Title
Attention capacity and task difficulty in visual search

Journal
Cognition, 94(3)

ISSN
0010-0277

Authors
Huang, L Q
Pashler, Harold

Publication Date
2005

Peer reviewed

Available online at www.sciencedirect.com
SCIENCE DIRECT®

COGNITION

Cognition 94 (2005) B101–B111

www.elsevier.com/locate/COGNIT

Brief article

Attention capacity and task difficulty in visual search

Liqiang Huang, Harold Pashler*

Department of Psychology 0109, University of California, San Diego, La Jolla, CA 92093, USA

Received 11 June 2004; accepted 27 June 2004

Abstract

When a visual search task is very difficult (as when a small feature difference defines the target), even detection of a unique element may be substantially slowed by increases in display set size. This has been attributed to the influence of attentional capacity limits. We examined the influence of attentional capacity limits on three kinds of search task: difficult feature search (with a subtle featural difference), difficult conjunction search, and spatial-configuration search. In all 3 tasks, each trial contained sixteen items, divided into two eight-item sets. The two sets were presented either successively or simultaneously. Comparison of accuracy in successive versus simultaneous presentations revealed that attentional capacity limitations are present only in the case of spatial-configuration search. While the other two types of task were inefficient (as reflected in steep search slopes), no capacity limitations were evident. We conclude that the difficulty of a visual search task affects search efficiency but does not necessarily introduce attentional capacity limits.

Keywords: Attention capacity; Visual search; Task difficulty

1. Introduction

The visual search task has become one of the most widely used measures in the study of visual perception and attention, beginning with the work of Treisman and Gelade (1980). Most of this work has aimed to characterize visual attention limitations by scrutinizing "search slopes".[1] The present article argues (in agreement with writers such as Duncan, 1980; Geisler, 1989; Palmer, 1994) that search slopes may offer a misleading

* Corresponding author. Tel.: +1 858 534 394; fax: +1 419 821 6421.
E-mail address: hpashler@ucsd.edu (H. Pashler).

[1] Here the terms display-set-size effect (or slope) will be used to refer to an increase in the response time for speeded detection tasks as the number of elements in the search display is increased.

0022-2860/$ - see front matter
doi:10.1016/j.cognition.2004.06.006

answer to this very fundamental question. We go on to employ a different and, we will argue, more incisive test of capacity limits, obtaining results suggesting that some, but not all, of the perceptual operations thought to exhibit capacity limitations are actually quite free of such limitations.

In speeded visual search tasks in which an observer has an opportunity to view a display as long as s/he would like, arranged so that the target differs from uniform distractors in only one feature dimension, search time usually does not increase with the number of distractors. In daily life, we experience essentially the same phenomenon from time to time, as when we find a person wearing red in a large crowd of people all wearing green. This "pop-out" effect is generally agreed to reflect spatially parallel processing (Treisman & Gelade, 1980). When the target/distractor difference is very subtle, however, even this sort of *singleton search* often shows a substantial display-set-size effect (Carter, 1982; Carter & Carter, 1981; Nagy & Sanchez, 1990). Thus, a person wearing red cannot be so readily picked out from a crowd of people all wearing a slightly pinkish red chartreuse. Thus, "difficulty" (in the sense of target–distractor similarity) significantly affects the *efficiency* (slope) of a visual search. A number of researchers have concluded that target–distractor similarity is critical in determining the attentional limitations evident in visual search (Duncan & Humphreys, 1989).

What is not yet clear, however, is what causes the inefficiency in singleton searches involving a subtle feature difference. One possibility is that the inefficiency reflects an attentional capacity limit (Nagy & Sanchez, 1990), a concept that will be discussed in more depth below. Another possibility is that as the processing of each item is impaired, a display-set-size effect arises for reasons other than attentional capacity limitation.

1.1. Attentional capacity limits and the serial/parallel dichotomy

The highly influential Feature Integration Theory of Treisman and Gelade (1980) argues that whereas feature singleton search is accomplished with parallel processing, search requiring detection of targets defined in a more complex fashion (e.g. by feature conjunction) requires serial processing. This theory stimulated a great deal of research. However, most researchers have abandoned it (e.g. Wolfe, 1998a), as evidence accumulated to suggest that conjunction search can sometimes be quite efficient (Nakayama & Silverman, 1986) and that (as mentioned above), feature search can be inefficient when targets and distractors are very similar (Carter, 1982; Carter & Carter, 1981; Nagy & Sanchez, 1990). In addition, as is well known, sizable display set size slopes might be predicted by either a serial search or a limited-capacity parallel search in which processing is always parallel, but worsens when display set size increases (Townsend, 1976, 1990; Wolfe, 1998a).

In this study we use the term "attentional capacity limit" to describe an aspect of processing that cannot be conclusively determined by search efficiency alone: is the quality or speed of processing (of each item) reduced when the number of items is increased? If processing is retarded or worsened, this is said to show an attentional capacity limit. If the quality and speed of processing remain constant, no attention capacity limit is present. This can be put another way: Suppose an array of stimuli is divided into two parts; if the processing of each part is independent of that of the other (the processing is the same for one part whether or not the other part is being processed simultaneously), then there is no attentional capacity limit; on the other hand, if the processing of each part proceeds at a cost to the processing of the other portion, then we say there is an attentional capacity limit. There is a large literature on the effects of display set size on search response times; but for reasons we mention below, this literature is not conclusive on the question of attentional capacity limitation. It is apparent that both singleton search (search for an element that differs from the background) and conjunction search can produce a broad range of efficiency or inefficiency. What is not clear,[2] for reasons discussed below, is the implication for attentional capacity limitations.

1.2. Display-set-size effect vs. attentional capacity limit

If the slope is negligible in a speeded visual search task (pop-out), then we can reasonably infer that there is no attentional capacity limit. On the other hand, the presence of a substantial slope is less conclusive. It may reflect an attentional capacity limit, or it may be caused by at least two other factors—statistical decisional noise or eye movements.

In a visual search experiment, if each item in the display is analyzed with some fixed degree of accuracy, as the total number of items in the display is increased, the overall probability of an error will rise due to a factor often termed *statistical decisional noise*. To see why this is so, suppose you are searching for a target in a display containing 1 object, and there is a probability of 0.1 that you will mistake a distractor for the target. If the display contains 100 objects, the probability of mistaking at least one distractor for the target would be $1-(1-p)^{100}=0.999973$! That is to say, even if the processing for each item is unaffected by display set size, your accuracy with such a very large display will be abysmal. This point has been developed in connection with threshold detection (Geisler, 1989; Geisler & Chou, 1995; Green & Swets, 1967; Palmer, 1994; Shaw, 1984). What are the implications for speeded search tasks with high accuracy rates? If the display remains until the subject responds, information will be collected for a longer period of time under conditions of greater uncertainty. Thus, both the RT and the error rate can increase with the number of items even if there is no capacity limitation (see Duncan, 1980; Pashler, 1998, for discussion).

Another potentially contaminating factor is subjects' tendency to make eye movements when scanning large displays. If the task becomes so difficult that processing the stimuli requires fixation, then even if the underlying mechanism is not limited by attention capacity, eye movements may be enough to cause RT increases with increased display size (Maioli, Benaglio, Siri, Sosta, & Cappa, 2001).

1.3. Successive/simultaneous comparison method

The method used here for testing capacity limitations was first introduced by Eriksen and Spencer (1969) and Shiffrin and Gardner (1972). It involves assessing the accuracy of judgments about a briefly presented display when its different parts are presented successively (SUCC), as compared to when the entire display is presented simultaneously

[2] This question of attentional capacity limit is related to, but not identical to, the serial/parallel distinction.

(SIM). In both types of displays, every element is presented for the same period of time and followed by a mask, so that the time available for processing each item is equated across the SUCC and SIM conditions. If processing is subject to a capacity limitation, there should be a substantial advantage in SUCC display, because it allows the limited resources to be allocated to just a subset of the display at any given time. If processing has unlimited capacity, on the other hand, there should be no difference in accuracy between SUCC and SIM displays.

The comparison between performance in SUCC and SIM displays allows one to investigate capacity allocation while holding display content—and thus statistical decision noise—constant (Duncan, 1980; Gardner & Joseph, 1975). In the experiments reported here, we used very brief stimulus presentation (about 200 ms or less) that can minimize the role of eye movements.

1.4. Outline of this study

We presented subjects with 16-item displays, sometimes in their entirety (SIM), and sometimes divided equally into two halves which were presented one by one (SUCC; see Fig. 1). Each of tasks 1–3 included two block types: the speeded blocks, measuring

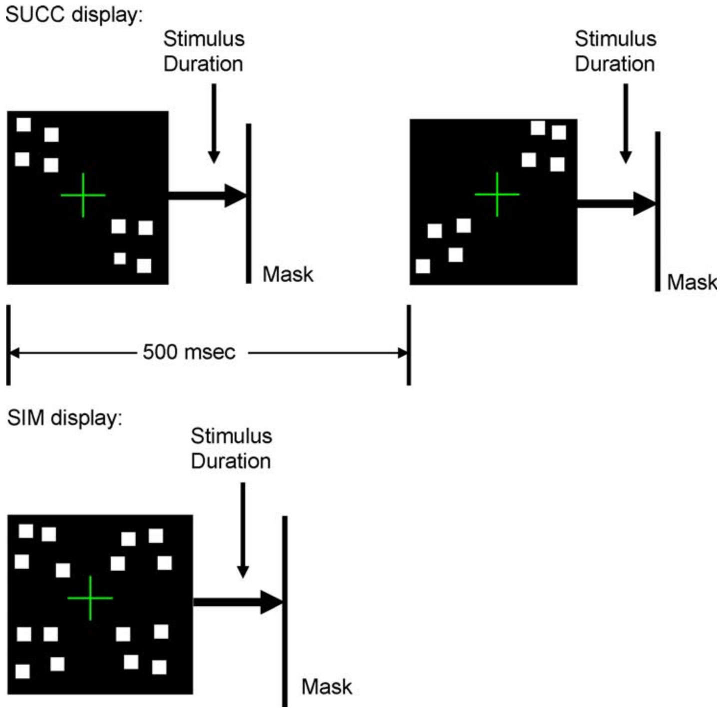

Fig. 1. Procedure and display used in this study. Top figure: in a SUCC display, 8 items appear in each of 2 corners; after 500 ms, eight more items appear in the other two corners. Bottom figure: in an SIM display, all 16 items appear at the same time in the four corners.

involved traditional speeded response time, allowed measurement of the display-set-size effect; the other, the SUCC/SIM blocks, was designed for SUCC/SIM comparison.

The stimuli used in the three tasks are shown in Fig. 2. In the first task, subjects searched for a size singleton; the target/distractor difference was very subtle, insuring that

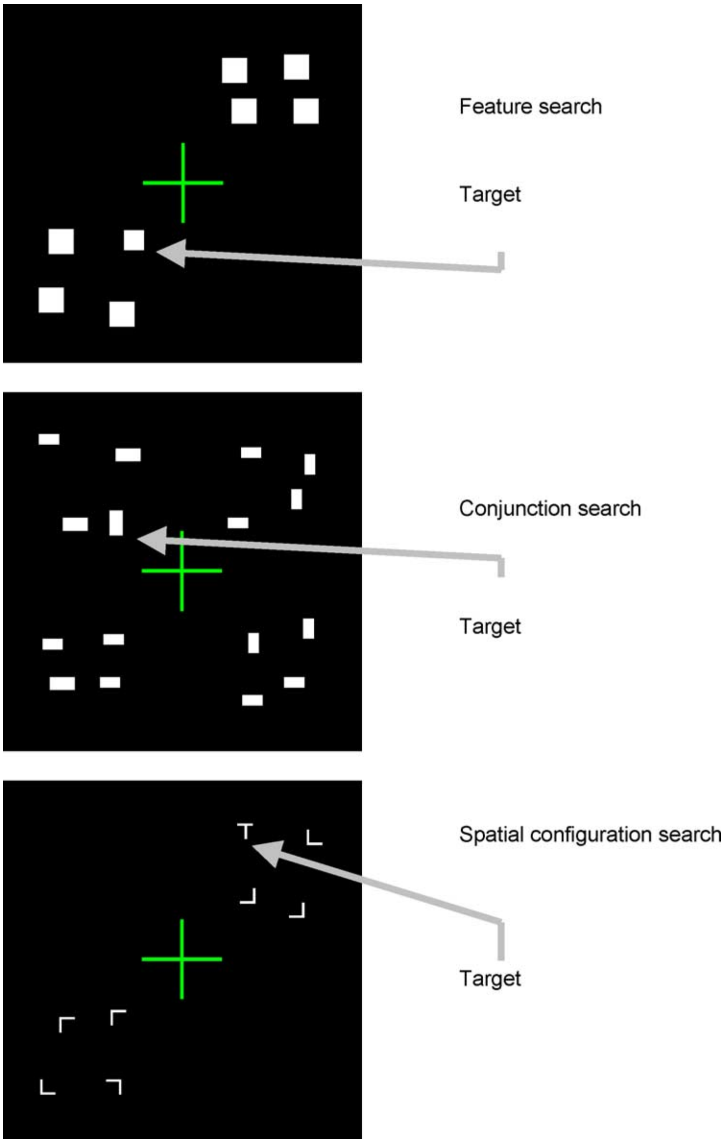

Fig. 2. Approximate appearance of stimuli in this study. Top figure: feature search. Target, 0.48 cm × 0.48 cm square; distractors, 0.59 cm × 0.59 cm squares. Middle figure: conjunction search. Target, 0.59 cm × 0.31 cm rectangle; distractors, 0.31 cm × 0.59 cm, 0.48 cm × 0.25 cm or 0.25 cm × 0.48 cm rectangles. Bottom figure: spatial configuration search. Target, "T" in all four orientations (consisted of two segments of 0.36 cm × 0.06 cm). Distractors, "L"s in all four orientations (consisted of two segments of 0.36 cm × 0.06 cm).

the search would be inefficient (high slope in the RT blocks). In the second task, subjects searched for a large vertical object among a set of objects that were large horizontal, small vertical, or small horizontal—a typical conjunction search task. Again, the target/distractor difference was subtle, insuring an inefficient search slope. In the third task, subjects searched for a letter **T** among a set of letter **L**s, with all letters randomly oriented; this is a spatial-configuration search (search for a target defined with spatial configuration) and it is known to be very inefficient (Wolfe, 1998b).

If the increasing difficulty of a search task (or its increasing inefficiency) indicates the presence of an attentional capacity limit, then we should expect the tasks in all three tasks to show attentional capacity limits since they are all predicted to be difficult and inefficient. On the other hand, it is possible that, if obstacles to efficiency depend chiefly on the contaminating factors mentioned earlier (statistical decisional noise and eye movement), processing efficiency and attention capacity may be dissociable. Suppose that the degree of a task's demand on the attention is determined only by the nature of the required search and not by the efficiency of the task. In that case, the perceptual operations required in tasks 1 and 2 should not be significantly limited by attention capacity, since we already know that both singleton and conjunction searches sometimes show no attentional capacity limit (near-zero display-set-size effect). But spatial configuration search might still significantly involve an attentional capacity limit; spatial-configuration search might by its very nature be limited by attentional capacity.

In summary, the present study seeks to shed new light on attentional capacity limitations in visual search by systematically comparing two modes of presentation: the simultaneous presentation of an entire display for a given duration versus the successive presentation of subsets of the display (each for the same duration). For the same tasks and subjects, standard measures of RT slopes are assessed in the traditional manner of Treisman and Gelade (1980).

2. Method

2.1. Subjects

Eight subjects were paid to participate in this study. All had normal vision, and none were aware of the purpose of this study.

2.2. Apparatus

Stimuli were presented on a high-resolution color monitor. Responses were recorded from two adjacent keys using a standard keyboard. The subjects viewed the displays from a distance of about 60 cm.

2.3. Design and Stimuli

There were two types of blocks: SUCC/SIM blocks and RT blocks. In a SUCC/SIM block, two types of display were randomly mixed and equally likely to occur: a SIM display containing 16 items, 4 in each of the 4 corners (Fig. 1), or a SUCC display consisting of 2 frames, each with 8 items shown in 2 corners (top-left and bottom-right, or bottom-left and top-right). The space between the edges of inside squares (the gap between the 4 corner groups) is 2.21 cm and the gap within a corner group is 0.81 cm before the jittering. In a speeded search block, each display contained 16 or 8 items, as in a SIM or SUCC display, respectively. All items were shown against a black background. Each item was randomly jittered within a range of ± 0.34 cm.

A pre-defined target was present in half the displays. The targets and distractors for all three tasks are illustrated in Fig. 2.

2.4. Procedure

Each trial began with a small green fixation cross presented in the center of the screen. Subjects were instructed to fixate the cross, which remained present for 400 ms. The cross was followed by a short blank interval (400 ms), which was then followed by the first display.

In a speeded search block, each display remained until the subject responded. In an SIM display, the entire display was presented for certain stimulus duration and then replaced by the mask. In an SUCC display, the two sub-displays were presented successively. Each sub-display was individually and locally masked after a stimulus duration equal to that of a SIM display. The interval between the first and second sub-displays was 500 ms, and the duration of the mask was 100 ms. The stimulus duration was adjusted for each session to get appropriate accuracy levels (mean stimulus durations were as follows: for feature search, 164 ms; for cojunction search, 146 ms; for spatial configuration search, 189 ms). These adjustments were conducted using a staircase: two successive correct responses led to a 1/11 decrease of stimulus duration; one erroneous response led to a 1/10 increase, with the restriction that the stimulus duration was never greater than 250 ms. This was to insure that the stimulus duration would not be long enough to encourage an eye movement.

In all the tasks, when the display was presented, subjects attempted to determine the presence or absence of the target. They responded by pressing one of two adjacent keys with fingers of right hand. They were told to respond as rapidly and accurately as possible in speeded search blocks and as accurately as possible in SUCC/SIM blocks. A positive or negative sound provided feedback on the accuracy of each response. In each session, a subject performed 20 blocks of 50 trials, with block order alternating. The first two blocks were excluded as practice, leaving six RT blocks and 12 SUCC/SIM blocks. Each subject completed three sessions of each task. Different tasks alternated and the tasks used in the initial session were counterbalanced across subjects.

3. Results

The results from RT blocks are shown in Fig. 3. The three tasks yielded similar results by this measure: all three task types showed very inefficient search slopes.

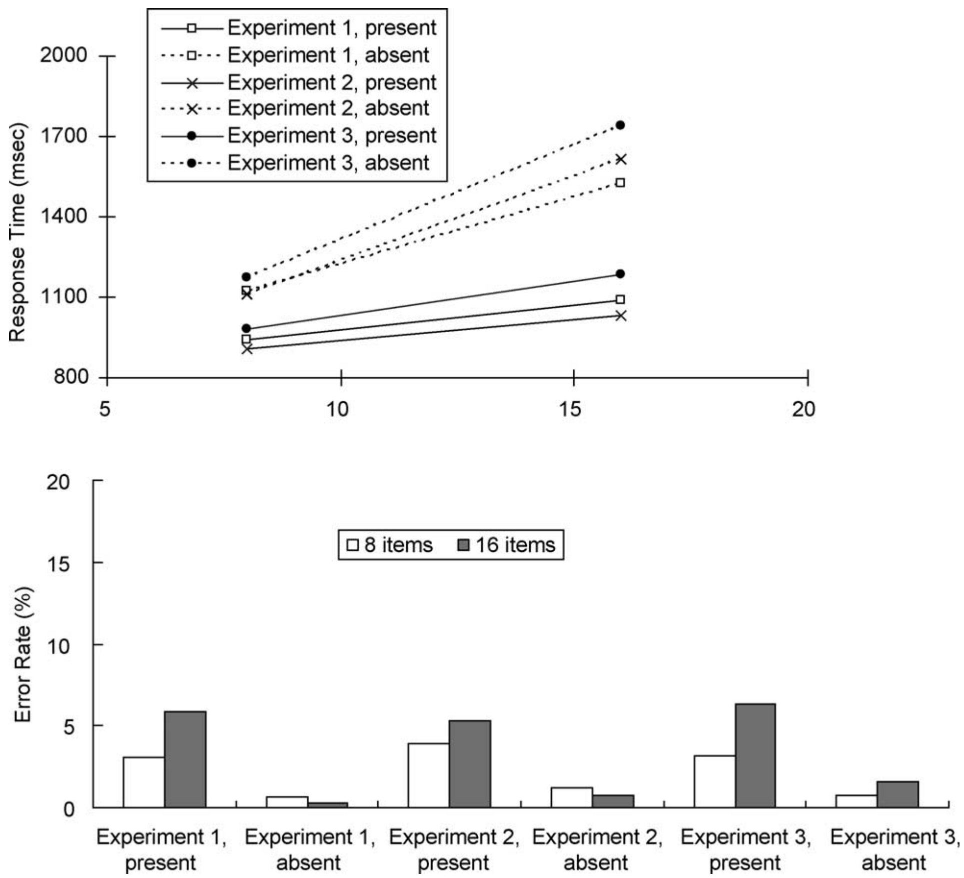

Fig. 3. Results of speeded search blocks. Top figure: response times for 8 or 16 items. Bottom figure: accuracy. All three tasks show typical very inefficient search slopes. The slopes of all three tasks (feature search, 35 ms/item; cojunction search, 39 ms/item; spatial configuration search, 48 ms/item) suggested that all three tasks are inefficient or serial.

The results of SUCC/SIM blocks (Fig. 4), however, reveal an important difference in performance between the various tasks. In tasks 1 and 2, subjects gained little accuracy advantage from the separation of displays into two sub-displays that could be processed at different times (SUCC advantage). In spatial configuration search, by contrast, this separation gave subjects a substantial advantage.

An ANOVA was performed to test the effects reported above. It confirmed that indeed there was little SUCC advantage in feature search ($F(1,7)=0.0077$) or cojunction search ($F(1,7)=1.37$); however, there was a significant SUCC advantage in spatial configuration search ($F(1,7)=20.48$; $p<0.003$). An across-task ANOVA analysis confirmed that the SUCC advantage is significantly larger in spatial configuration search than in both feature search ($F(1,7)=6.58$; $p<0.05$) and cojunction search ($F(1,7)=5.92$; $p<0.05$).

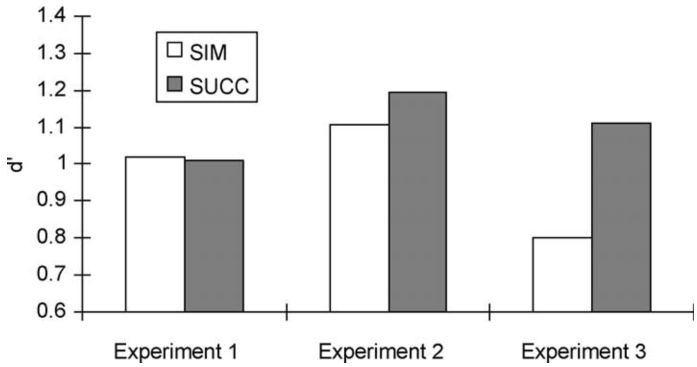

Fig. 4. Results of SUCC/SIM blocks. Figure shows d'. There is no SUCC advantage in tasks 1 and 2, but a significant SUCC advantage in spatial configuration search.

4. Discussion

Our results suggest that although the difficulty of a search task (similarity between target and distractors) determines search slopes, substantial slopes do not always indicate the presence of an attentional capacity limit. Singleton and feature conjunction searches can show little attentional capacity limitation even when the tasks are very difficult by the standard criteria (e.g. Treisman & Gelade, 1980). It seems, therefore, that attention capacity is not affected by search task efficiency, but rather is determined mainly by the nature of the task. Unlimited-capacity processing that is fast and efficient in easy tasks, when faced with harder tasks, may become slow, inefficient, and perhaps inaccurate, but it still fails to show an attentional capacity limitation (i.e. the processing of different parts of the display may be slow or inaccurate, but it is still independent of processing in other parts of the display.)

It should be noted that the RT slope of spatial configuration search (48 ms/item) is greater than that of tasks 1 (35 ms/item) and 2 (39 ms/item). This fact points to one conceivable objection to our above interpretation: it is possible that the influence of an attentional capacity limit is determined by difficulty (search slope) in a highly non-linear fashion; perhaps we see no influence in tasks 1 and 2 only because their efficiency is less than some threshold reflecting the point where attentional capacity limits come into play (e.g. a search slope over 40 ms/item). This explanation, though not impossible, appears unlikely, since naturally the intercept of attention capacity vs. slope function should be zero if the attentional capacity limit effect monotonically increases with the slope (the attention capacity should be non-zero when the slope is non-zero).

Another conceivable objection is that the search might have no attention capacity limit when the set size is smaller than a certain number, but still show attention capacity limits when the set size exceeds that number (see, e.g. Pashler, 1987, for a related suggestion). The current results may be explained by assuming that the number is greater than 16 for feature search but smaller than 16 for spatial-configuration search. While this interpretation cannot be ruled out, it seems somewhat implausible.

For one thing, previous proposals for boundaries of this kind have typically suggested that they emerge at much smaller numbers of items (e.g. 4–5, Fisher, 1982, 1984).

These alternative accounts, in any case, would not overturn our conclusion that a typical very inefficient singleton search or conjunction search exhibits no attentional capacity limit, thus showing that efficiency and attention capacity can be dissociated from each other.

The results also show a significant SUCC advantage in spatial-configuration search (spatial configuration search). Therefore, it appears that this kind of search is indeed fundamentally different from the other two kinds tested here (Wolfe, 1998b). Inefficient feature search arises when the system need to accumulate information for a longer period of time in order to distinguish signal from noise (Eckstein, Thomas, Palmer, & Shimozaki, 2000; Lu & Dosher, 1998). In spatial configuration search, however, the inefficiency is in the selection time. However, even for this type of search, the extra d' gained in SUCC displays was fairly modest (from 0.78 to 1.09, an increase of 40%). If the underlying search process were strictly serial (one by one), the d' should be approximately doubled when the two sub-displays were processed at non-overlapping times, as the SUCC condition permits. Therefore, it seems that even for this type of search, the underlying mechanism is not likely serial, but rather parallel with limited capacity.

In summary, there is no doubt that the pioneering research of Treisman and Gelade (1980) revealed robust and important functional distinctions between different kinds of visual search tasks. However, the present results (along with those of Eckstein et al., 2000; Geisler & Chou, 1995, and others) suggest that a rather different hierarchy of tasks may characterize the relative visual-attention demands of different kinds of visual search tasks.

Acknowledgements

The research was supported by grant MH45584 from the National Institute of Mental Health.

References

Carter, R. C. (1982). Visual search with color. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 127–136.

Carter, E. C., & Carter, R. C. (1981). Color and conspicuousness. *Journal of Optical Society of American*, *71*, 723–729.

Duncan, J. (1980). The demonstration of capacity limitation. *Cognitive Psychology*, *12*, 75–96.

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. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception and Psychophysics*, *62*, 425–451.

Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology Monograph*, *79*(Part 2), 1–16.

Fisher, D. L. (1982). Limited channel models of automatic detection: Capacity and scanning in visual search. *Psychological Review*, *89*, 662–692.

Fisher, D. L. (1984). Central capacity limits in consistent mapping visual search tasks: Four channels or more? *Cognitive Psychology*, *16*, 449–484.

Gardner, G. T., & Joseph, D. J. (1975). Parallel perceptual channels at "deep" processing levels. *Bulletin of the Psychonomic Society*, *6*, 658–660.

Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review*, *96*, 267–341.

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

Green, D. M., & Swets, J. A. (1967). *Signal detection theory and psychophysics*. New York: Wiley.

Lu, Z.-L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research*, *38*, 1183–1198.

Maioli, C., Benaglio, I., Siri, S., Sosta, K., & Cappa, S. (2001). The integration of parallel and serial processing mechanisms in visual search: Evidence from eye movement recording. *European Journal of Neuroscience*, *13*, 364–372.

Nagy, A. L., & Sanchez, R. R. (1990). Critical color differences determined with a visual search task. *Journal of Optical Society of American—A*, *7*, 1209–1217.

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, *320*, 264–265.

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

Pashler, H. (1987). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception and Psychophysics*, *41*, 191–201.

Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.

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. 109–121). Hillsdale, NJ: Erlbaum.

Shiffrin, R. M., & Gardner, G. T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology*, *93*, 78–82.

Townsend, J. T. (1976). Serial and within-stage independent parallel model equivalenceon the minimum completion time. *Journal of Mathematical Psychology*, *14*, 219–238.

Townsend, J. T. (1990). Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science*, *1*, 46–54.

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

Wolfe, J. M. (1998a). Visual search. In H. Pashler (Ed.), *Attention*. London: University College London Press.

Wolfe, J. M. (1998b). What can 1,000,000 trials tell us about visual search? *Psychological Science*, *9*, 33–39.