK AND ZENON PYLYSHYN
UNIVERSITY OF WESTERN ONTARIO

According to Pylyshyn's FINST hypothesis(in press) distinctive feature clusters, locations that pop out in search tasks (Treisman and Gellade, 1980), compete for a small number of spatial reference tokens or FINSTS. A FINST individuates a given feature cluster, making it distinct from others. Once assigned, a FINST remains attached to its respective cluster despite changes in the cluster's position. These spatial indices allow attentive processes to access selected clusters for further analysis though the retinal position of the cluster change through object or eye movement. It is argued that *subitizing*, the rapid apprehension of number in the 1-4 range, exploits this preattentive mechanism; therefore, subitizing should not be possible when the counting task is embedded in one in which subjects have to compute a spatial relation that requires the attentional focus, such as inside(Ullman,1984). Subjects were required to count either concentric rectangles, which implicitly entails computing the inside relation, or count rectangles of uniform or varying sizes spread across the screen. Trend analysis of the counting latencies revealed no sign of subitizing when subjects were required to count items that were one inside another, although subitizing emerged as usual in the other two conditions.

In constructing robots that can move and manipulate objects in a complex, dynamic visual world we are faced with a problem. The properties of objects in an image change from moment to moment as a result of changes in the object or in the object's projection or lighting: an objects's position in the visual field changes with object or camera movement; its projected shape may change if it changes shape or rotates in depth; its projected size may change as it grows or shrinks, approaches or retreats. How can an item be individuated from others so that it can retain its identity despite changes in its properties? People routinely solve this problem; we can easily pick out an item in an image and then move our retinal or attentional focus towards it, compensating even as the item changes position, shape or size. Moreover, we can do this even if we do not recognise the object. We appear to be able to keep track of objects automatically and effortlessly. This ability has been called *indexing*(Ullman,1984). The ability to index is prerequisite for visual motor coordination; we could not touch or capture the things we want to manipulate, or dodge the things we want to avoid if we could not distinguish a particular item from the rest, focus on it and keep track of it. The ability to index is also prerequisite for object description because spatial attention is thought necessary for combining features(Treisman & Gellade, 1980) and computing spatial relations between parts(Ullman, 1984). Without the ability to index a location it would not be possible to move the attentional focus to where we want it to go, since in order to move attention to a particular point we must first be able to specify *which* point. The objective of this research is to understand indexing by studying visual counting, a process that by its nature requires item individuation. First, however, it is necessary to discuss counting in the context of visual processing in general.

Visual processing is thought to have two stages. The first is an automatic *preattentive stage* that employs local parallel operations to derive features (eg. color, line orientation, depth). The second is a goal driven *attentive* stage that employs a serial spatial processing focus to combine features at a location (eg. combine "red" and "vertical" for a red vertical line, Treisman & Gellade,1980), and derive global relations such as inside and connected that cannot be computed by local parallel units(Ullman,1984). Coordination of parallel (preattentive) and serial (attentive) stages is assumed to involve a bottleneck; a small number of preattentively derived feature clusters must be individuated, assigned unique internal reference tokens (FINSTs, Pylyshyn, in press) so they can serve as destinations for the attentional focus. FINSTs, short for FINgers of INSTiation, provide a way of indexing, "pointing to", a cluster without specifying retinal coordinates or properties so that the cluster's identity could be preserved though the cluster moved and changed. Although items may be FINSTed automatically or in response to goals, according to the theory only FINSTed locations can be accessed by attentional or motor commands. A small number of feature clusters can be FINSTed simultaneously; for example, there is evidence that up to 5 independently moving targets can be tracked at once in a field of identical moving distractors(Pylyshyn & Storm, in press). The FINST mechanism is thus parallel but limited capacity; there are only a small number of reference tokens or FINSTs. Differential processing of small and large numbers of items is hence predicted. For this reason the research on subitizing and counting is important. It has been suspected for over a hundred years that the enumeration of small numbers of items employs different processes than the enumeration of large numbers(Jevons, 1871). *Subitizing*, the process of enumerating up to 4 items, is rapid (60 msec/item), effortless and accurate; *counting proper*, the process of enumerating more than 4 items, is slow (300 msec/item) effortful and error-prone. The question remains, however: Why are there two enumeration processes? Why can't we subitize any number of items? I would like to argue that subitizing is parasitic on FINSTS, the limited capacity preattentive indexing system, whereas counting proper involves moving the attentional focus as suggested by Ullman(1984).

One way to support this contention is to show that subitizing of small numbers of items cannot be accomplished in situations in which spatial attention is required to distinguish one item from others. One such situation is when subjects are required to count concentric items, items that are one inside another. 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(Marr, 1982), primarily proximity in this case. When objects are spread across the screen, as in the most counting experiments, these groupings 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 rectangles 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 would also have the same orientation. Thus, there would be a tendency to group the wrong contours on the basis of both the proximity and similarity. Attention would be required to properly establish which edges belong to which objects. 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. Regardless, subjects need to move the attentional focus in order to count concentric objects. Consequently, subitizing should not be possible in this situation.

Given this prediction it is interesting that one of the few studies that failed to produce evidence of subitizing had subjects counting concentric circles (Saltzman and Garner, 1948). The characteristic "bend" in the reaction time curve caused by the change in slope after four, the trademark of the shift from subitizing to counting, was not evident in this study. Unfortunately, trend analysis was not in use at the time and the authors had different interests; this result was not pursued. At this point it is necessary to replicate their finding and establish why subitizing was not evident in their study. Their concentric circle task differed from typical dot enumeration tasks in three ways. First, subjects were presented with objects, circles, instead of points of light. Second, these objects were of different sizes. Third, the objects were concentric. I would like to argue that it is the fact that the items had a common centre, rather than that they were objects of different sizes, that made subitizing impossible in their study.

There were three conditions in the experiment. In the *Same size* 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 than the others. Finally, in the *Concentric* condition subjects were required to count concentric rectangles, thus implicitly computing the inside relation, which requires the focus of spatial attention, according to Ullman(1984). If subitizing is only possible when items were spread across the screen then there should be evidence of slope discontinuities between the 1-4 and 5-8 ranges in the *Same size* and *Different size* conditions but not the *Concentric* condition.

METHOD

Subjects

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

Apparatus and Materials

An Apple II+ computer was used to generate the displays and record the data. Oral response latencies were measured using a Gerbrands G1341 voice activated relay.

Displays were comprised of up to eight 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 .26 X .16, .60 X .42, or 1.01 X .78 degrees visual angle. Rectangles could be located in any of 24 positions. The closest horizontal and vertical neighbours were 1.2 and .94 degrees away from each other, respectively. The minimal distance between diagonal neighbours was .18 degrees, however. The maximal distance between items was 8.33 degrees for small squares in diagonal corners. At 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 .26 X .16 to 7.25 X 5.71 degrees visual angle. For the inner six rings the minimal distance between items was .21 degrees horizontal and .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 .29 and .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.[1]

Procedure

The experiment was conducted in a slightly darkened room. Subjects were seated 110 cm from a video screen, with a computer keyboard within easy reach. Their task was to say 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 also given 24 practise trials.

---

[1] Recently this experiment has been replicated with an increase in the number of ring sizes, and thus maximal distance between contours, in the *Concentric* condition, and a decrease in the inter-item distances for the *Same size* and *Different size* conditions. The same basic results obtained although there was a relative inflation in the time to count 1 item in the *Concentric condition*.

RESULTS

The counting latency data was analyzed in three ways. First, analysis of variance was performed in order to determine if the configuration of the stimuli, *Concentric* as opposed to *Same size* and *Different size*, had an effect on counting latencies. Second, trend analysis on averaged and individual datasets was done in order to determine if there was evidence of subitizing in the counting latency function for the three conditions. Finally, slopes for the subitizing and counting functions were calculated using regression.

Analysis of variance revealed that condition had an effect ($F(2,22)=114.6, p<.001$) as did number of items ($F(7,77)=261.9, p<.001$). See figure 1. Newman Keuls analysis revealed that latencies for the *Concentric* condition were significantly greater than the other two conditions starting at 2($p<.05$). Finally, there was a significant interaction between condition and number($F(14,154)=13.2, p<.001$), with number having an overall greater effect on latencies in the *Concentric* condition than the other conditions.

Trend analysis over averaged data

The primary difference between the subitizing and counting processes is the speed with which they can be carried out; for subitizing the reaction time increase with number is slight whereas for counting the reaction time increase is substantial. For this reason it is important to look for slope changes in the function that relates counting latencies to the number of items. These slope changes are the principal evidence that different processes are being employed for small than large numbers of items. At points where slope changes occur, trend analysis will register significant deviations from linearity.

As predicted, number seemed to produce a more uniform effect on latencies in the *Concentric* condition than it did in the other two, as would be predicted if subjects could no longer subitize when objects were concentric. In order to determine more precisely if subitizing occured trend analysis was performed on the entire range (1-8) for the three conditions, to find out if significant non-linear trends emerged. If the reaction time function showed no significant deviation from linearity then it was assumed that subitizing did not occur. If there were significant deviations from linearity, however, it was necessary to find out where the trend emerged and if it was in the right direction. The point at which there was an upward turn in the latency curve and the function began to show significant deviations from linearity 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 $F(6,88)=7.1$, $p<.0001$ and $F(6,88)=7.8$, $p<.0001$ for *Same size* and *Different size* conditions respectively as compared to $F(6,88)=1.5$, $p>.05$ for the *Concentric* condition). Given that non-linear trends indicate the change from one enumeration process to another, it would seem that the same enumeration process is being used for both small and large numbers in the *Concentric* condition. In fact, considering the magnitude of the latencies, it seems probable that counting proper is occuring.

Trend analysis over individual datasets

Given that there are individual differences in how high people can subitize(Akin and Chase, 1978), averaging across subjects could obscure slope changes in the latency functions. Consequently, the data for each subject were also analyzed separately. All subjects had non-linear trends in the in *Same size* and *Different size* conditions. (See table 1). Only three of the twelve subjects showed significant trends in the *Concentric* condition, however; this represents a significantly smaller proportion ($\chi^2(2)=6.0, p<.05$) than in the other two conditions. Further analyses were performed on the individual datasets to ascertain where the non-linear trend emerged. For both the *Same size* and *Different size* conditions, most subjects subitized to 4. Of the few that showed non-linear trends in the *Concentric* condition, most subitized to 3.

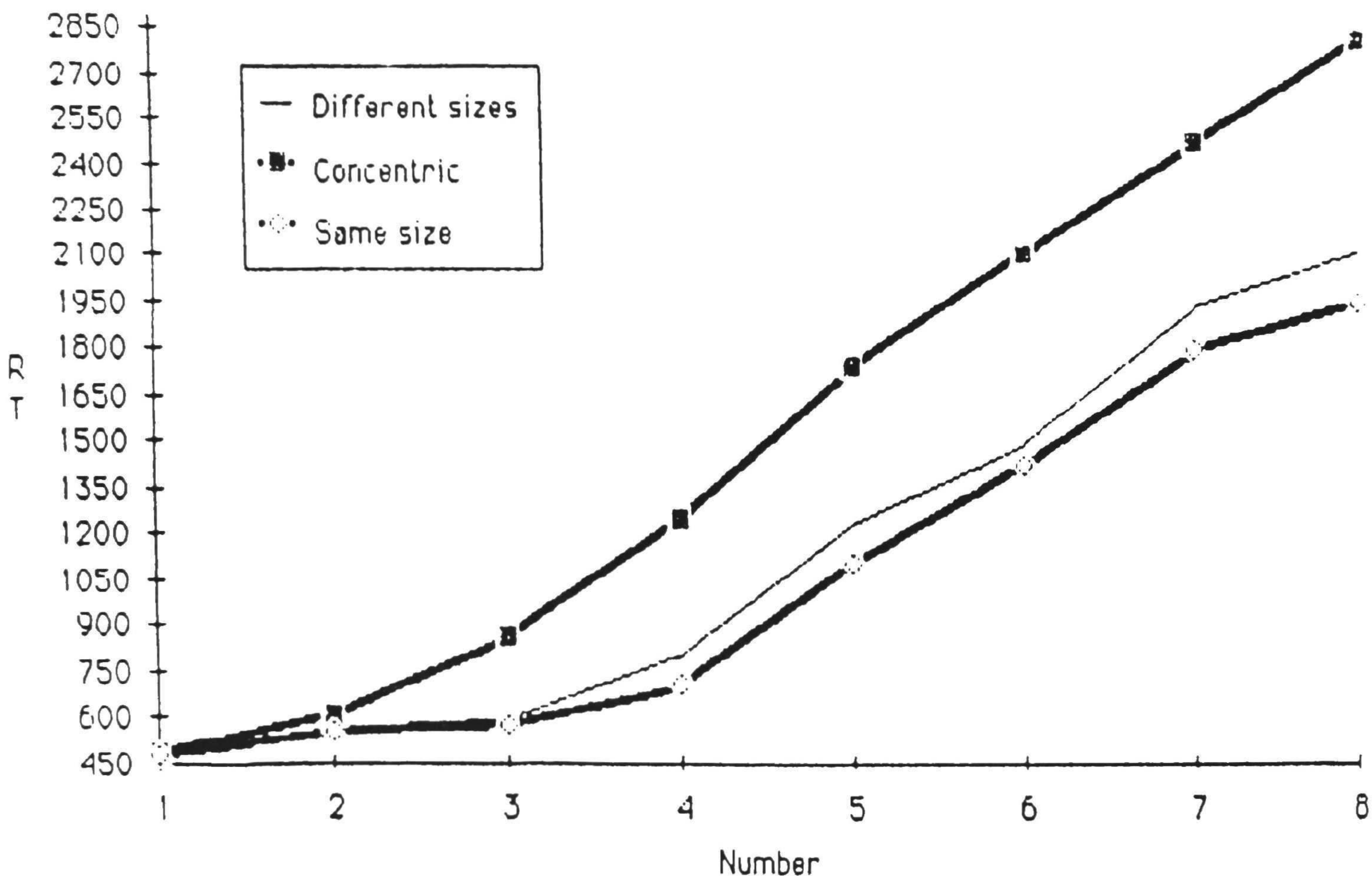

Figure 1: Counting latencies for Concentric study

TABLE 1

Trend analysis of individual datasets

Number of subjects showing evidence of subitizing

Same size	12/12
Different size	12/12
Concentric	3/12

Number of subjects subitizing to each number

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

Regression analysis of averaged counting latencies

	Slope
SUBITIZING RANGE (1-3)	
Same size	55.6
Different size	66.3
Concentric*	198.8
(2-3)	276.2
COUNTING RANGE (5-8)	
Same size	300.7
Different size	330.3
Concentric(1-8)	346.3

*Only the data from subjects who showed evidence of subitizing were included in this analysis(N=3).

Slope analysis

Regression was performed on the averaged data in order to calculate slopes. Although most subjects subitized to 4 subitizing slopes were calculated in the 1-3 range to avoid inflating slopes with latencies from trials in which subitizing did not occur. As can be seen from table 1, the slopes for the 1-3 range were 55.6 and 66.3 for the *Same* and *Different* size conditions, respectively. The 95% confidence intervals for the slopes overlapped in these conditions so there were no significant differences, however. In contrast, for the three subjects that showed evidence of subitizing in the *Concentric condition* the slope in the 1-3 range fell outside these confidence intervals, at 198.8 msec/item. Notice that the slope in the 1-3 range of the *Concentric* condition is somewhat lower than for the 2-3 range; latencies to count 2 in this condition were atypical of the rest of the range. Perhaps subjects are more adept at counting 2 because of frequent exposure to concentric rectangles in objects such as picture frames. Slopes for the 5-8 range for the *Same size* and *Different size* conditions and the 1-8 range in the *Concentric condition* condition are in excess of 300 msec. All slopes fell within each others 95% confidence interval, and differed significantly from the slopes in the non-concentric conditions for the 1-3 range.

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 counting proper. 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 had a common centre, 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. Subitizing was not possible 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, perhaps four for the number of corners, or one for the centre of the radiating pattern.

In contrast, 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 ascertaining the number of assigned reference tokens--performing a FINST role call.

REFERENCES

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

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

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

Pylyshyn, Z. (In press). The role of location indexes in spatial perception: A sketch of the FINST spatial index model. Cognition.

Pylyshyn, Z. & Storm, R. (In press). Tracking multiple independent targets: Evidence for both serial and parallel stages. Spatial vision.

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

Treisman, A. & Gellade, G. (1980). A feature integration model of attention. Cognitive Psychology, 12, 97-136.

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