AD-A278 538

PAGE

2

1 AGEN

3 REPORT TYPE AND DATES COVERED
ANNUAL 1 Mar 93 TO 28 Feb 94

4 TITLE AND SUBTITLE

STOCHASTIC MODELS OF ATTENTION AND SEARCH

5 FUNDING NUMBERS

F49620-92-J-0186
61102F
2313
BS

6 AUTHOR(S)

DR STEVEN YANTIS

7 PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Dept of Psychology
Johns Hopkins University
Baltimore, MD 21234

8 PERFORMING ORGANIZATION REPORT NUMBER

9 SPONSORING MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/NL
110 DUNCAN AVE SUITE B115
BOLLING AFB DC 20332-0001

Dr John F. Tangney

11 SUPPLEMENTARY NOTES

12a DISTRIBUTION AVAILABILITY STATEMENT

Approved for public release;
distribution unlimited

13 ABSTRACT (Maximum 200 words)

Six lines of experimentation have been continued. In the first set of experiments, the PI has tested and rejected a two-process model of visual attention allocation. He has proposed an alternative perceptual sampling model and performed stochastic simulations of the model to show that it can account for certain aspects of human performance in cued visual search tasks. In the second set of experiments, the PI has found evidence that observers perceive occluded objects across time, a finding that complements an analogous ability to perceptually complete partially occluded objects across space. Several lines of experimentation have been carried our using a bistable apparent motion display (the Ternus display) as a tool to explore the assignment of object identity over time. For example, the PI has found evidence that a common mechanism may underlie the perception of bistable apparent motion and the capture of visual attention in certain visual search tasks, and he has discovered that perceptual grouping by proximity can precede the assignment of motion correspondences in bistable apparent motion. In the fifth project, the PI has shown that visual salience is not sufficient to produce attentional capture; a deliberate state of attentional readiness is required to

14 SUBJECT TERMS

15 NUMBER OF PAGES

16 PRICE CODE

17 SECURITY CLASSIFICATION OF REPORT	18 SECURITY CLASSIFICATION OF THIS PAGE	19 SECURITY CLASSIFICATION OF ABSTRACT	20 LIMITATION OF ABSTRACT
(U)	(U)	(U)	(U)

guide attention according to salient stimulus attributes. Finally, in the last set of experiments, the PI is examining the attentional demands of perceptual grouping and visual "pop-out."

Annual Technical Report

Submitted to
Air Force Office of Scientific Research

"Stochastic Models of Attention and Search"

AFOSR Grant F49620-92-J-0186
Reporting Period: 1 Mar 1993 - 28 Feb 1994

Principal Investigator:

Steven Yantis
Department of Psychology
Johns Hopkins University
Baltimore, MD 21234
voice: (410) 516-5328
fax: (410) 516-4478
email: yantis@jhuvms.hcf.jhu.edu

Summary

Six lines of experimentation have been continued. In the first set of experiments, the PI has tested and rejected a two-process model of visual attention allocation. He has proposed an alternative perceptual sampling model and performed stochastic simulations of the model to show that it can account for certain aspects of human performance in cued visual search tasks. In the second set of experiments, the PI has found evidence that observers perceive occluded objects across time, a finding that complements an analogous ability to perceptually complete partially occluded objects across space. Several lines of experimentation have been carried out using a bistable apparent motion display (the Ternus display) as a tool to explore the assignment of object identity over time. For example, the PI has found evidence that a common mechanism may underlie the perception of bistable apparent motion and the capture of visual attention in certain visual search tasks, and he has discovered that perceptual grouping by proximity can precede the assignment of motion correspondences in bistable apparent motion. In the fifth project, the PI has shown that visual salience is not sufficient to produce attentional capture; a deliberate state of attentional readiness is required to guide attention according to salient stimulus attributes. Finally, in the last set of experiments, the PI is examining the attentional demands of perceptual grouping and visual "pop-out."

Status of the Research

This report describes the technical progress achieved in "Stochastic Models of Attention and Search," which is a project supported by AFOSR Grant F49620-92-J-0186.

1. *Tests of a two-process model of visual attention allocation* (Johnson & Yantis, submitted)

I reported on this project last year, and we have carried out additional experiments and simulations in order to explore the properties and consequences of the two process model of visual attention allocation. This work is described in detail in the enclosed manuscript, which has been submitted for publication to *Perception & Psychophysics*.

In the experiments, subjects detected the presence or absence of a prespecified target element appearing in an array of four elements. A spatial cue appeared 200 ms before the onset of the display, indicating a location in the upcoming array. In separate blocks of trails, the cue was 100% valid, 50% valid, or 25% valid with respect to predicting the location of the upcoming target. In the 100% valid condition, the cue was completely predictive of the target's location, and in the 25% valid condition, it was independent of the target's location. In the 50% valid condition, the cue provided an intermediate degree of predictive validity. We were interested in testing an attention allocation model in which subjects could allocate perceptual samples to spatial locations according to the information provided by the cue. According to the model, samples are taken at a location in proportion to the probability that that location contains the target. Evidence about the identity of the element in each location accumulates as a random walk over time, with evidence accumulating more rapidly (i.e., a higher drift rate in the random walk) for locations from which a larger number of perceptual samples have been taken.

The predictions of this parallel model were compared to those of a serial probability matching model (Jonides, 1983). This model holds that observers may select in advance of each trial one location to which they direct their attention, or they may divide their attention over the entire display; subjects' selection strategy is assumed to vary from trial to trial according to the location and predictive validity of the cue. If the cue is 100% valid, observers always attend first to the cued location, and if the cue is only 25% valid, the observers ignore the cue. When the cue is 50% valid, observers attend to the cued location on 50% of the trials and they distribute their attention evenly within the display on the remaining trials. According to the model, the reaction time distribution from the 50% valid condition should be a mixture of the distributions from the 100% valid and 25% valid conditions, respectively. Although the mean RT across conditions was consistent with the mixture prediction, a more complete distributional analysis (Yantis, Meyer, & Smith, 1991) ruled out the mixture hypothesis, thus undermining the viability of the serial probability-matching model. We carried out additional control experiments in which we monitored subjects' eye movements to ensure that the effects we measured were due to covert shifts of attention rather than to overt eye movements.

We assessed the viability of the parallel perceptual-sampling model by running stochastic simulations of the model under the conditions used in the experiments. In the simulations, perceptual samples were allocated to spatial locations according to the validity of the cue, and perceptual evidence accumulated at each location in a random walk until a criterion amount of information for the target was acquired (information can be interpreted as perceptual features of the elements to be identified). The simulations yielded reaction-time distributions whose shapes and relative positions were compared to those obtained from human subjects. The simulations suggest that the parallel perceptual sampling model provides a viable alternative to the serial probability-matching model.

Our most recent experiments involved recording observers' eye position during the experiment to ensure that the effects we observed were due to covert shifts of attention and not overt shifts of the eyes.

2. *Amodal completion across time and space* (Yantis, in preparation)

Another series of experiments that we are currently conducting also concerns the possible role of object representations in the perception of apparent motion in the Ternus display. In a well known perceptual phenomenon known as *amodal completion*, a partly occluded object is perceived as complete. Recent investigations of completion suggest that it is a truly perceptual phenomenon that may rely on some of the same mechanisms as those responsible for illusory contours (e.g., Kellman & Shipley, 1992; Sekuler & Palmer, 1992). The role of amodal completion in perception is to provide a coherent and continuous representation of objects in the world given a discontinuous and fragmentary retinal image (due to occlusion, for example). All previous studies of amodal completion have involved completion in space.

The experiments I have carried out explore the possibility that amodal completion may occur across time as well as across space. I used a bistable apparent motion display first described by Ternus (1939 1926). There are many versions of the Ternus display; I will describe one common one here. Two elements (e.g., disks) are displayed in a subset of three predetermined locations in the visual field (see Figure 1a). The three locations are horizontally arrayed about 1 degree of visual angle apart. In frame 1 of the display, lasting 200 ms, one element appears in the right location and the other object in middle location. In frame 2, also 200 ms in duration, the objects occupy the middle location and the left location. There is a brief blank interval (the interstimulus interval or ISI) between frames.

Usually several such cycles are presented. The Ternus display is perceptually ambiguous: it can be perceived as involving either element motion (the element that begins on the right hops end to end over the stationary element in the middle location) or group motion (the two elements move back and forth together, as a group, so the identity of the element in middle location switches back and forth from one frame to the next). Figure 2 illustrates the motion correspondences under each type of percept.

The probability that observers perceive group motion depends on the magnitude of the ISI (see function labeled "No Occluder" in Figure 3): at short ISIs (less than about 10 ms), element motion is almost always perceived, and at long ISIs (more than about 100 ms), group motion is almost always perceived. The probability of perceiving group motion increases monotonically between 0 and 100 ms with the exact location and slope of the psychometric function depending on factors like the size of the objects, their spatial separation, and their duration.

Observers viewed a standard Ternus display like the one depicted in Fig. 1a on some trials, and a Ternus display like the one shown in Fig. 1b on other trials. In the latter displays, which are termed "occluded Ternus displays," segments of the surrounding rectangle above and below the middle location are removed during the ISI. This produces two vertically oriented illusory contours which support the perception of a "virtual occluder" covering the middle location of the Ternus array.

The probability of group motion reports as a function of ISI for the occluded Ternus display (as in Figure 1b) is shown in Fig. 3 as the function labeled "Occluder"; the function labeled "No Occluder" shows performance on trials in which there was no occluder (as in Figure 1a). The main result is that the location of the transition from element motion perception to group motion perception is centered at about 40 ms for the standard display and much longer for the occluded display.

This finding can be explained by assuming that the occluded element was amodally completed during the ISI when it was occluded (and not, of course, when it was not occluded). In other words, the occluder provided an "explanation" for the disappearance of the middle element: the element did not disappear, it was merely occluded. Consequently, the element that was occluded still "persisted" during the ISI, and its continuity lead to the perception of element motion even at very long ISIs. Because its phenomenal persistence was prolonged, when the center disk physically reappeared, it was perceived as being the same object as the one appearing in that location during Frame 1, even though its physical absence was long enough for any trace of visible persistence to have faded.

I have carried out additional experiments using a virtual occluder presented upon a semitransparent ("paint-splattered") surface floating above the moving elements using stereoscopic viewing techniques. The virtual occluder in these experiments is much more salient in this version of the experiment and provides a more robust demonstration of time-based amodal completion.

3. *Object continuity in visual attention and motion perception* (Yantis & Gibson, in press)

We have recently argued (see Yantis & Hillstrom, 1994) that the appearance of a new perceptual objects in the visual field appears to capture visual attention. Experiments from our lab over the last several years had shown that the abrupt onset of a visual element captures attention; Yantis and Hillstrom, using equiluminant visual elements, found that it was not the luminance increment accompanying these appearances that was responsible for attentional capture, but the appearance of a new perceptual object. In the experiments described below, we sought to further explore the hypothesis that the visual system is predisposed to attend to the appearance of new perceptual objects. The goal is to link the phenomenon of attentional capture to the perception of group and element motion in the Ternus (1939/1926) display.

One can view the perception of the Ternus display as reflecting the parameters that determine what constitutes a perceptual object. If the duration of the ISI is long enough, then the perceptual continuity of the object in the middle location can be disrupted, allowing a change in the identity of that object, which is required for group motion to be perceived (see Figure 2, top). In contrast, if the duration is quite short, then that object's identity remains intact, and it can only be interpreted as a single, coherent, continuously-present perceptual object, leading to the perception of element motion (Figure 2, bottom). If the ISI at which the transition between element-motion perception and group-motion perception in the Ternus display represents the duration required to disrupt the continuity of a perceptual object, then that same interval might also be sufficient, when inserted into an element in visual search, to capture attention (recall that attention is captured when a new perceptual object appears).

To test this idea, we have run several visual search experiments in which one element in the visual display is briefly erased and redisplayed during the trial (Figure 4). The duration of the gap varies over the same range as the ISIs in the Ternus display. Attention is captured in this experiment to the extent that the duration of the gap in the target element disrupts the object representation (see Figure 5, closed symbols and right ordinate): if the duration of the gap is very short, then it does not capture attention (attentional capture is indexed by the relative slope of the display-size function: flat or very shallow slopes indicate complete attentional capture, steep slopes indicate the absence of attentional capture). If the duration of the gap is very long, then it captures attention almost completely (i.e., the slope of the visual search function is near zero when the target is the element exhibiting a gap), and when the duration of the gap has an intermediate duration, then capture is partially successful. The timecourse of attentional capture is similar to the timecourse of the transition from element to group motion perception in the Ternus experiment (Figure 5, open symbols, left ordinate). This result suggests that a common mechanism may underlie attentional capture in my visual-search task and the perception of element and group motion in the Ternus display.

4. *Perceptual grouping and the perception of bistable apparent motion (Ternus) displays* (Kramer & Yantis, submitted)

It has been thought for a long time that apparent motion perception is insensitive to the form of the objects rendering the motion. Recent studies have shown, however, that there are small but measurable effects of element identity on disambiguating apparent motion perception. It is less clear whether configurational or grouping effects significantly influence motion perception; available experimental evidence is equivocal.

In this series of experiments, we examined whether perceptual grouping by proximity could influence the perception of element and group motion in a Ternus display. If figural properties of a display are in fact computed after a motion signal is extracted, then one would not effect perceptual grouping to influence motion perception in a Ternus display. We implemented the test with displays like the one shown in Figure 6. In each panel, two displays are shown, labeled "1" and "2"; these are the two frames of the Ternus display, with a blank interstimulus interval of variable duration placed between them. Panels A and

B show conditions in which the Ternus display is embedded in a context. In Figure 6A, the context produces strong vertical grouping, and in Figure 6B, the context produces strong horizontal grouping. If grouping does affect motion perception here, then we would expect to observe more element motion in conditions like Figure 6A than those like Figure 6B, because the context in 6A should serve to "anchor" the middle points in place. Figures 6C and 6D represent control conditions without context, used to control for variations in horizontal separation in Figures 6A and 6B. If the claim that figural effects are computed after the motion signal is extracted is correct, then we would expect displays like Figure 6A to produce psychometric functions like those in Figure 6C.

Figure 7 shows the results from one representative subject in this experiment. The four functions correspond to the four display types shown in Figure 6. There is a small effect of spacing (compare the top two functions), replicating previous findings. The most dramatic effect, however, is the comparison between the vertical grouping condition in which little group motion is perceived, with the horizontal grouping condition, in which much more group motion is reported. Clearly there is a strong effect of grouping on bistable apparent motion perception. The accompanying manuscript describes our results in detail.

5. *Visual salience and attentional capture* (Yantis & Egeth, 1994)

When observers engage in visual search for a target defined by a salient visual attribute (e.g., the target is red and the nontargets are blue, or the target is a vertical bar and the nontargets are horizontal bars), search is typically extremely efficient (as indexed, for example, by the insensitivity of response time to the number of elements to be searched). This visual "pop-out" phenomenon is widely believed to reflect stimulus-driven or (bottom-up) capture of visual attention by the salient element. Indeed, salience is sometimes *defined* as the ability of an attribute to capture attention. In virtually all studies that have examined this question, however, the salient attribute is explicitly relevant in the task set for the observer; this design feature makes it impossible to know whether attention is captured by the attribute in question, because there is an inherent top-down or intentional component to the distribution of attention in these cases. In this project, we have shown clearly that salient visual attributes (e.g., red among blue) can be used to guide attention efficiently when they are known to be relevant to the task, but when they are irrelevant to the observer's perceptual goals, salient elements *do not* capture attention.

We demonstrated this simple result in a visual search task requiring observers to find a vertical bar among slightly tilted bars (a fairly difficult visual search task). In the "color-singleton" version of the experiment, three groups of subjects participated in three slightly different tasks. The no-singleton group simply searched for the vertical target; all stimuli were blue. In this condition, response time increased linearly with display size (Figure 8). The "target always singleton" group searched for a vertical target and they were (correctly) informed that if the target is present, it would always be the single red bar among multiple blue bars. Response times in this condition were uniformly fast and did not depend on the number of elements in the display (Figure 9), showing that this salient attribute can be used to guide search efficiently. This is what is usually called the "pop-out effect." In the "target 10% singleton" condition, a red singleton was present on every trial, but the target was the singleton on only 10% of the trials; in other words, the singleton was slightly *less* likely to be the target than any of the other elements. If color singletons really capture attention, then response times in this condition when the target happens to be the singleton should be just as fast as in the "target always singleton" condition. Instead, response times to singleton targets were if anything slightly slower than to nonsingleton targets (Figure 10). This result shows that when the salient attribute is not part of the observer's attentional set, it does not have any particular control over attention.

6. Perceptual Grouping and Attentional Capture (Kramer & Yantis, 1994)

Texture segregation and grouping have been assumed to be closely related processes by some researchers (e.g., Beck, Prazdny, & Rosenfeld, 1983; Treisman & Gormican, 1988), while other researchers assume they are qualitatively different (e.g., Ben-Av, Sagi, & Braun, 1992). Some researchers assume texture segregation and grouping are both preattentive (Julesz, 1986), some assume that both require dispersed attention. Ben-Av et al. (1992), on the other hand, set out to show that an aggregative process like grouping requires attention, but that a segregative process like texture segregation does not.

Ben-Av et al. performed a series of experiments in which performance on grouping tasks and segregation tasks were compared under single and dual task conditions. In the dual task condition, subjects performed an attention-demanding primary task requiring identification, along with a grouping or segregation task. The hypothesis was that the attention demanding identification task would significantly affect performance on the grouping task, because this task was assumed to require attention, but that performance on the segregation task would not be impaired, because this task does not require attention. This is exactly what Ben-Av et al. (1992) observed.

Unfortunately, claims that perceptual grouping per se is more or less difficult than texture segregation per se are very difficult to sustain, because it is difficult to equalize the difficult of the tasks used to probe these two psychological processes. We hypothesize that by manipulating the strength of perceptual grouping in tasks similar to Ben-Av et al's we will find that grouping requires attention under difficult grouping conditions, but that grouping does not require attention when grouping is easy. Similarly, by manipulating the difficulty of texture segregation we expect to find that segregation does not require attention when the task is easy, but that it does require attention when the task is hard.

In our experiments we manipulate the strength of grouping by introducing various degrees of rotation and jitter in the elements to be grouped, ranging from no rotation or jitter (easy grouping) to much rotation and jitter (hard grouping). We manipulate the difficulty of texture segregation by varying the similarity between the target and the background texture.

The results of these experiments will reveal whether texture segregation and perceptual grouping are qualitatively different by removing task difficulty as a potentially confounding factor.

References

Beck, J., Prazdny, K., & Rosenfeld, A. (1983). A theory of textural segmentation. In J. Beck & A. Rosenfeld (Eds.), *Human and Machine Vision* (pp. 1-38). New York: Academic Press.

Ben-Av, M. B., Sagi, D., & Braun, J. (1992). Visual attention and perceptual grouping. *Perception & Psychophysics, 52,* 277-294.

Jonides, J. (1983). Further toward a model of the mind's eye's movement. *Bulletin of the Psychonomic Society, 21,* 247-250.

Julesz, B. (1986). Texton gradients: The texton theory revisited. *Biological Cybernetics, 54,* 464-469.

Kellman, P. J., & Shipley, T. F. (1992). Perceiving objects across gaps in space and time. *Current Directions in Psychological Science, 1,* 193-199.

Sekuler, A. B., & Palmer, S. E. (1992). Perception of partly occluded objects: A microgenetic analysis. *Journal of Experimental Psychology: General, 121,* 95-111.

Ternus, J. (1939). The problem of phenomenal identity. In W. E. Ellis (Ed. and Trans.), *A sourcebook of Gestalt psychology* (pp. 149-160). New York: Harcourt Brace and Co. (Original work published 1926)

Treisman, A. & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95,* 15-48.

Yantis, S. & Hillstrom, A. P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant visual objects. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 95-107.

Publications reporting research supported by this grant (*copies enclosed)

*Yantis, S., & Gibson, B. S. (in press). Object continuity in visual attention and motion perception. Special Issue on Visual Attention, *Canadian Journal of Experimental Psychology.*

Yantis, S. (in press). Attention. In R. J. Sternberg (ed.), *Encyclopedia of Human Intelligence.* New York: MacMillan Publishing Co.

Yantis, S. (forthcoming). Attentional capture in visual search. Invited chapter to appear in A. Kramer (Ed.), *Converging Operations in the Study of Visual Selective Attention.* Washington, DC: American Psychological Association.

*Johnson, D. N., & Yantis, S. Allocating visual attention: Tests of a two-process model. Submitted to *Journal of Experimental Psychology: Human Perception and Performance.*

*Kramer, P., & Yantis, S. Figural identity and bistable apparent motion. Submitted to *Perception & Psychophysics.*

Yantis, S. Amodal completion across time and space. To be submitted to *Nature.*

Yantis, S., & Egeth, H. E. Visual salience and attentional capture. In preparation for *Vision Research.*

Collaborators

Howard Egeth, Department of Psychology, Johns Hopkins University.

Douglas N. Johnson, former graduate student. PhD, Johns Hopkins University, 1992. Dissertation entitled "Representation and item-specific learning in lexical decision."

Bradley S. Gibson, postdoc. PhD, University of Arizona, 1992. Dissertation entitled "Representations of shape in memory."

Peter Kramer, graduate student.

Barry Vaughan, graduate student.

Professional Presentations and Interactions

Member, Perception and Cognition Research Review Committee, NIMH, 1993-present.

Yantis, S. (1993, June). Visual attention and perceptual objects. Invited tutorial presented at the SISSA Theoretical Cognitive Neuropsychology Seminar on Attentional Processes and their Disorders (A. Caramazza & T. Shallice, organizers), Trieste, Italy. Conference schedule enclosed.

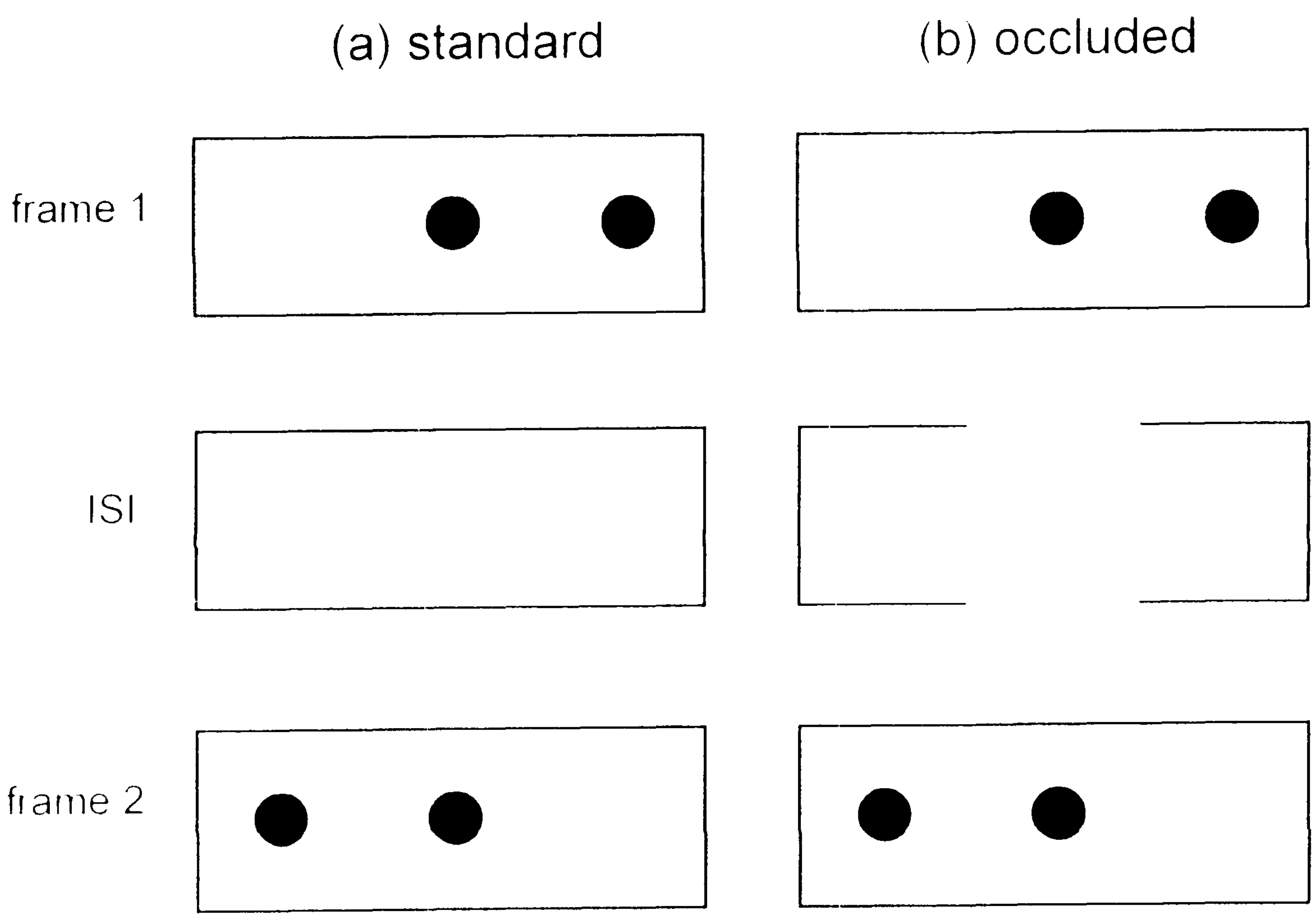
Yantis, S. (1993, November). Object continuity in motion perception and attention. Paper presented at the 34th Annual Meeting of the Psychonomic Society, Washington, DC. Abstract enclosed.

Yantis, S. (1994, January). Perceptual objects and attentional capture. Paper presented at the Annual Interdisciplinary Conference, Jackson Hole, WY. Conference schedule enclosed.

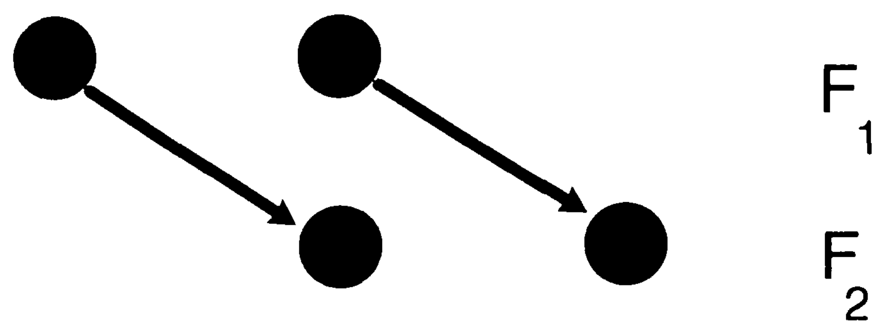
Kramer, P., & Yantis, S. (1994, May). Perceptual grouping and apparent motion. Poster to be presented at the 34th Annual Meeting of the Association for Research in Vision and Ophthalmology, Sarasota, FL. Abstract enclosed.

Yantis, S., & Egeth, H. E. (1994, May). Visual salience and stimulus-driven attentional capture. Poster to be presented at the 34th Annual Meeting of the Association for Research in Vision and Ophthalmology, Sarasota, FL. Abstract enclosed.

Yantis, S. (1994, May). Attentional capture in visual search. Invited tutorial to be presented at "Converging Operations in the Study of Visual Selective Attention" (A. Kramer, organizer), University of Illinois, Champaign, IL. Conference schedule enclosed.


(a) standard
(b) occluded
frame 1
ISI
frame 2

Group Motion Correspondences



Element Motion Correspondences

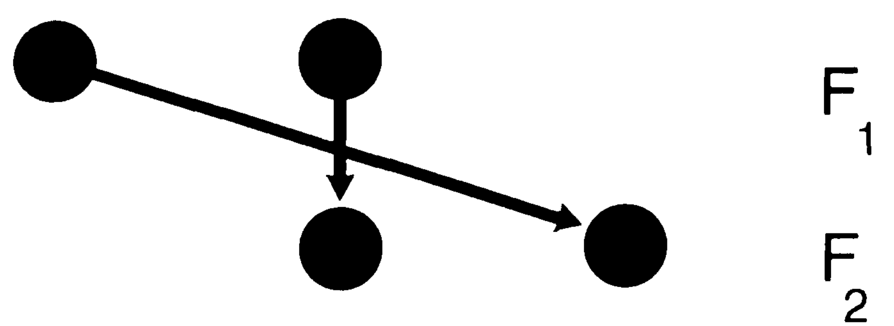
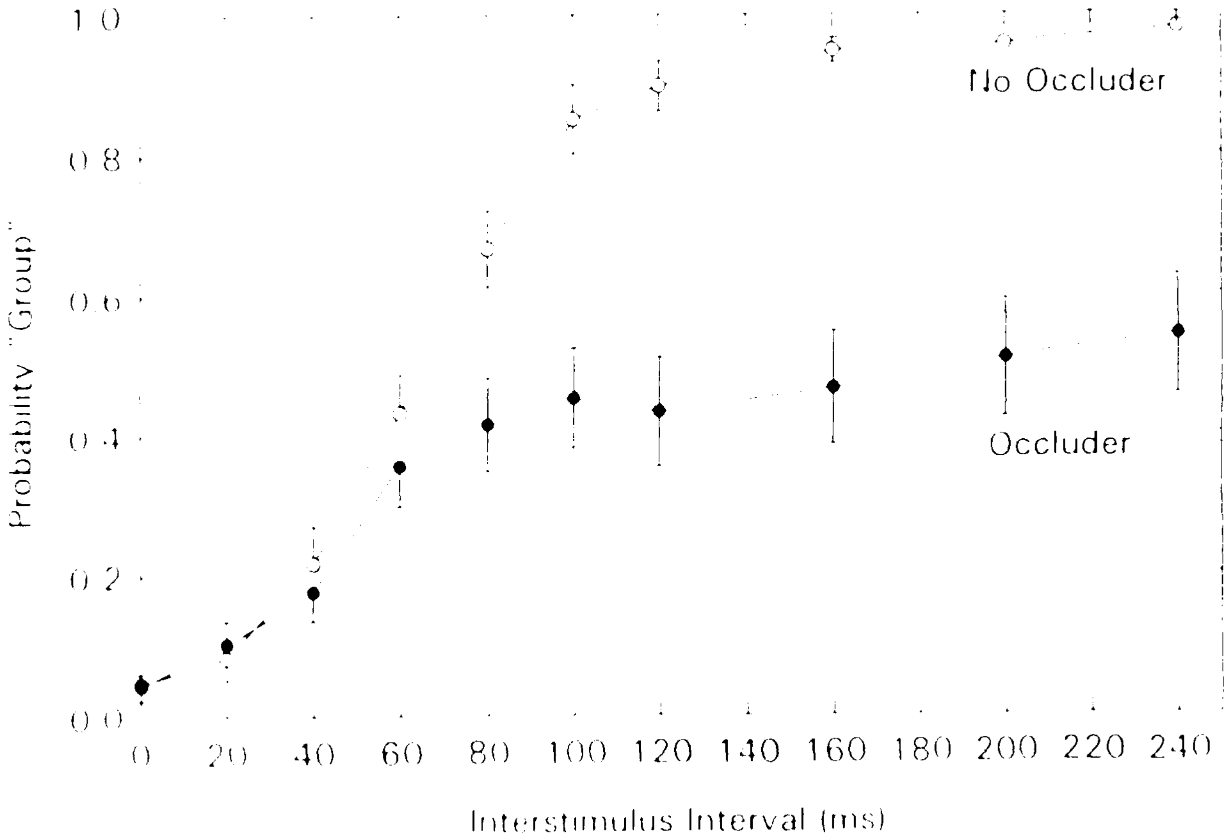
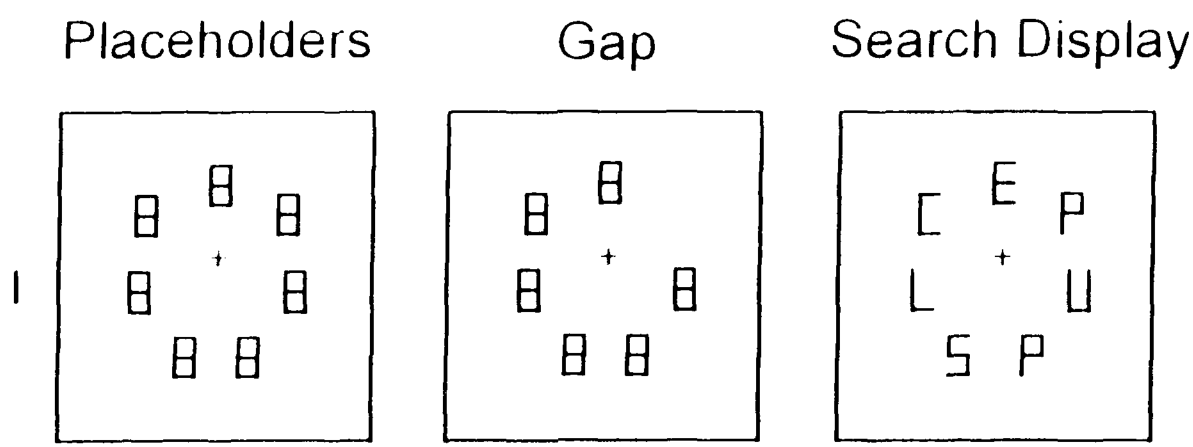
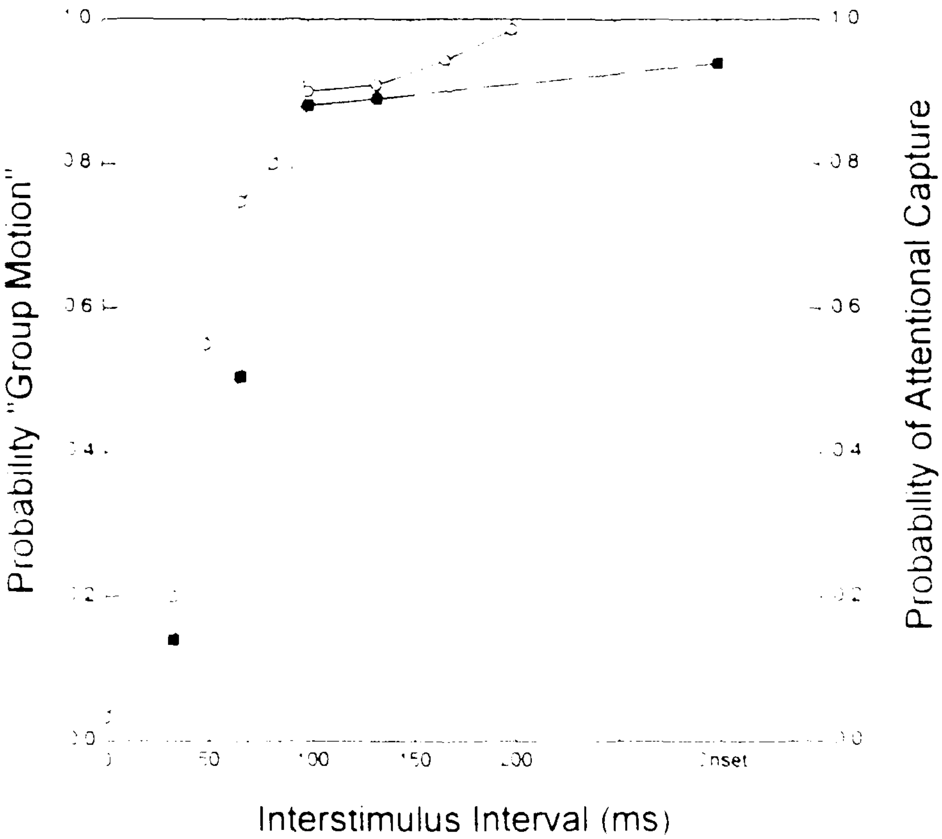


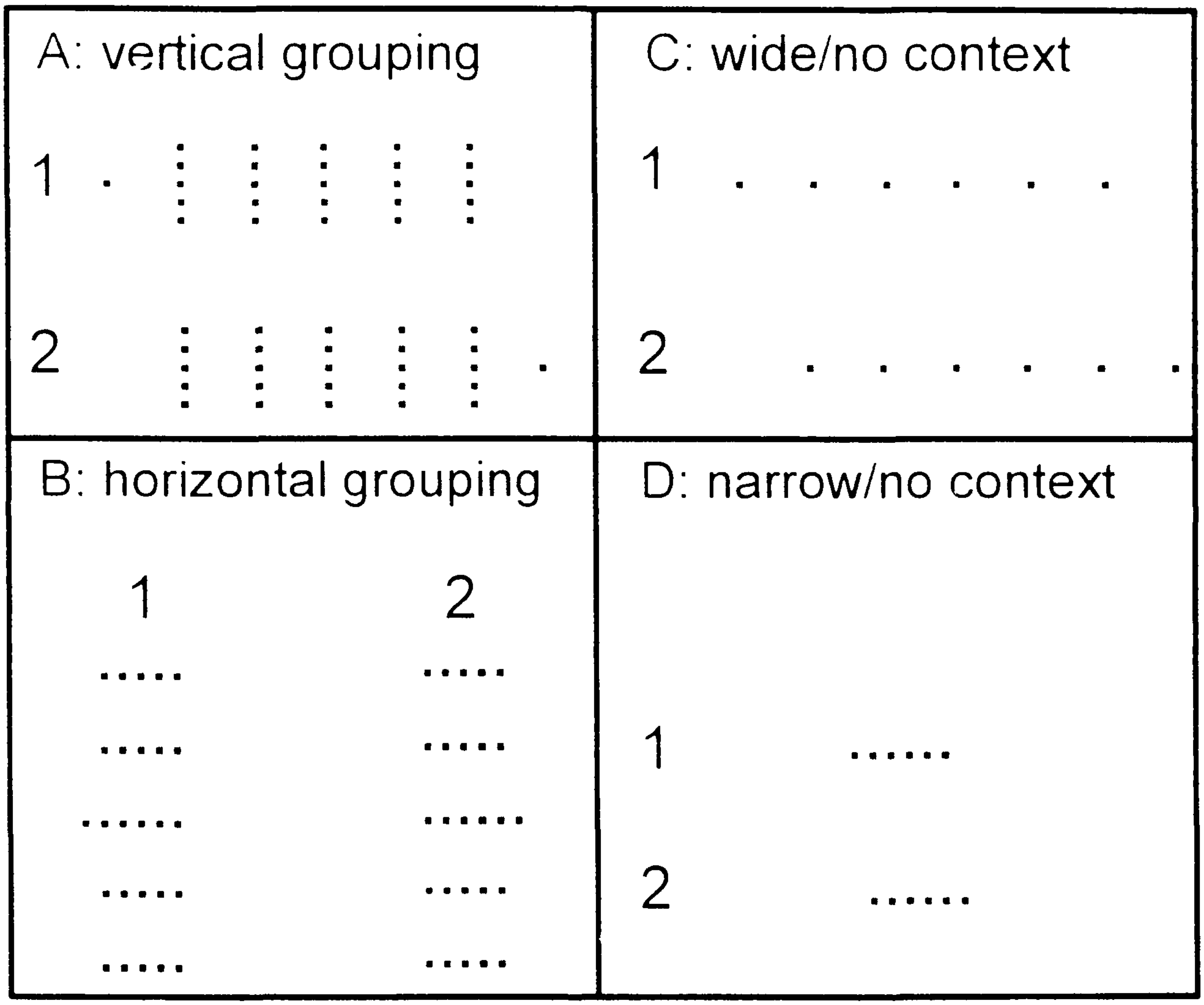
fig 2



15 observers


Placeholders
Gap
Search Display
C E P
L U
S P


Probability "Group Motion"
Probability of Attentional Capture
Interstimulus Interval (ms)
1.0
0.8
0.6


A: vertical grouping
1
2
C: wide/no context
1
2
B: horizontal grouping
1
2
D: narrow/no context
1
2

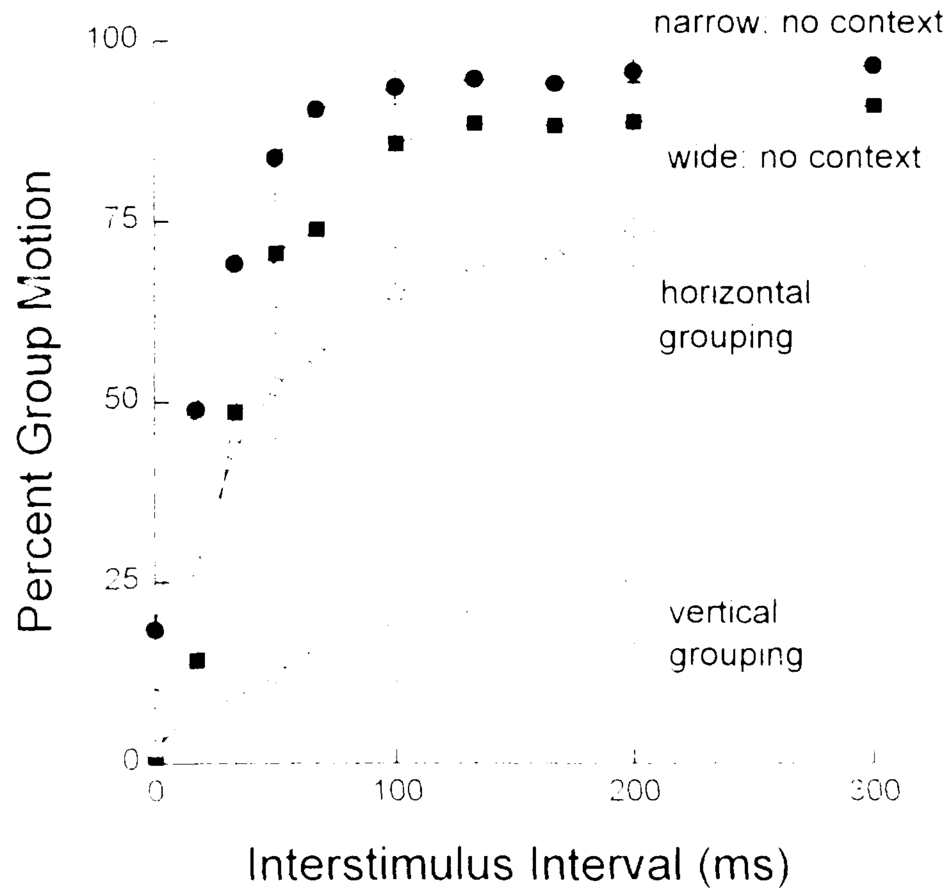
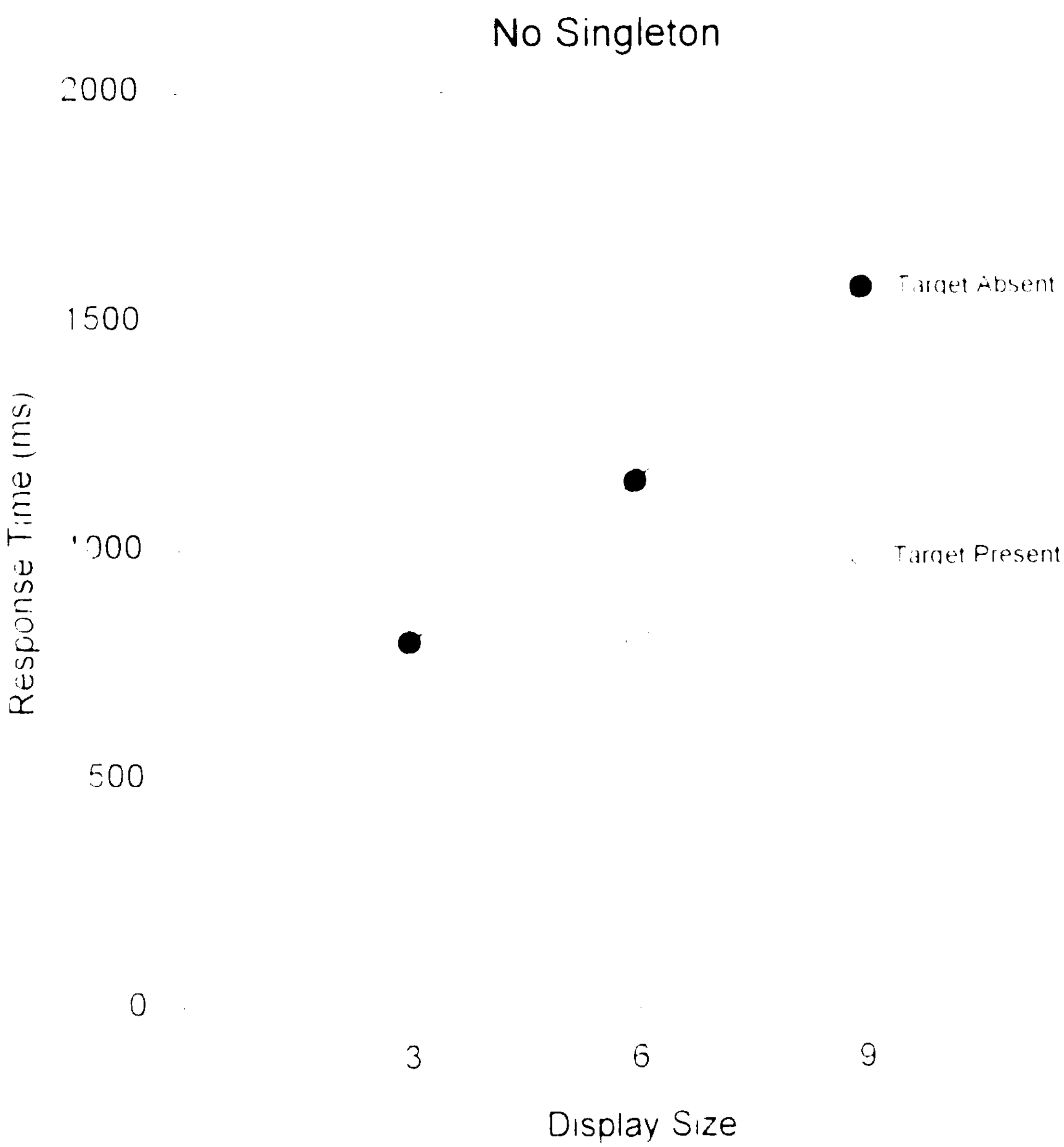
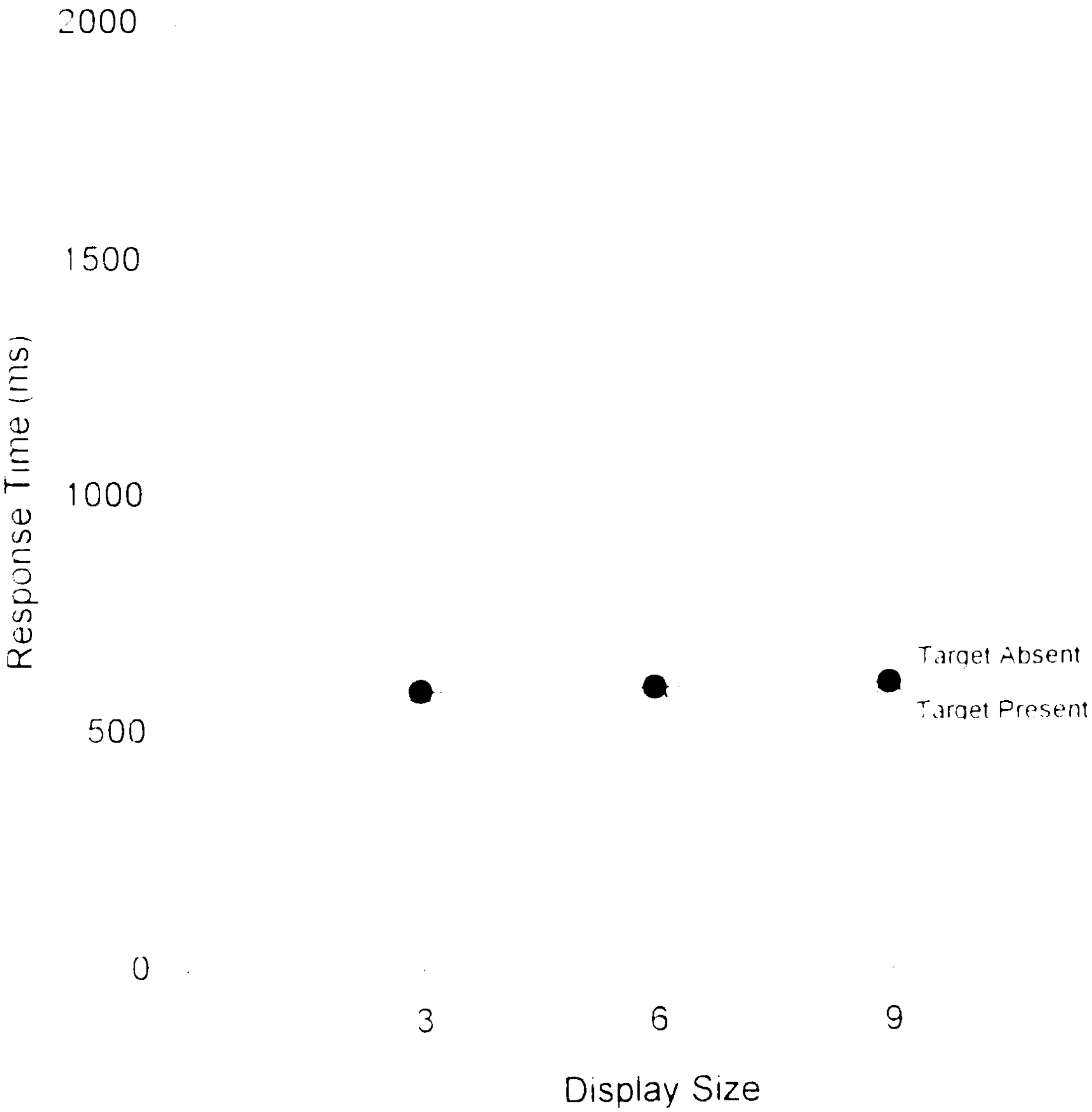
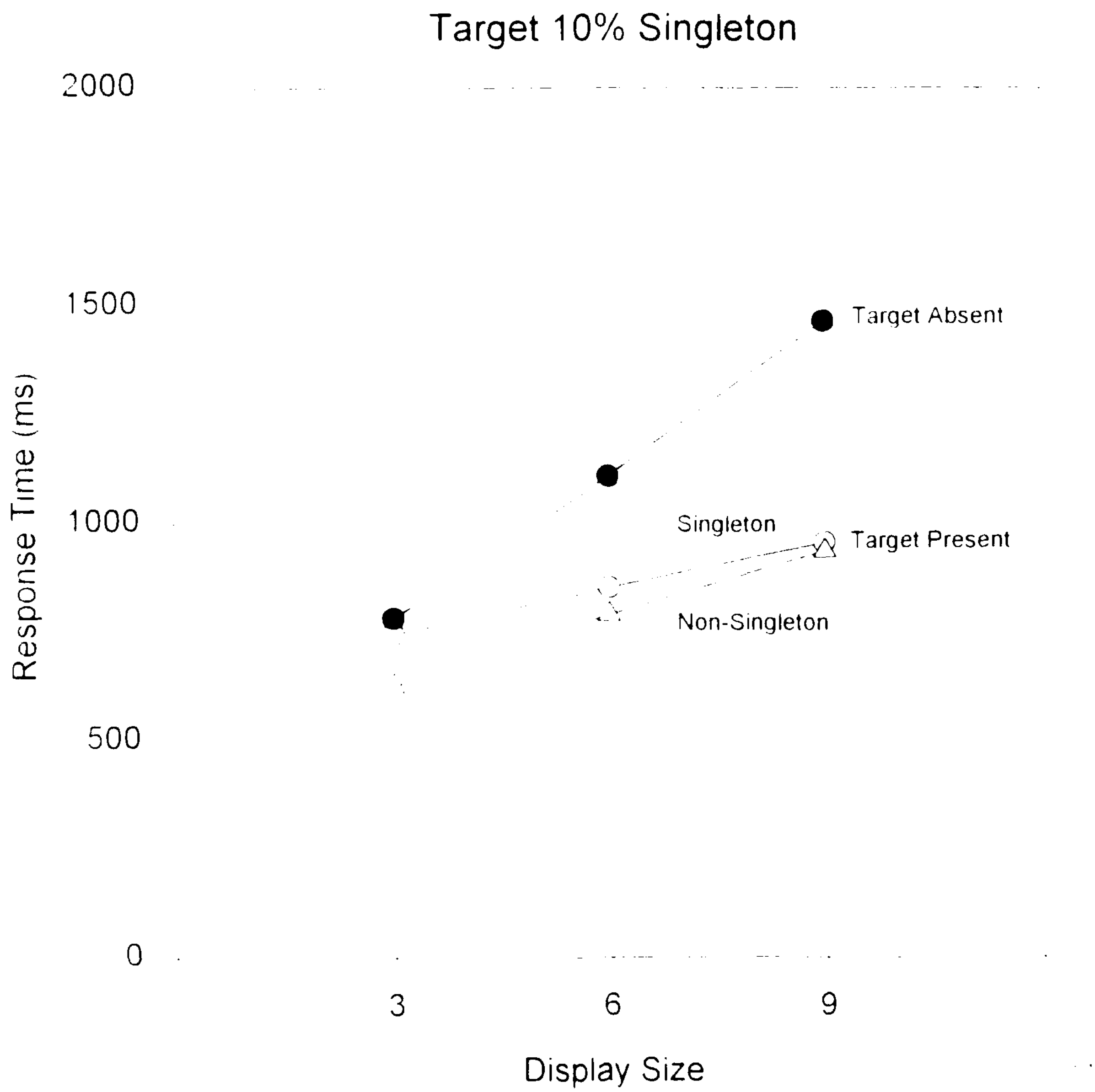

narrow, no context
wide: no context
horizontal grouping
vertical grouping
100
75
50
25
0
0
100
200
300
Percent Group Motion
Interstimulus Interval (ms)

fig 7


No Singleton
2000
1500
1000
500
0
Response Time (ms)
Target Absent
Target Present
3
6
9
Display Size

Target Always Singleton




Target 10% Singleton
2000
1500
1000
500
0
Response Time (ms)
Target Absent
Singleton
Target Present
Non-Singleton
3
6
9
Display Size