

Research Article

THE LOWER VISUAL SEARCH EFFICIENCY FOR CONJUNCTIONS IS DUE TO NOISE AND NOT SERIAL ATTENTIONAL PROCESSING

Miguel P. Eckstein

NASA Ames Research Center and Department of Medical Physics & Imaging, Cedars Sinai Medical Center, Los Angeles, California

Abstract—Models of human visual processing start with an initial stage with parallel independent processing of different physical attributes or features (e.g., color, orientation, motion). A second stage in these models is a temporally serial mechanism (visual attention) that combines or binds information across feature dimensions. Evidence for this serial mechanism is based on experimental results for visual search. I conducted a study of visual search accuracy that carefully controlled for low-level effects: physical similarity of target and distractor, element eccentricity, and eye movements. The larger set-size effects in visual search accuracy for briefly flashed conjunction displays, compared with feature displays, are quantitatively predicted by a simple model in which each feature dimension is processed independently with inherent neural noise and information is combined linearly across feature dimensions. The data are not predicted by a temporally serial mechanism or by a hybrid model with temporally serial and noisy processing. The results do not support the idea that a temporally serial mechanism, visual attention, binds information across feature dimensions and show that the conjunction-feature dichotomy is due to the noisy independent processing of features in the human visual system.

The first stage in models of human visual processing is the independent parallel processing of different physical attributes such as color, orientation, motion, and form (Neisser, 1967; Treisman & Gelade, 1980; Wolfe, Cave, & Franzel, 1989). Evidence for this parallel processing comes from physiological discoveries of many separate visual areas that appear to specialize in coding different properties (Maunsell & Newsome, 1987). A fundamental question is, how does the human visual system put together these different properties or features to form an object? The most prominent model, feature integration theory, suggests that a serial mechanism, visual attention, binds information across feature dimensions (Treisman & Gelade, 1980). A major source of evidence for this model is a fundamental finding in visual search for a target among a set of distractors: When the target differs from the distractors along one feature dimension, such as luminance, orientation, or color (feature displays), the response time to find the target is relatively constant with increasing number of distractors (parallel search). In contrast, when the target can be differentiated from the distractors only by combining information across multiple feature dimensions (conjunction displays), the search time increases approximately linearly with number of distractors. Feature integration theory attributes the larger set-size effects in conjunction displays to the serial allocation of visual attention needed to bind information across feature dimensions.

However, many investigators have found that other lower level factors affect visual search. The effect of target-distractor similarity on

visual search has been studied extensively. This work has demonstrated that increasing target-distractor similarity in a feature display increases set-size effects (Duncan & Humphreys, 1989; Palmer, 1994; Verghese & Nakayama, 1994). It is also well known that human visual contrast sensitivity decreases with increasing retinal distance from the fovea. Therefore, if the retinal eccentricity of the elements in the display increases with increasing set size, performance will decrease with increasing set size (Carrasco, Evert, Chang, & Katz, 1995; Geisler & Chou, 1995). Increasing element density has also been shown to decrease performance because of an increase in lateral inhibition and lateral masking (Carrasco et al., 1995). The number of eye movements increases with increasing target-distractor similarity (Zelinski, Sheinberg, & Bülthoff, 1993). As a result, under free viewing (eye movements allowed), it becomes hard to separate set-size effects due to increasing number of eye movements from those caused by the serial allocation of visual attention.

Many of the search-time studies have not carefully controlled for all these low-level factors that are confounded with possible set-size effects due to the serial allocation of attention. One possible method to control for low-level sources of set-size effects is to use a visual search accuracy task (Bergen & Julesz, 1983a, 1983b; Eckstein, Thomas, Palmer, & Shimozaki, 1996; Eckstein, Thomas, Shimozaki, & Whiting, 1995; Palmer, 1994). The display is flashed briefly, precluding the observer from making eye movements during the trials. Target-distractor similarity is held constant in the feature and conjunction conditions by matching the physical differences between the target and distractors (Treisman, 1991). Element eccentricity is preserved by locating the elements on fixed locations along the circumference of a circle centered at a fixation point (Palmer, 1994). Element density is preserved by manipulating set size by cuing each relevant element with a surrounding black-box cue (Palmer, 1994). The effect of the observer's internal noise has been shown to contribute to set-size effects in visual search in feature displays (Eckstein et al., 1995, 1996; Palmer, 1994; Pavel, 1990; Shaw, 1980). Noise in the internal responses of the observer arises from the variability in the firing of neural cells (Tolhurst, Movshon, & Dean, 1982). Unlike the other low-level effects, the effect of the observer's internal noise cannot be neutralized with any experimental design. However, one can generate model predictions for the expected set-size effect due to the presence of neural noise in the human visual system and compare these predictions with the experimental data. These models are based on classical signal detection theory (SDT; Green & Swets, 1966).

The goal of this study was to test with an experimental paradigm that carefully controls for low-level effects whether the larger set-size effects for conjunction versus feature displays are better predicted by a noise-based model (SDT model), a temporally serial model, or a hybrid model (noisy and serial processing).

SDT-BASED MODEL (NOISE MODEL)

The SDT-based model for visual search accuracy in feature displays (Eckstein et al., 1995, 1996; Palmer, 1994; Pavel, 1990) assumes

Address correspondence to M.P. Eckstein, Department of Medical Physics & Imaging, Cedars Sinai Medical Center, 8700 Beverly Blvd., Davis 6065, Los Angeles, CA 90048-1865; e-mail: miguel@vision.arc.nasa.gov.

Lower Visual Search Efficiency for Conjunctions

that each element in the display elicits a noisy independent response and that the observer chooses the interval that elicits the largest response. Noise will make the internal neural response of the observer to a given element (target or distractor) fluctuate from trial to trial. Therefore, the response to the distractor will exceed the response to the target every so often, and the observer will make the wrong selection. If the internal noise is assumed to be Gaussian and stochastically independent, then the observer's responses to the target and distractor can be represented as two Gaussian distributions. The observer's internal discriminability between the target and distractor can be described by the difference in mean internal responses to the target and distractors divided by the standard deviation of the response (due to noise). This measure is known as d' , the index of detectability.

Performance in a search accuracy study in which a display is flashed briefly will degrade with increasing set size owing to the increasing probability of the observer's neural response to any one distractor being larger than the neural response to the target (Eckstein et al., 1995; Eckstein & Whiting, 1996; Palmer, 1994; Pavel, 1990; Shaw, 1980). Also, the set-size effect will be larger as target-distractor discriminability, d' , decreases (Eckstein et al., 1995). Figure 1 illustrates the effect of noise and increasing set size on visual search accuracy.

This simple model can be extended to displays that include multiple feature dimensions (conjunction displays; Aiken & Palmer, 1992; Eckstein et al., 1995; Eckstein & Whiting, 1996; Pavel, 1990). In such a case, the model assumes that each feature dimension is processed by a separate group of cells that have statistically independent noise. In the conjunction display, the observer is assumed to monitor two noisy internal responses per element, corresponding to the two feature dimensions. The observer then combines the responses along the individual feature dimensions for each element using a linear operation (averaging) and selects the element that elicits the largest response along the new (averaged) decision variable. This model predicts an increased set-size effect for the conjunction condition compared with the feature condition because the conjunction condition involves the combination of information across two noisy encoders, only one of which provides information to distinguish the target from a given distractor.

For example, for a conjunction display of contrast and orientation, the target differs from one distractor along the orientation dimension and from the other distractor along the contrast dimension. Rows 1 and 2 of Table 1 specify the mean, standard deviation, and target-distractor discriminability (d') along both (contrast and orientation) feature dimensions for the target and the two types of distractors consistent with a conjunction display of this type. Responses for all elements along both dimensions are assumed to have equal standard deviations (noise).

The third row of the table shows the mean and standard deviation of the internal response for each element after averaging internal responses across feature dimensions. Finally, the table also shows the target-distractor discriminability for the target and each distractor type along the "averaged" internal response dimension. The net effect is to reduce target-distractor discriminability (defined as the difference in mean responses to target and distractor divided by the standard deviation of the responses) along the new decision variable by a factor of $\sqrt{2}$. This reduction in effective target-distractor discriminability will produce larger set-size effects (even though the physical differences between the target and distractor are matched in the feature and conjunction displays). Appendix A develops in detail the mathematical expressions to predict percentage correct as a function of set size for the SDT model in feature and conjunction displays for a two-interval forced-choice task (2IFC).

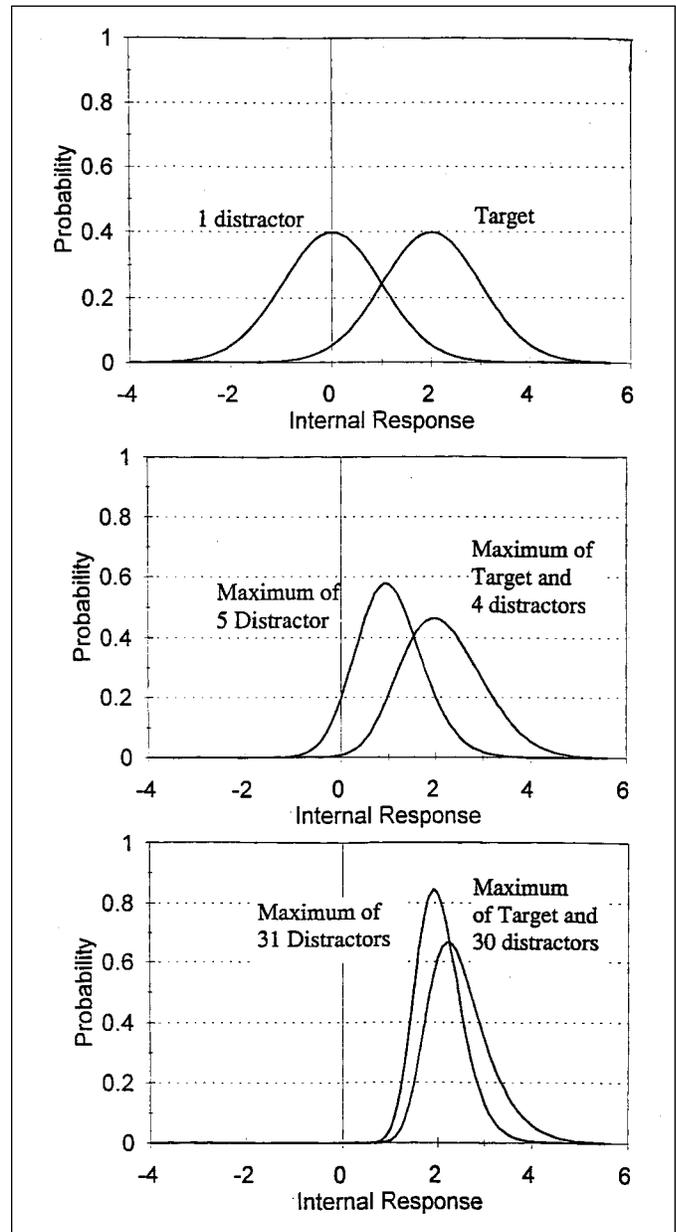


Fig. 1. Effect of noise on visual search accuracy in a two-interval forced-choice (2IFC) task. Each element in the search display is assumed to elicit a noisy Gaussian-distributed response. Percentage of correct identification of the interval containing the target is the probability that the maximum response to the target-present interval exceeds the maximum response in the target-absent interval. A qualitative measure of percentage correct is the overlap between the expected distributions of the maximum of the target-present interval and the maximum of the target-absent interval. For the case in which one interval contains the target and the other a single distractor, the distributions are simply two Gaussian distributions (top panel). When the number of distractors increases, the distributions for the maximum of n distractors (target-absent interval) and the maximum of a target and $n - 1$ distractors (target-present interval) come closer together (middle and bottom panels). Therefore, percentage correct decreases with increasing number of distractors (set size).

Table 1. Effects of averaging the internal responses across feature dimensions

| Dimension | Target | | Distractor 1 | | d' | Distractor 2 | | |
|-------------|---------------------|------------------|-----------------|------------------|----------------|-----------------|------------------|----------------|
| | M | SD | M | SD | | M | SD | d' |
| Orientation | μ_o | σ | μ_o | σ | 0 | 0 | σ | d'_o |
| Contrast | μ_c | σ | 0 | σ | d'_c | μ_c | σ | 0 |
| Averaged | $(\mu_o + \mu_c)/2$ | $\sigma/2^{1/2}$ | $(\mu_o + 0)/2$ | $\sigma/2^{1/2}$ | $d'_c/2^{1/2}$ | $(0 + \mu_c)/2$ | $\sigma/2^{1/2}$ | $d'_o/2^{1/2}$ |

Note. $d'_o = \mu_o/\sigma$ is the discriminability between the target and Distractor 2 along the orientation dimension, and $d'_c = \mu_c/\sigma$ is the discriminability between the target and Distractor 1 along the contrast dimension.

TEMPORALLY SERIAL VISUAL ATTENTION

The human data can also be compared with the set-size effect expected from a temporally serial attentional mechanism that is supposedly needed to bind information across feature dimensions (Treisman & Gelade, 1980). In this article, I test the ability of the serial search model previously used by Bergen and Julesz (1983b) to predict the set-size effect for conjunctions. In the serial search model, the observer can perfectly process k items or elements per presentation time (100 ms in the present experiment). When the display contains fewer than k elements, the observer performs perfectly (100%). When there are more than k elements in the display, the observer will process k random elements of the total n elements in the display (without returning twice to the same element). At this point, performance of the temporally serial model will degrade with increasing set size because there will be an increasing probability of not processing the target in the limited presentation time. Appendix B summarizes the mathematical expressions to predict percentage correct as a function of set size for the temporally serial model (Bergen & Julesz, 1983a) for a 2IFC task.

**HYBRID MODEL
(NOISY TEMPORALLY SERIAL PROCESSING)**

Human performance can also be compared with a more contemporary version of the temporally serial model. To do so, I introduce a hybrid model that includes temporally serial and noisy processing for the conjunction search. In this way, the model retains its serial processing for conjunctions (Treisman & Gelade, 1980) but will also be affected by target-distractor similarity (Duncan & Humphreys, 1989) because of the presence of noise. For feature search conditions, the search is parallel and noisy, and therefore the hybrid model generates predictions identical to those of the SDT model. For conjunction search conditions, the hybrid model has two different regions of behavior. When the set size (n) is less than the maximum number of elements that can be processed serially (k) during the display duration, the hybrid model performance is degraded only by noise and generates predictions identical to those of the SDT model. However, when the set size (n) is larger than the maximum number of elements that can be processed serially (k) during the display duration, the hybrid model performance is degraded by noise and by the temporally serial processing. For these conditions, the hybrid model will predict larger performance degradation than the SDT model.

Appendix C describes the hybrid model in more detail and summarizes the mathematical expressions to predict percentage correct as a function of set size for a 2IFC task.

METHODS

Two feature conditions and a conjunction condition were included in the experiment (Fig. 2). In the first (contrast) feature condition, the target was a high-contrast vertical ellipse among low-contrast vertical ellipses. In the second (orientation) feature condition, the target was a diagonally oriented low-contrast ellipse among vertical low-contrast ellipses. In the conjunction condition, the target was a high-contrast, diagonally oriented ellipse among high-contrast, vertically oriented ellipses and low-contrast, diagonally oriented ellipses. The ellipses subtended a visual angle of $19.8' \times 11.9'$.

In all conditions, 12 elements were located at equal distances (1.9°) along the circumference of a circle centered on a fixation point. The retinal eccentricity of all element locations was 4.9° . Set size was manipulated

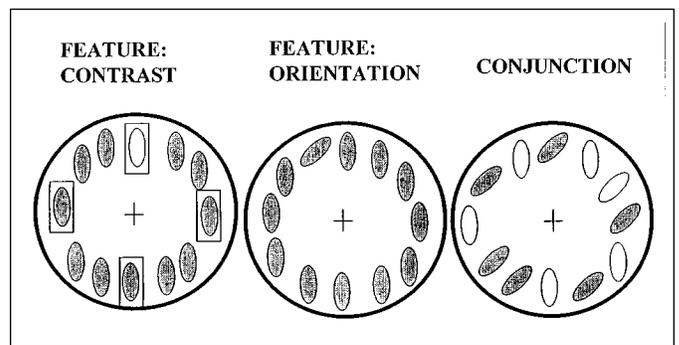


Fig. 2. Displays used. In the contrast feature condition, the target was a high-contrast vertical ellipse among low-contrast vertical ellipses (left panel). In the orientation feature condition, the target was a diagonally oriented low-contrast ellipse among vertically oriented low-contrast ellipses (middle panel). In the conjunction condition, the target was a high-contrast, diagonally oriented ellipse among low-contrast, diagonally oriented and high-contrast, vertical ellipses (right panel). Set size was manipulated by cuing each possible target location with the outline of a black box. The left-most display illustrates cuing for a set size of four.

Lower Visual Search Efficiency for Conjunctions

lated by cuing with black boxes the n relevant elements of the total 12 elements displayed. For example, when the set size was four, four black boxes surrounded the four relevant locations (Fig. 2, left panel). Two different images were presented, one after the other, to the observer. The observer had to decide which image contained the target (2IFC).

For each feature condition, the difference (in contrast or orientation) between the target and distractors was manipulated for each observer individually to achieve performance at a level of approximately 95% correct with a set size of 12. In the conjunction condition, half of the distractors were those used in the contrast feature condition, and the other half were those used in the orientation feature condition. Therefore, the physical difference in contrast between the target and half of the distractors was matched to the target-distractor difference in the contrast feature condition, and the physical difference in orientation between the target and the other half of the distractors was matched to the target-distractor difference in the orientation feature condition (see Fig. 2).

The procedure for each trial was as follows. A fixation cross was presented for 1 s. Next, a first display was displayed for 100 ms, followed by a mask for 1 s. Finally, the second display was presented for 100 ms, and another mask for 1 s. The observer had 600 ms to respond. The observer chose the interval that contained the target by pressing the "1" or "2" key on the computer keyboard. Each session consisted of 100 trials. Feedback was provided for correct responses. Set size is defined as the number of cued elements in the display.

The author and 2 naive observers participated in the study. All observers were males between ages 20 and 30 and had normal vision. The observers participated in five set-size conditions ($n = 1, 2, 4, 6,$ and 12) for each feature condition and three set-size conditions ($n = 4, 6,$ and 12) for the conjunction condition. Observers participated in four sessions per experimental condition. Images were displayed on a 15-in. Sony color monitor. The mean luminance of the background of the display was 14.7 cd/m^2 .

Percentage correct was computed for each condition. The SDT model was simultaneously fit to the feature and conjunction conditions for each observer separately with two adjustable parameters. The temporally serial model was fit to the data for each observer with one adjustable parameter, and the hybrid model was fit with three adjustable parameters.

RESULTS

Figure 3 shows percentage correct as a function of set size for 3 subjects for the contrast feature condition, the orientation feature condition, and the conjunction condition. The conjunction condition resulted in lower performance and larger set-size effects for all 3 observers.

The dashed lines in Figure 3 are the best fit (with the number of elements processed during the presentation time, k , as the single free parameter) of the temporally serial model to the conjunction data. For all 3 observers, the serial model tends to underestimate performance degradation at small set sizes ($n = 4$) and overestimate degradation at large set sizes ($n = 12$). Table 2 shows, for each observer, the value of k for the best fit and the chi-square goodness-of-fit measure for the serial model. The serial model can be rejected for all 3 observers at a 5% significance level.

The solid lines in Figure 3 are a simultaneous fit of the SDT model to the feature and conjunction data with two free parameters (target-distractor discriminability along the contrast and orientation dimensions; see Appendix A). For the 3 observers, the SDT model provides a much better fit than the serial model. Table 2 shows, for each observer, the values of the free parameters for the best fit and the chi-square goodness-of-fit measure for the SDT model.

Finally, the hybrid model was fit to the data with three free parameters: the maximum number of elements (k) processed in the display presentation time and the target-distractor discriminability along the contrast and orientation dimensions. For 2 observers, the best fit of the hybrid model resulted in a k equal to the total number of elements in the largest set-size condition ($k = 12$), effectively nullifying the effect of temporal serial processing on performance. For these 2 observers, the target-distractor discriminability values were identical to those obtained by fitting the SDT model (Table 2). In other words, for these observers, the hybrid model's best fit defaulted to the SDT model. Therefore, predictions for the hybrid model are identical to those for the SDT model and are not graphed in Figure 3 for observers K.G. and G.N.

For the 3rd observer (M.E.), the hybrid model resulted in slightly different results than the SDT model. The best value for k was 11, which means that the temporal serial processing was able to process all the elements in all set-size conditions except when the set size was 12. Therefore, for observer M.E., serial processing affected performance only when the set size was 12. The dotted line in Figure 3 (bottom panel) is the best fit of the hybrid model for the conjunction data of observer M.E. (the predictions of the hybrid model for the feature search condition were very similar to the predictions of the SDT model and are therefore not shown).

Table 2 lists the chi-square goodness-of-fit measure for the hybrid model for each observer. For observer M.E., the chi-square goodness of fit is somewhat smaller for the hybrid model than the SDT model. However, a test (with one degree of freedom) on the difference in chi-squares for the hybrid and SDT models ($\Delta\chi^2$) showed that the difference in goodness of fits was not significant.

DISCUSSION

The findings show that once the lower level factors that affect visual search (physical similarity of target and distractors, element eccentricity and density, and eye movements) are controlled for experimentally, set-size effects for conjunction versus feature displays are not consistent with a temporally serial mechanism that binds information across feature dimensions. These results disagree with those of Bergen and Julesz (1983a), who reported that results for search for an L among T s were consistent with a temporally serial mechanism, visual attention. A more contemporary hybrid model that included serial and noisy processing was also used to fit the human data in the present experiment. For 2 observers, the best fits of the hybrid model effectively nullified the temporally serial component of the model and preserved the noisy processing component, defaulting to the SDT model. For the 3rd observer, the best fit of the hybrid model resulted in a serial processing parameter that affected performance only when the set size was 12 and did not improve the fit significantly. Together, these results show that the SDT (noise) model provides the best account for the set-size effects in the human data. The results agree with experiments by Palmer (1994), who reported that results for search for an L among T s were consistent with a parallel noisy model (SDT model).

Geisler and Chou (1995) showed that the rank order of search accuracy with eccentricity in a two-alternative forced-choice task can predict the variation of search times in a high-spatial-uncertainty task. They therefore concluded that the feature-conjunction difference is explained by low-level factors. The present results extend this concept by specifically attributing the feature-conjunction difference to the existence of internal noise and a specific neural computation. The results are quantitatively accounted for by the parallel processing of feature dimensions in the human visual system, the existence of neural

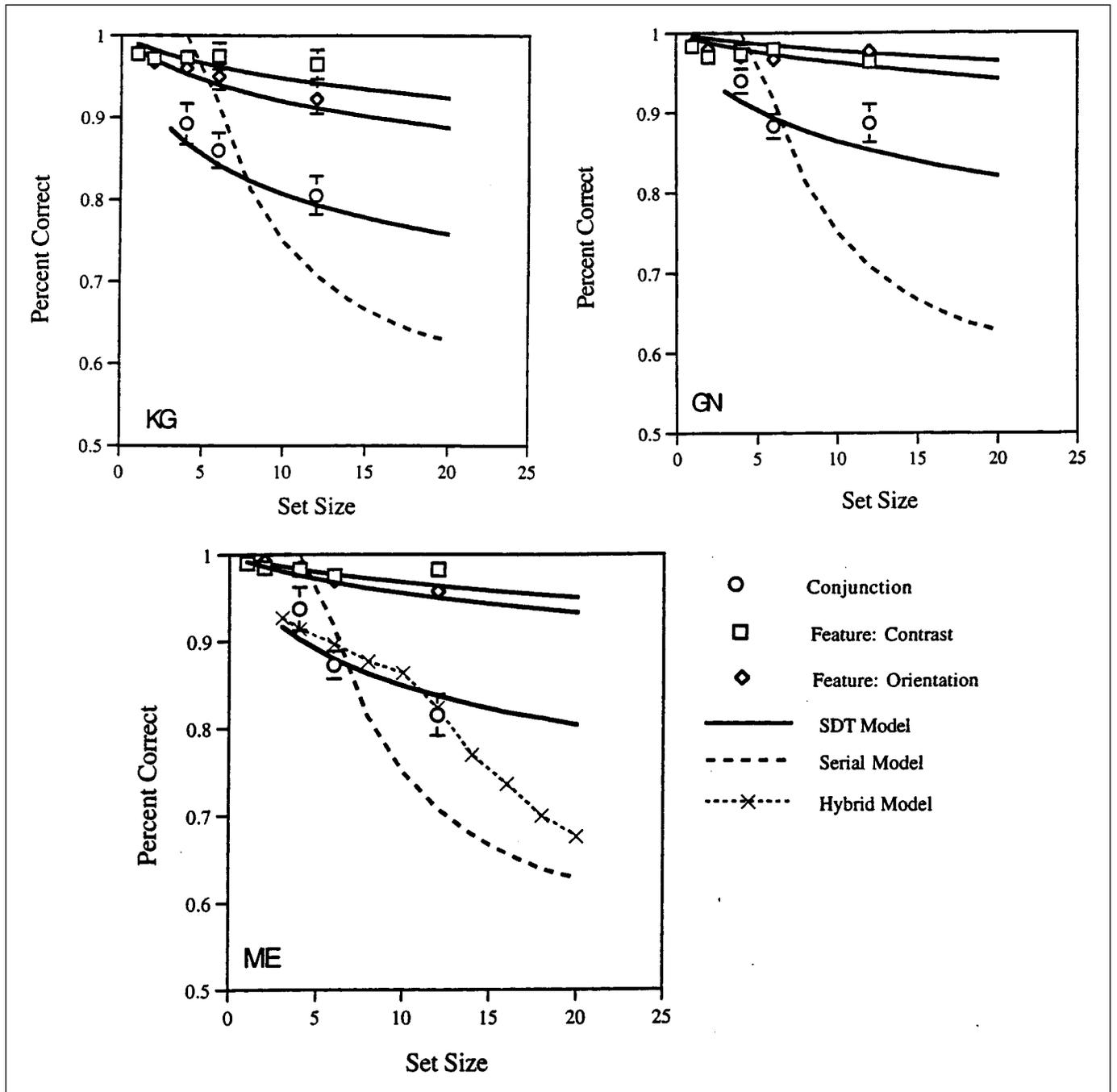


Fig. 3. Percentage correct identification of the interval containing the target as a function of set size for 3 observers in the contrast feature condition, orientation feature condition, and conjunction condition. Data points are based on average performance across four sessions of 100 trials each. Error bars are standard errors across four sessions. Solid lines correspond to the best fit of the model based on signal detection theory (SDT). Dashed lines correspond to the best fit of the temporally serial model to the conjunction data. The dotted line (observer M.E. only) corresponds to the best fit for the hybrid model consisting of serial and noisy processing. The fits of the hybrid model for the other 2 observers resulted in curves equal to those of the SDT model and are therefore not graphed.

noise, and the assumption that a linear combination is used to combine information across feature dimensions. The two former properties are well supported by physiological recordings of cells in the visual cortex (Maunsell & Newsome, 1987; Tolhurst et al., 1982). This experiment

cannot discard the possibility that other models, such as the guided search model (Wolfe, 1994; Wolfe et al., 1989), would fit the data. In the guided search model, a parallel system provides guidance to the serial allocation of attention. However, in its present form, the guided

Lower Visual Search Efficiency for Conjunctions

Table 2. Model fits

| A. Parameters for best fit | | | |
|----------------------------|---|---------------------------------|---|
| Observer | Signal detection theory model (2 free parameters) | Serial model (1 free parameter) | Hybrid model (3 free parameters) |
| K.G. | $d'_c = 3.4; d'_o = 3.0$ | $k = 4$ (25 ms/item) | $d'_c = 3.4; d'_o = 3.0;$ $k = 12 (< 8.3 \text{ ms/item})$ |
| G.N. | $d'_c = 3.5; d'_o = 3.8$ | $k = 4$ (25 ms/item) | $d'_c = 3.5; d'_o = 3.8;$ $k = 12 (< 8.3 \text{ ms/item})$ |
| M.E. | $d'_c = 3.6; d'_o = 3.5$ | $k = 4$ (25 ms/item) | $d'_c = 3.8; d'_o = 3.5;$ $k = 11 (9 \text{ ms/item})$ |

| B. Goodness of fit: χ^2/df (reduced chi-square) ^a | | | |
|---|---|-----------------------|------------------------|
| Observer | Signal detection theory model (df = 11) | Serial model (df = 2) | Hybrid model (df = 10) |
| K.G. | 0.44 | 6.6* | 0.49 |
| G.N. | 1.77 | 8.8* | 1.95* |
| M.E. | 0.85 [†] | 23.8* | 0.571 [†] |

^aThe degrees of freedom is defined as the number of data points minus the number of constraints. The goodness-of-fit measure was calculated for each model as follows (Taylor, 1982):

$$\frac{\chi_o^2}{df} = \frac{1}{(n - c)} \sum_{i=0}^n \frac{(P_{\text{model}} - P_{\text{data}})^2}{\sigma_e^2},$$

where n is the number of data points to fit, c is the number of constraints (or free parameters), P_{model} is the percentage correct predicted by the model, P_{data} is the measured percentage correct for the observer, and σ_e^2 is the expected variance.

* $P_d(\chi^2 \geq \chi_o^2) < 5\%$ (probability of obtaining a χ^2 equal to or larger than the observed χ_o^2 given that the observed data followed the model is less than .05). Models are rejected.

[†] $P_d(\Delta\chi^2 \geq \Delta\chi_o^2) > 5\%$ (probability of obtaining a difference of χ^2 , $\Delta\chi^2$, between the signal detection theory and hybrid models equal to or larger than the observed difference in χ_o^2 , $\Delta\chi_o^2$, given that the models fit the data equally well is larger than .05). The hypothesis that the models fit the data equally well is not rejected.

search model 2.0 generates predictions for response time studies and not search accuracy studies, contains many more free parameters than the SDT model (Wolfe, 1994), and does not have closed-form mathematical expressions to calculate performance.

Previous studies have shown that parallel noise models can successfully predict quantitatively set-size effects in search for a target in a feature display (Palmer, 1994), localization of a luminance increment (Shaw, 1984), and search for a contrast-defined target on a variety of backgrounds (Eckstein & Whiting, 1996; Swenson & Judy, 1981). The present results extend the success of the model in predicting set-size effects to multidimensional displays (with more than one relevant feature dimension). However, there is evidence for larger set-size effects than expected from the noise model for complex tasks such as point orientation judgments (Palmer, 1994), letter search (Shaw, 1984), and tasks involving memory (Palmer, 1990).

The fact that a temporally serial attention mechanism does not fit the data does not imply that visual attention has no role in visual search. In this experimental paradigm, the relevant elements in the task are cued with a high-contrast box. Observers are therefore required to attend to the internal responses arising from the cued locations and ignore those arising from the noncued locations. In this sense, the task requires visual attention as a selective process to monitor the relevant

items (Graham, 1989). If the observers had been unable to attend to the relevant items, performance would have been constant as a function of set size (which was manipulated by cuing possible target locations).

To conclude, given that feature integration theory is largely based on the fundamental difference in visual search for features and conjunctions, the present results seriously call into question the temporal serial allocation of visual attention as a mechanism that binds information across multiple feature dimensions.

Acknowledgments—The author would like to thank Brent Beutter, Preeti Verghese, Albert Ahumada, Jr., Andrew B. Watson, and James P. Thomas for comments and helpful suggestions at different stages during the research. This research was supported in part by NASA RTOP 199-06-12-39.

REFERENCES

Aiken, D., & Palmer, J. (1992). *Pursuing an alternative theory of the difference between conjunctions and disjunctions in visual search*. Unpublished master's thesis, University of Washington, Seattle.
 Bergen, J.R., & Julesz, B. (1983a). Parallel versus serial processing in rapid pattern discrimination. *Nature, London*, *303*, 696–698.
 Bergen, J.R., & Julesz, B. (1983b). Rapid discrimination of visual patterns. *IEEE Transactions on Systems, Man and Cybernetics*, *13*, 857–863.

Carrasco, M., Evert, D.L., Chang, I., & Katz, S.M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics*, *57*, 1241–1261.

Duncan, J., & Humphreys, G.W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.

Eckstein, M.P., Thomas, J.P., Palmer, J., & Shimozaki, S.S. (1996). Further predictions of signal detection theory on visual search accuracy: Conjunction, disjunction and triple conjunctions [Abstract]. *Investigative Ophthalmology & Visual Science*, *36*(Suppl.), 15.

Eckstein, M.P., Thomas, J.P., Shimozaki, S.S., & Whiting, J.S. (1995). Signal detection theory explains many aspects of visual search efficiency [Abstract]. *Investigative Ophthalmology & Visual Science*, *37*(Suppl.), 2994.

Eckstein, M.P., & Whiting, J.S. (1996). Visual signal detection in structured backgrounds: I. Effect of number of possible spatial locations and signal contrast. *Journal of the Optical Society of America A*, *9*, 1777–1787.

Geisler, W., & Chou, K.L. (1995). Separation of low-level and high-level factors in complex tasks: Visual search. *Psychological Review*, *102*, 356–378.

Graham, N.V.S. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

Green, D.M., & Swets, J.A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Maunsell, J.H.R., & Newsome, W.T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, *10*, 363–401.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Palmer, J. (1990). Attentional limits on the perception and memory of visual information. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 332–350.

Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research*, *34*, 1703–1721.

Pavel, M. (1990, November). *A statistical model of preattentive visual search*. Paper presented at the annual meeting of the Psychonomic Society, New Orleans.

Shaw, M.L. (1980). Identifying attentional and decision making components in information processing. In R.S. Nickerson (Ed.), *Attention and performance VIII* (pp. 277–296). Hillsdale, NJ: Erlbaum.

Shaw, M.L. (1984). Division of attention among spatial locations: A fundamental difference between detection of letters and detection of luminance increments. In H. Bouma & D.G. Bouwhuis (Eds.), *Attention and performance X* (pp. 106–121). Hillsdale, NJ: Erlbaum.

Swenson, R.G., & Judy, P.F. (1981). Detection of noisy visual targets: Models for the effects of spatial uncertainty and signal-to-noise ratio. *Perception & Psychophysics*, *29*, 521–534.

Taylor, J.R. (1982). *An introduction to error analysis*. Mill Valley, CA: University Science Books, Oxford University Press.

Tolhurst, D.J., Movshon, J.A., & Dean, A.F. (1982). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research*, *23*, 775–785.

Treisman, A. (1991). Search, similarity and integration of features between and within dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 652–675.

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

Verghese, P., & Nakayama, K. (1994). Stimulus discriminability in visual search. *Vision Research*, *34*, 2453–2467.

Wolfe, J.M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202–238.

Wolfe, J.M., Cave, K.R., & Franzel, S.L. (1989). Guided search: An alternative to the modified feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 419–433.

Zelinsky, G., Sheinberg, D., & Bülhoff, H. (1993). Eye movements during visual search [Abstract]. *Investigative Ophthalmology and Visual Science*, *34*(Suppl.), 1236.

APPENDIX A: SDT-BASED MODEL

Feature Search (2IFC)

In the feature conditions, the observer monitors one noisy response per relevant element in each interval and then chooses the interval containing the element that elicited the highest internal response. If we assume that the noise is Gaussian distributed and statistically independent for each element, the probability of correctly identifying the target-present interval is given by

$$P_c(n, d') = \int_{-\infty}^{+\infty} [g(x)G^{2n-1}(x+d') + (n-1)g(x+d')G^{2n-2}(x+d')G(x)]dx, \tag{A.1}$$

where

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{A.2}$$

is the probability that the internal response to the target takes a value x ,

$$G(x) = \int_{-\infty}^x g(y)dy \tag{A.3}$$

is the cumulative probability that the internal response to the target takes a value less than x , $g(x+d')$ is the probability that the response to the distractor takes a value x (defined similarly to Equation A.2), $G(x+d')$ is the probability that the response to the distractor takes a value less than x (defined similarly to Equation A.3), n is the total number of elements, and d' is a measure of the observer's internal discriminability between the target and distractor, defined to be the distance between the target and distractor distributions in standard deviation units.

Equations A.1 through A.3 can be used to predict performance in the orientation and contrast feature conditions by replacing d' by d'_c (d' along the contrast dimension) and d'_o (along the orientation dimension). A mathematical expression to calculate percentage correct in the feature condition was first derived by Shaw (1980) and is mathematically equivalent to Equation A.1.

Conjunction Search (2IFC)

In the conjunction condition, the observer monitors two noisy feature responses per element, averages the two responses, and then chooses the interval that contains the element that elicits the highest response. The effective discriminability between the target and the two types of distractors (d'_{d1} and d'_{d2}) after the averaging operation can be shown to be (see Table 1 for more detail)

$$d'_{d1} = \frac{1}{\sqrt{2}} [d'_c + 0] = \frac{1}{\sqrt{2}} d'_c \tag{A.4.a}$$

and

$$d'_{d2} = \frac{1}{\sqrt{2}} [d'_o + 0] = \frac{1}{\sqrt{2}} d'_o, \tag{A.4.b}$$

where d'_{d1} and d'_{d2} are the effective target-distractor discriminabilities along the new decision variable after combining information across feature dimensions. Percentage correct for finding the target among $n/2$ distractors of Type 1 and $n/2$ distractors of Type 2 for the conjunction condition is given by (Eckstein et al., 1995)²

$$P_{SDT} = \int_{-\infty}^{+\infty} g(x)G^n(x+d'_c)G^{n-1}(x+d'_o) + int(n/2)g(x+d'_c)G^{n-1}(x+d'_c)G^{n-1}(x+d'_o)G(x) + int\left(\frac{n-1}{2}\right)g(x+d'_o)G^{n-2}(x+d'_o)G^n(x+d'_c)G(x)dx, \tag{A.5}$$

1. Note that the notation convention is different from standard SDT notation, which uses $g(x)$ for the probability density function of the distractor and $g(x-d')$ for the probability density function of the target. Instead, I use $g(x)$ for the target and $g(x+d')$ for the distractor to make the notation consistent with the conjunction equations, which have more than one type of distractor and cannot be described with the more standard SDT notation.

2. When an interval presented an odd number of distractors, the last distractor was randomly picked from one of the two types. In this way, half of these trials presented $int(n/2) + 1$ distractors of Type 1 and $int(n/2)$ distractors of Type 2 and the other half of these trials presented $int(n/2)$ distractors of Type 1 and $int(n/2) + 1$ distractors of Type 2. The mathematical expression (Equation A.5) given corresponds to the former element configuration. The correct calculation of percentage correct also requires calculation of performance for the other element configuration. However, when d' is similar along the two dimensions, Equation A.5 gives results virtually identical to those of the full calculation.

Lower Visual Search Efficiency for Conjunctions

where d'_c and d'_o are the target-distractor internal discriminabilities along the individual feature dimensions, the int function rounds off to the lowest integer, n is the set size, and the functions $g(x)$ and $G(x)$ are as previously defined (Equations A.2 and A.3). Equations A.1 through A.3 (for feature conditions) and Equations A.4 and A.5 (for conjunction conditions) were simultaneously fit to the data for three conditions with two free parameters, d'_c (target-distractor discriminability along the contrast dimension) and d'_o (target-distractor discriminability along the orientation dimension).

APPENDIX B: TEMPORALLY SERIAL ATTENTION MODEL

The temporally serial model assumes that the observer can serially process k elements in the display presentation time (100 ms in the present experiment). When there are less than k items in the display, the observer performs perfectly (100% correct). When there are $n > k$ elements in the display, the observer is assumed to process any k random elements from the total n elements. If none of the k randomly sampled elements is the target, then the observer guesses and selects one of the two intervals at random. Performance selecting the interval containing the target for such an observer is given by

$$P_{\text{serial}} = \frac{k}{n} + guess = \frac{k}{n} + \left(1 - \frac{k}{n}\right)0.5 = \frac{k}{2n} + 0.5, \quad (\text{B.1})$$

where k is the number of items processed in the display duration and n is set size in the display. Equation B.1 was fit to the data with k as the only free parameter.

APPENDIX C: HYBRID MODEL (SERIAL AND NOISY PROCESSING)

Feature Search (2IFC)

For feature search conditions in which visual search is parallel (Treisman & Gelade, 1980), the hybrid model consists solely of noisy processing. Therefore,

accuracy predictions are identical to those of the SDT model and are given by Equations A.1 through A.3.

Conjunction Search (2IFC)

In conjunction search, the hybrid model is limited by both temporally serial processing of the elements in the display and noisy encoding. When the maximum number of elements that can be processed (k) is larger than the total number of elements in the display (n), the observer processes all elements in the display, compares the noisy responses, and chooses the interval containing the largest internal response. In other words, the model behaves as the SDT model because the serial processing does not affect performance. However, when the maximum number of elements that can be processed (k) is less than the number of elements in the display, the observer randomly samples k of the n elements per interval. The observer then compares the k processed noisy responses from one interval with the k processed noisy responses from the second interval and chooses the interval that contains the largest response.

It can be shown that percentage correct for such an observer in the conjunction search is given by

$$P_{\text{hybrid}} = P_{\text{SDT}}(n, d'_c, d'_o) \quad \text{for } k \geq n \quad (\text{C.1.a})$$

and

$$P_{\text{hybrid}} = (1 - k/n)0.5 + (k/n)P_{\text{SDT}}(k, d'_c, d'_o) \quad \text{for } k < n, \quad (\text{C.1.b})$$

where k is the maximum number of elements that can be processed, n is the set size in the display, and $P_{\text{SDT}}(k, d'_c, d'_o)$ is the probability that the observer chooses the target-present interval given that one of the processed k items per interval is the target (Equation A.5). The $(1 - k/n)0.5$ term corresponds to the guessing rate for those trials on which the observer did not process the target.

The hybrid model was simultaneously fit to the feature and conjunction data with three free parameters: k , the maximum number of elements that can be processed; d'_c , the internal discriminability along the orientation dimension; and d'_o , the discriminability along the contrast dimension.

(RECEIVED 4/25/97; ACCEPTED 8/18/97)