

# Spacing affects some but not all visual searches: Implications for theories of attention and crowding

Lavanya Reddy

California Institute of Technology, Computation and  
Neural Systems, Pasadena, CA, USA



Rufin VanRullen

Centre de Recherche Cerveau et Cognition, CNRS,  
Université Paul Sabatier, Toulouse, France



We investigated the effect of varying interstimulus spacing on an upright among inverted face search and a red–green among green–red bisected disk search. Both tasks are classic examples of serial search; however, spacing affects them very differently: As spacing increased, face discrimination performance improved significantly, whereas performance on the bisected disks remained poor. (No effect of spacing was observed for either a red among green or an L among + search tasks, two classic examples of parallel search.) In a second experiment, we precued the target location so that attention was no longer a limiting factor: Both serial search tasks were now equally affected by spacing, a result we attribute to a more classical form of crowding. The observed spacing effect in visual search suggests that for certain tasks, serial search may result from local neuronal competition between target and distractors, soliciting attentional resources; in other cases, serial search must occur for another reason, for example, because an item-by-item, attention-mediated recognition must take place. We speculate that this distinction may be based on whether or not there exist neuronal populations tuned to the relevant target–distractor distinction, and we discuss the possible relations between this spacing effect in visual search and other forms of crowding.

Keywords: visual search, serial search, crowding, spacing effects, attention, biased competition, feature integration

Citation: Reddy, L., & VanRullen, R. (2007). Spacing affects some but not all visual searches: Implications for theories of attention and crowding. *Journal of Vision*, 7(2):3, 1–17, <http://journalofvision.org/7/2/3/>, doi:10.1167/7.2.3.

## Introduction

In a typical visual search experiment, a target stimulus is presented among a varying number of distractor items, and subjects are required to locate the target. Performance on such tasks is usually measured in terms of the amount of time required to detect the target (reaction time [RT]). Classical studies of visual attention such as Treisman and Gelade's (1980) influential feature integration theory put forth the idea that visual search is largely of one of two kinds: Parallel search ("pop out") is distinguished by the fact that RT is constant regardless of the number of distractors that are present, whereas serial search is characterized by steep search slopes, indicating that subjects require increasing amounts of time to find the target as more and more distractors are added to the display. Note, however, that there is no strict dichotomy between parallel and serial search slopes but rather a continuum of attentional requirements (Eckstein, 1998; Geisler & Chou, 1995; Joseph, Chun, & Nakayama, 1997; McElree & Carrasco, 1999; Nakayama & Joseph, 1998; Norman & Bobrow, 1975; Palmer, 1994; Treisman & Gormican, 1988; Vergheze & Nakayama, 1994; Wolfe, 1998).

According to the feature integration theory, visual search efficiency depends on whether individual stimuli require attention to be discriminated. Parallel search occurs when target and distractors can be discriminated

in the absence of attention (e.g., when stimuli are defined in terms of simple features, like color or orientation); on the other hand, search is serial when each item in the display requires attention to be recognized (as is the case for more complex stimuli, such as color–orientation conjunctions). In the latter case, attention is required to "bind" the individual features of an object together and create a coherent whole: Thus, it is proposed that attention explores the visual display, focusing on each stimulus in turn until the target object is located and identified (Treisman & Gelade, 1980; Wolfe, Cave, & Franzel, 1989). As the number of items in the display increases, there is a corresponding increase in RT: Hence, the linear search slopes typical of a serial search process.

This particular view has, over the years, become widely accepted as the classical explanation of attention's role in serial search. However, binding is not necessarily the only reason why attention might be required. When multiple stimuli fall into a single neuron's receptive field, target and distractors compete to dominate the neuron's response; according to Desimone and Duncan's (1995) "biased competition" model, attention is needed to resolve this competition. Potentially, this would also result in serial exploration of the visual display, especially when the visual display is highly cluttered. On the other hand, if target and distractors fall into distinct fields, there is no neuronal competition and, hence, presumably no need for attention: Visual search should therefore occur in parallel.

To summarize, attention might be required in serial search tasks for one of two nonexclusive reasons, which, for simplicity, we will term in this paper “attention for recognition,” and “attention against competition”: According to the former idea, the attentional requirements of a task are determined by the properties of individual stimuli (attention aids object recognition), whereas according to the latter, attentional requirements depend on properties of the entire display (attention resolves competition between target and distractors). It follows from the latter that if target and distractors can be separated enough so that they fall into distinct receptive fields, the attentional requirements of the corresponding task should be reduced; on the other hand, when attention is required solely for the purposes of recognition, interstimulus spacing should have no effect on visual search performance.

In our current study ([Experiment I](#)), we investigated spacing effects in two visual search tasks: an upright versus inverted face discrimination and a red–green versus green–red bisected disk discrimination. It has been shown that target detection in both tasks is a serial search process (face discrimination task: [Brown, Huey, & Findlay, 1997](#); [Nothdurft, 1993](#); [Purcell, Stewart, & Skov, 1996](#); but see also [Hershler & Hochstein, 2005](#); [VanRullen, 2006](#); bisected disks task: [Ramachandran, 1988](#)). As mentioned earlier, visual search is typically investigated by varying the number of stimuli presented in the display. Now, we wanted to test the effect of not only varying the number of stimuli displayed but also the distance between individual stimuli. If stimuli are close together, does the resulting “clutter” of the visual scene make the task harder? If the spacing between stimuli is increased, is there an improvement in search performance? We found that as interstimulus distance was increased, visual search performance improved only in the case of the face discrimination task; performance on the bisected disks task was unaffected. Thus, it is possible that the face search takes place according to the “attention against competition” idea, whereas the bisected disks search might fit with an “attention for recognition” scheme.

Is it safe to assume that the obtained spacing effect for the face search task reveals some form of attentional requirement, or could it be just a variant of spatial crowding, which, according to certain authors ([Pelli, Palomares, & Majaj, 2004](#)), occurs independent of attentional allocation? To distinguish these two alternatives, we precued the spatial location of the target in [Experiment II](#) (which was, as in [Experiment I](#), only presented on half of the trials). The cue allowed observers to allocate attentional resources at the corresponding location on every trial. Under these conditions, both the face and the bisected disk discrimination tasks showed a similar influence of spacing; the extent of this spacing effect was compatible with classical observations of crowding ([Bouma, 1970](#); [Intriligator & Cavanagh, 2001](#); [Pelli et al.,](#)

[2004](#); [Toet & Levi, 1992](#)). Thus, our main result, the discrepancy between spacing effects for the face and the disk discrimination tasks in visual search ([Experiment I](#)), was due to attentional resources being divided between the various items in the search display. In the [Discussion](#) section, we evaluate the potential implications of this finding for the current theoretical debate regarding the relation between attention and crowding.

For completeness, we also tested (in [Experiment I](#)) two tasks, which are classical examples of parallel search (color discrimination and “L vs. +” discrimination), and found that interstimulus distance had no significant effect on either task.

## Experiment I: Visual search

We designed a visual search experiment in which stimuli were arranged at varying distances from each other. The main focus of our study was the set of serial search tasks: upright versus inverted face discrimination and red–green versus green–red bisected disk discrimination. For completeness, we also tested two search tasks that have been shown to result in parallel visual search (pop out): color discrimination ([Treisman & Gelade, 1980](#)) and rotated “L” versus rotated “+” discrimination ([Bergen & Julesz, 1983](#)).

## Methods

### Subjects

Ten naive subjects and the two authors performed two sessions daily for 6 days. Within a session, the tasks were presented in eight successive blocks, which were randomly ordered. Each block consisted of 96 trials.

### Apparatus

Subjects were seated in a dark room, 120 cm from a computer monitor connected to a Macintosh G4. The refresh rate of the monitor was 75 Hz, and the display was synchronized with the vertical retrace of the monitor.

### Procedure

[Figure 1](#) is a schematic of the timeline for one trial. Stimuli were arranged around an imaginary circle of radius 1.5°, at any of 12 possible positions (similar to hour markings on a clock; however, to avoid having stimuli positioned along the vertical and horizontal meridians, all positions were shifted by 15° of arc or, in the “clock” analogy, by 30 min of time). Both set size and interstimulus spacing (hereafter “spacing”) varied across trials. Note that spacing refers to the straight-line (rather

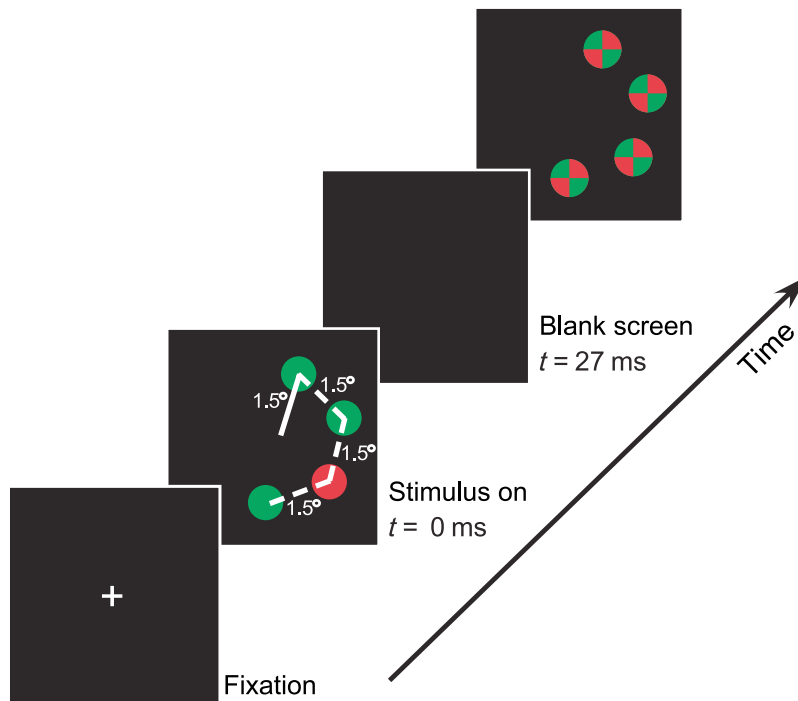


Figure 1. Schematic timeline for one trial in [Experiment I](#) (visual search). Stimulus eccentricity was fixed at  $1.5^\circ$ , illustrated by the solid white line on the stimulus display. Interstimulus spacing represents the straight-line distance between adjacent stimuli: It is indicated by the dashed white line. (Note that the lines representing spacing and eccentricity were not displayed during the actual experiment; they are shown here only in the interests of clarifying our experimental protocol.) Fifty percent of all trials contained one target stimulus. At the end of the trial, subjects used the mouse to report whether or not a target stimulus was present. In this particular case, the task illustrated is color discrimination, but the same experimental protocol was employed for all four tasks. For this trial, the set size is four and the spacing is  $1.5^\circ$ . (Both these parameters varied across trials. See [Figure 2](#) for more details.) All stimuli were masked individually following presentation. SOA was adjusted individually for each subject and each task to avoid saturation in performance. The resulting SOAs, consistently below 320 ms for all subjects and tasks, prevented extensive exploration of the arrays using eye movements. Figure not to scale.

than arc-length) distance between adjacent elements. Search arrays were displayed for 27 ms.

All stimuli were masked after presentation, the type of mask presented being particular to each task. In training sessions prior to the real experiment, stimulus onset asynchrony (SOA) was adjusted individually for each subject and for each task: That is, SOA was decreased when performance on a 96-trial block exceeded 90%, until performance eventually stabilized. This procedure was designed to avoid saturation in performance. SOAs varied between 40 and 320 ms on different tasks and for different subjects (see below). This range of SOAs, together with our short display time (27 ms), can be expected to limit the exploration of the search arrays using eye movements.

Half of the trials contained one target stimulus. Subjects responded to each trial in a “go/no-go” fashion: That is, they held down the mouse button to start trials, releasing it (as fast as possible) only if they detected the target stimulus. Thus, RTs were recorded only for the “target-present” trials. Note that subjects were not instructed to favor RTs over accuracy. In fact, because the visual displays were masked, RT was only expected to vary

mildly between the different conditions; accuracy was thus our main dependent variable.

The small size of the imaginary circle on which the stimuli could appear ( $1.5^\circ$  radius) was deliberately chosen based on preliminary experiments because it allowed us to investigate effects at very small spacings. The number of stimuli displayed on any given trial could be two, three, four, or six. Depending on the set size chosen, the distance between adjacent stimuli could be one of the following:  $0.8^\circ$ ,  $1.5^\circ$ ,  $2.1^\circ$ ,  $2.6^\circ$ ,  $2.9^\circ$ , or  $3^\circ$ . The number of possible spacings was constrained by the number of stimuli to be displayed: For a set size of six, stimuli could only be spaced  $0.8^\circ$  apart or  $1.5^\circ$  apart, as these were the only two distances that would accommodate all six stimuli around the circle. For a set size of two, however, all six distances were viable. Although spacing and set size could not be fully independently manipulated in our experimental design, note that our estimates of spacing effects are performed separately for each set size; hence, there can be no confound between these two independent variables in our results.

Note, finally, that within one trial, the same spacing was used between any two adjacent stimuli, except for the

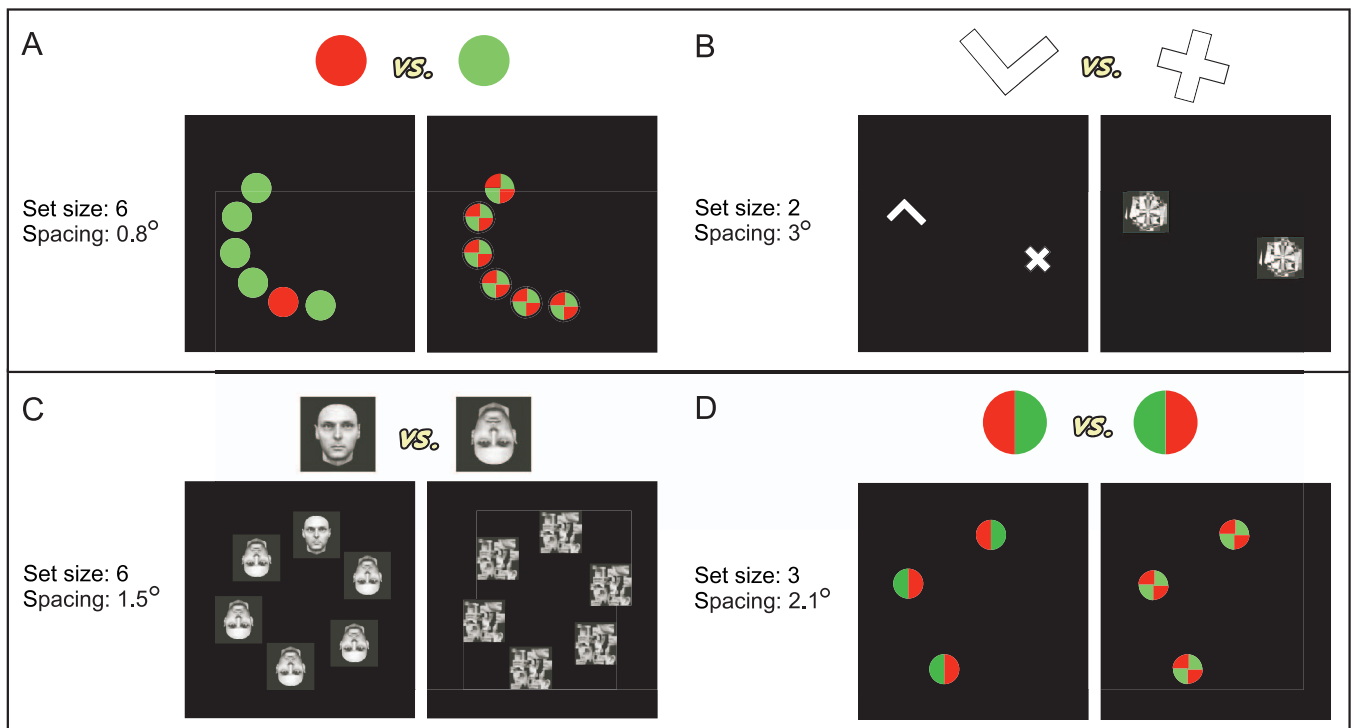


Figure 2. Stimuli and masks used for all four tasks. The tasks are organized in two rows, according to whether they lead to parallel (top row) or serial (bottom row) visual search. Each of the four displays indicates one particular target-present trial for each task, at a certain set size and spacing. Note, however, that these displays are just examples meant to illustrate the experimental protocol: In reality, (i) the target was only present on 50% of all trials and (ii) set size and spacing varied across trials for all tasks. Stimuli are arranged around a circle at any of 12 possible positions (similar to hour markings on a clock). The distance between two adjacent positions (or 1 “hour” on the clock) is about  $0.8^\circ$ . Set size can be two, three, four, or six items, and spacing can take on six values:  $0.8^\circ$ ,  $1.5^\circ$ ,  $2.1^\circ$ ,  $2.6^\circ$ ,  $2.9^\circ$ , and  $3^\circ$  (corresponding to placing stimuli 1, 2, 3, 4, 5, or 6 “hours” apart in the clock analogy). Note that spacing is constrained by set size: for example, a set size of six (Panels A and C) can only be displayed with two possible spacings— $0.8^\circ$  (1 “hour”) or  $1.5^\circ$  (2 “hours”). Within a trial, the spacing between any two adjacent stimuli is the same, except for items at the outer edges: For example, in Panel A, the spacing is  $0.8^\circ$ , except for the two outermost stimuli between which spacing is obviously greater. Figure not to scale.

instance when the stimuli displayed formed an arc—rather than a complete circle—as in Figure 1: In this case, the stimuli located at either end of the arc were separated by a different (larger) spacing. Figure 2 is a diagrammatic representation of these concepts.

### Tasks

We investigated two serial search tasks and two parallel search tasks.

#### Serial tasks

- **Face discrimination.** Subjects had to indicate whether or not an upright face was present among inverted faces. The target stimulus was obtained from the Max Planck Institute for Biological Cybernetics, Tübingen, Germany (Troje & Bühlhoff, 1996; <http://faces.kyb.tuebingen.mpg.de>). This face was converted to gray levels and then rotated through  $180^\circ$  to obtain the distractor stimulus. Face stimuli subtended roughly  $0.7^\circ$ . The mask was a scrambled

face. Across the 12 subjects who performed the main experiment, SOAs ranged from 173 to 320 ms (average, 294 ms).

- **Bisected two-color disks.** Subjects had to discriminate whether or not a red–green bisected disk was present among green–red bisected disks. The axis along which the color changed was always vertical. Stimuli subtended  $0.5^\circ$  in diameter. The mask, a combination of target and distractor symbols, resembled a pie with four equal sections, which were alternately colored red and green. Across subjects, SOAs ranged from 133 to 320 ms (average, 241 ms).

#### Parallel tasks

- **Color discrimination.** Subjects had to indicate whether a single target (red disk) was or was not present among the distractors (green disks). The red and green equiluminant patches subtended  $0.5^\circ$  in diameter. The mask was the same as that used in the bisected



disks discrimination task (described previously). Across the 12 subjects, SOAs ranged from 47 to 87 ms (average, 67 ms).

- (Rotated) “L” versus (rotated) “+” discrimination. Subjects had to discriminate whether or not a randomly rotated letter “L” was present among randomly oriented “+” signs. Target and distractors subtended  $0.6^\circ$  and were constructed by placing the same two bars perpendicular to each other. These stimuli were then randomly rotated at display time. The mask was a combination of gray polygons of differing luminances and was larger than the masked stimulus ( $0.8^\circ$ ). Across subjects, SOAs ranged from 53 to 200 ms (average, 89 ms).

### Performance measurement

We estimated performance in two ways: using RTs (a typical measure of visual search performance) as well as  $d'$  (a measure of detectability; Green and Swets, 1966), which is defined as follows:

$$d' = z(\text{HIT}) - z(\text{FA}),$$

where HIT denotes the proportion of trials on which the target was correctly detected, FA denotes the proportion of false alarms (trials on which subjects reported the presence of a target although no target was present), and  $z$  denotes the inverse function of the cumulative normal distribution.

## Results

Subjects performed four visual search tasks. As described in the [Methods](#) section, both set size and interstimulus spacing varied from trial to trial. Consequently, in our analysis, we first investigated the *overall* effects of set size and spacing and then examined each of these factors individually.

### Overall effects of set size and spacing

[Figure 3](#) is a plot of performance (Panel A, in terms of  $d'$ ; Panel B, in terms of RTs) as a function of both set size and spacing, for each of the four tasks. The row-by-row arrangement of tasks reflects whether visual search is “parallel” (top row of each panel) or “serial” (bottom row of each panel).

A two-way ANOVA (Set Size  $\times$  Spacing) reveals the following.

#### Main effect of set size on $d'$ and RT

For the face and bisected disks tasks,  $d'$  clearly decreases and RT increases as set size is increased,  $d'$ :

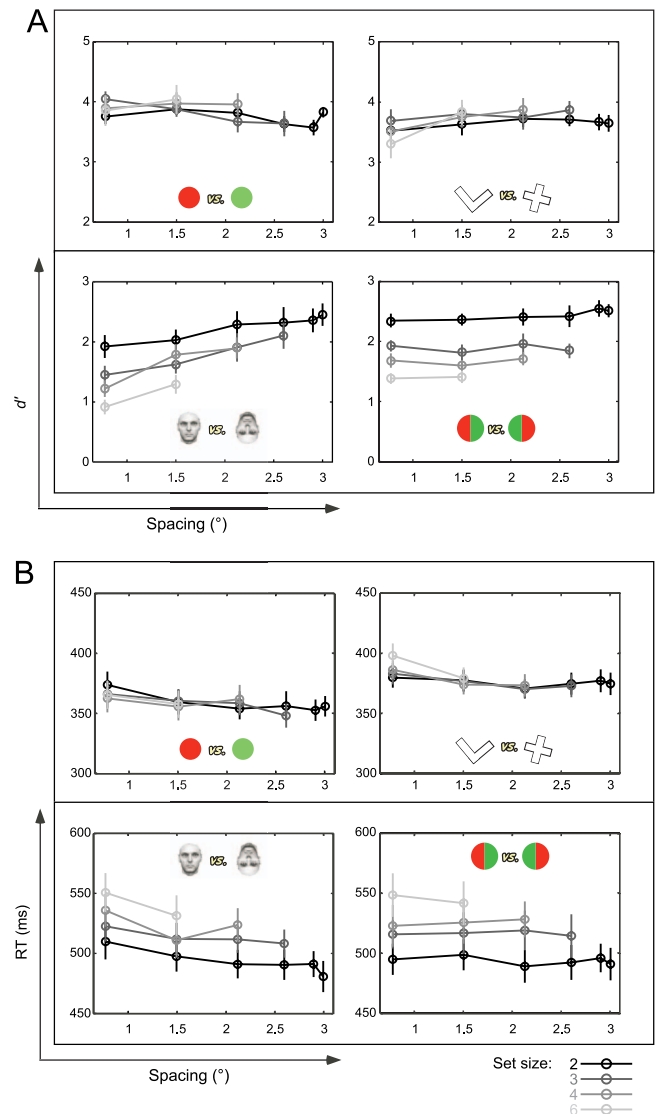


Figure 3. Search performance as a function of spacing and set size, for each of the four tasks. Error bars represent standard error of the mean. In Panels A and B, the search tasks are organized according to their defining characteristics (parallel or serial) as in [Figure 2](#). (A) Performance in terms of  $d'$ . The effect of set size on performance is immediately evident for the face discrimination and bisected disks tasks, for which  $d'$  decreases with increasing set size (as we would expect for serial search tasks). Also, for the face task, performance improves significantly as spacing is increased—in keeping with our prediction that if local competition is reduced (by increasing spacing between stimuli), visual search performance should improve. There is no such effect for the other three tasks. (B) Performance in terms of RTs. The effect of set size on RT is clear for the face and bisected disks search tasks, for which RTs increase significantly with increasing set size (as expected for serial search tasks). There is no significant main effect of spacing on RT for any of the four tasks.

$F(3,165) > 7, p < .001$ ; RT:  $F(3,165) > 3.5, p < .01$ . For the color task,  $d'$  increases slightly with set size; however, there is no significant change in RT,  $d'$ :  $F(3,165) = 5.8, p < .001$ ; RT:  $F(3,165) = 0.06, p > .1$ . For the L versus + task, neither  $d'$  nor RT shows a significant change with set size,  $d'$ :  $F(3,165) = 0.4, p > .05$ ; RT:  $F(3,165) = 0.49, p > .1$ . As will be discussed in the next section, these results confirm that, as expected, the color discrimination and the L versus + tasks show parallel search behavior, whereas the face discrimination and the bisected disks tasks show serial search behavior.

#### Main effect of spacing on $d'$ and RT

$d'$  increases with spacing,  $F(5,165) = 3.7, p < .005$ , only in the case of the face task. There is no effect of spacing on  $d'$  for any of the other three tasks,  $F(5,165) < 1, p > .05$ . Thus, performance on one of the serial search tasks—face discrimination—improves as interstimulus spacing increases, whereas the other—bisected disks discrimination—is unaffected by spacing. This is the major finding of this study. If increasing spacing can, in some cases, reduce local neuronal competition, then it seems that only the face task benefits here. In the sections below, we investigate these results in more detail.

Spacing does not significantly affect RTs for any of the tasks,  $F(5,165) = 0.97, p > .05$ . However, this is not unexpected: As will be described later, variations in RT are severely restricted in our protocol (by use of a go/no-go paradigm and backward masking).

#### Set Size $\times$ Spacing

None of the interactions are significant,  $d'$ :  $F(6,165) = 0.4, p > .05$ ; RT:  $F(6,165) = 0.17, p > .05$ . However, this might be a result of the unbalanced task design: Not all set sizes exist at every spacing (e.g., set size of six is defined only at the first two spacings), and thus, some data points are missing in the two-way ANOVA calculation.

We now study the effects of set size and spacing individually and in more detail.

#### Set size effects

We first examine how performance for each task is affected by set size. In so doing, we can also verify that performance matches the behavior described in the visual search literature: that is, that the so-called serial and parallel tasks really behave in the expected way under the particular conditions of our experiment.

Typically, set size effects are determined in response-terminated visual search (where search arrays remain visible until the subject responds) by calculating search slopes: that is, the average change in RT corresponding to an increase in set size (Treisman & Gelade, 1980; Wolfe, 1998). However, we used a go/no-go response procedure and backward-masked arrays, both of which tend to limit the variability of RTs (VanRullen, Reddy, & Koch, 2004). In this case, percentage correct and  $d'$  are more appro-

priate indicators of search performance (Bergen & Julesz, 1983; VanRullen et al., 2004; Wolfe, 1998). Indeed, numerous studies have used  $d'$  as a measure of set size effects (see Verghese, 2001, for a review). Therefore, we report set size effects in terms of  $d'$  (a metric we term “set size score”), but for the sake of completeness, we present the corresponding RT data (i.e., search slopes) as well.

#### Set size score

Set size score reflects the average change in  $d'$  that corresponds to an increase in set size. Because all possible set sizes (two, three, four, or six items) are viable only at the two smallest spacings ( $0.8^\circ$  and  $1.5^\circ$ ), set size score is based on  $d'$  averaged only over these two spacings. As mentioned in the Introduction section, we would like to test the idea that serial search behavior can (sometimes) result from neuronal competition at small spacings. It thus behooves us to investigate set size effects at these smaller spacings first and later examine the changes in performance associated with larger spacings.

Figure 4A is a representation of set size score ( $d'/\text{item}$ ). Observe the performance on the two tasks expected to be serial: upright versus inverted face discrimination and bisected disks. For these tasks,  $d'$  decreases,  $t(11) > 9.3, p < .001$ , with increasing set size. On the other hand, in keeping with the behavior of parallel search tasks (for a more sophisticated perspective, however, see Egeth, 1977; Eriksen & Spencer, 1969), there is no reduction in  $d'$  for the “L versus +” task,  $t(11) = 0.1, p > .05$ , whereas, in the case of the color discrimination task,  $d'$  even increases slightly,  $t(11) = 4.2, p < .002$ . Thus, serial and parallel behaviors are exhibited in our experimental conditions as expected from the visual search literature.

#### Search slopes

Figure 4B represents search slope data (RT/item). These results are equivalent to those obtained in terms of set size score (above): For the face and bisected disks tasks, RT increases,  $t(11) > 4.2, p < .005$ , with increasing set size, as is representative of serial search; there is no significant increase in RT for the color discrimination task,  $t(11) = 1.0, p > .05$ , and although it appears that RT increases significantly for the “L versus +” task,  $t(11) = 3.1, p < .01$ , the search slope itself is only about 3 ms/item, which is indicative of parallel search (Wolfe, 1998). These findings once again confirm the parallel/serial nature of these tasks.

Note that the search slopes reported here (about 10–12 ms/item) are lower than typical serial search slopes (30–40 ms/item and higher; Wolfe, 1998). However, this is not entirely unexpected because our calculations are based on only target-present trials (go/no-go paradigm), for which search is expected to terminate after half the items (on average) have been inspected. Further, visual search slopes are known to be greatly restricted by short presentation times (McElree & Carrasco, 1999) and backward masking (VanRullen et al., 2004). The average

RTs (and standard errors of the mean) for 12 subjects (target-present trials) were as follows: 360 ( $\pm 11.1$ ) ms for the color task, 379 ( $\pm 9.0$ ) ms for the L versus + task, 519 ( $\pm 13.6$ ) ms for the face discrimination task, and 521 ( $\pm 15.9$ ) ms for the bisected disks.

**Spacing effects**

Having investigated the effect of set size on performance, we can now examine spacing effects in more

detail. We have already noted that increasing spacing has a significant effect on performance only for the face task (cf. the two-way ANOVA in the [Main effect of spacing on  \$d'\$  and RT](#) section; see also [Figure 3](#)). But what is the size of this effect?

*Spacing effects in terms of  $d'$*

We consider the average *change* in  $d'$  as a function of spacing, using our smallest spacing ( $0.8^\circ$ ) as a reference point. That is, we express all subsequent  $d'$  values as a percentage of the  $d'$  at this reference point. [Figure 5](#) is a plot of this “normalized  $d'$ ” versus spacing for all four tasks, organized in the familiar row-by-row arrangement (parallel or serial). It is obvious from the figure that normalized  $d'$  increases strongly with spacing in the case of the face task: On average, increasing spacing by  $1^\circ$  results in a 50% increase in visual search performance. For the other tasks, increasing spacing does not lead to a similarly strong improvement in normalized  $d'$ . As mentioned earlier, both the face discrimination and the bisected disks tasks are typical examples of serial search, and thus, one might expect that the effect of increasing interstimulus spacing would be similar for both tasks; however, when spacing is increased, the improvement in the face task is an order of magnitude higher than that observed for the bisected disks.

Indeed, when repeating the previous two-way ANOVA but using this normalized  $d'$  data, we find a significant main effect of spacing only for the face and the L versus + tasks,  $F(5,165) = 5.6, p < .001$  and  $F(5,165) = 4, p < .005$ , respectively; recall, however, that for the latter, the corresponding effect on non-normalized  $d'$  was not

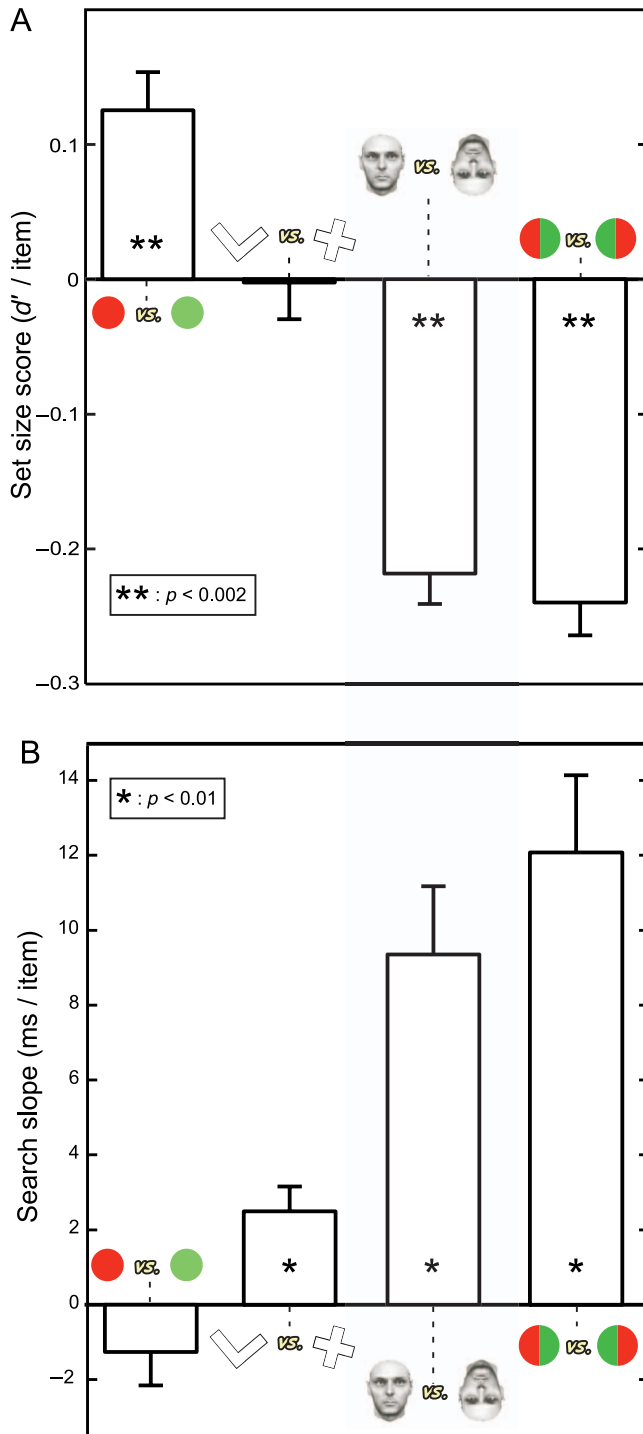


Figure 4. (A) Set size score. To study the effects of set size on search performance (and, thus, test whether a particular task exhibits parallel or serial search behavior), we introduced a new measure, similar to the search slope: set size score. Set size score ( $d'$ /item) reflects how much detectability ( $d'$ ) is gained or lost with every increase in set size and is, thus, analogous to search slopes. Error bars represent standard error of the mean. For the two tasks toward the far right (upright vs. inverted face discrimination, bisected disks),  $d'$  decreases significantly with increasing number of items, representative of serial search. For the color discrimination and the rotated L versus + tasks, there is little or no reduction in  $d'$  with increasing set size, as is expected for parallel search. In fact, in the case of the color task,  $d'$  even increases with increasing set size. \*\*Set size score for the task in question is significantly different from zero at the  $p < .002$  level. (B) Search slopes. A more classical way to study set size effects on search performance is to use RTs, rather than  $d'$ . Search slopes (ms/item) reveal a similar separation between parallel and serial tasks. The relatively low amplitude of the observed search slopes can be explained by the fact that we use backward-masked visual search and record RTs for target-present trials only. \*Search slope for the task in question is significantly different from zero at the  $p < .01$  level.

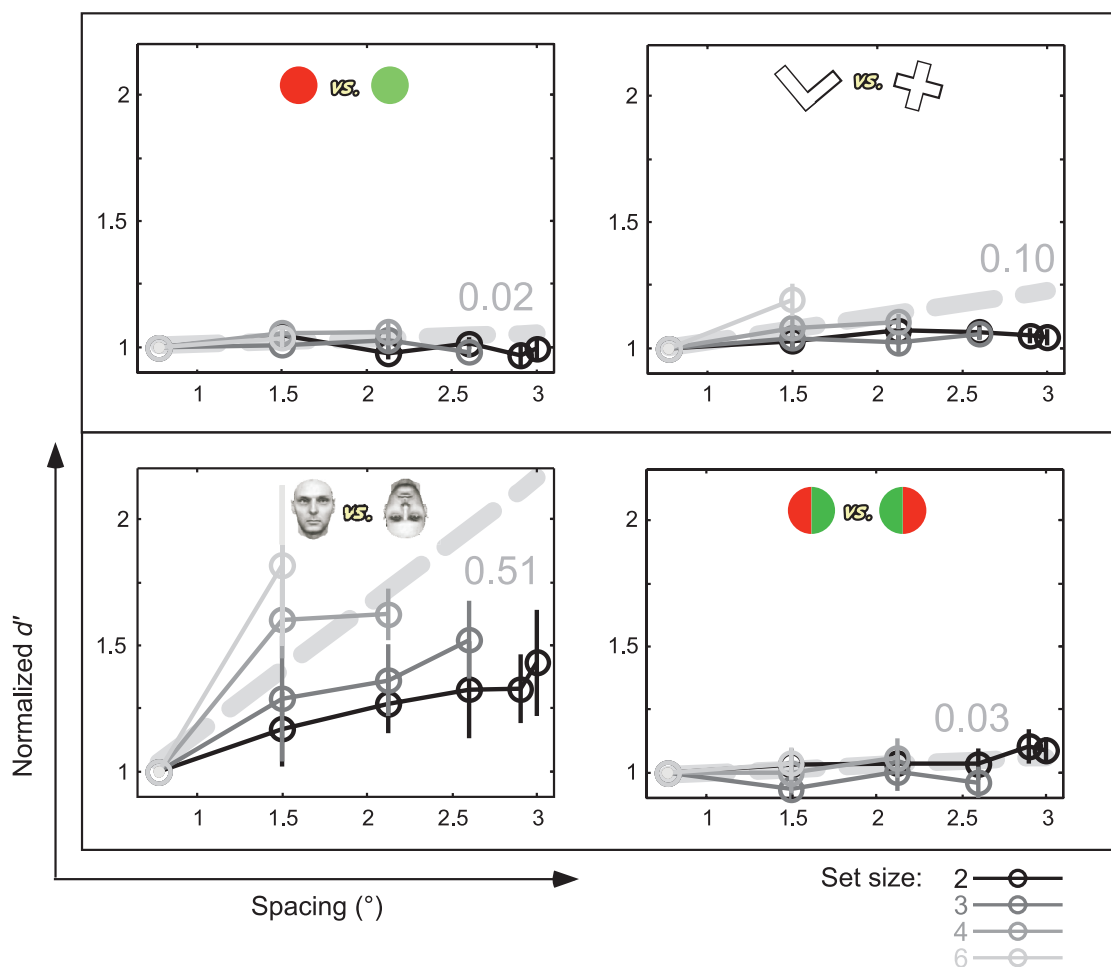


Figure 5. Normalized search performance. Normalized  $d'$  (i.e., the average change in  $d'$  compared with the  $d'$  value at the smallest spacing) as a function of spacing. The dashed line in each quadrant represents the average linear fit of the four set size curves (obtained by averaging the slopes from the four curves). The slope of this fit is indicated next to the line. Note that, although both the face discrimination and the red–green versus green–red disk discrimination are typical examples of serial search tasks, the effects of varying interstimulus spacing are strikingly different: For the face task, normalized  $d'$  increases dramatically with spacing, whereas for the bisected disks, the effect of spacing is negligible. (The two parallel search tasks [top row] are likewise unaffected by spacing.)

significant (see the [Main effect of spacing on  \$d'\$  and RT](#) section).

The curves for the different set sizes in the case of the face search task ([Figure 5](#), bottom left) reveal another interesting point: The effect of spacing appears to be larger with higher set sizes. This is not surprising, however, and not incompatible with our hypothesis developed in the [Introduction](#) section because for a given value of spacing, the total amount of competition in the display is larger when more elements are present.

#### Spacing effects in terms of RT

Although it is clear that spacing has a significant effect on  $d'$  performance for the face task, we did not see this main effect repeated in the case of RTs (see two-way ANOVA in the [Main effect of spacing on  \$d'\$  and RT](#) section). As mentioned earlier, this is not surprising given

that RT effects are minimized by the use of masked displays and the go/no-go paradigm. However, for completeness, we can quantify the effect of spacing on RTs as follows: For any given task, for each individual set size, we calculate a “spacing slope”; this is simply the average change in RT that corresponds to each  $1^{\circ}$  increase in spacing. (Note that this type of slope is, thus, different from a search slope because it measures the effects of spacing rather than set size.) These (four) slopes are then averaged over all (four) set sizes. The slopes obtained ( $M \pm SEM$ ) are as follows:  $-7.39 \pm 1.31$  ms/deg for the color discrimination task,  $-11.29 \pm 2.01$  ms/deg for the L versus + task,  $-14.55 \pm 3.30$  ms/deg for the face discrimination task, and  $-2.05 \pm 3.31$  ms/deg for the bisected disks. The most negative slope is thus obtained, as predicted from the  $d'$  results in the [Spacing effects in terms of  \$d'\$](#)  section, for the face discrimination task. This suggests that as spacing is increased, RTs decrease faster



for this task than for any other task. A one-way ANOVA shows a significant effect of the task on these slopes,  $F(3,44) = 4.2$ ,  $p < .01$ . A post hoc test (multiple comparisons) reveals that slopes for the face and bisected disks tasks are significantly different; other pairings of tasks reveal no significant difference.

## Experiment II: Spatially cued search

Although the face and bisected disks discriminations are very similar (in that they are both typical examples of serial search tasks), increasing spacing has markedly different effects on visual search performance for these tasks. Based on ideas developed in the [Introduction](#) section, our interpretation of these different spacing effects could be that attention serves different purposes in the two serial search tasks: “attention for recognition” in the case of the bisected disks task and “attention against competition” for the face task. However, spacing effects are known to occur also in situations in which attention *can* be directly allocated to the target (e.g., Bouma, 1970; Intriligator & Cavanagh, 2001; He, Cavanagh, & Intriligator, 1996; Toet & Levi, 1992), which led certain authors to propose that such “crowding” effects are wholly independent of attention (Pelli et al., 2004). Could the face search task simply suffer from this type of spacing limitation, independent of attentional allocation? How could we explain, in this case, that the bisected disk search task appears immune to this limitation? To address these questions, we performed another series of experiments with the upright versus inverted face and the bisected disks discrimination tasks, now precueing the potential location of the target in advance of the trials. This way, attentional allocation was no longer a limitation for either task because, presumably, the subjects could allocate attention at the relevant location on virtually every trial. If the spacing effect obtained for the face search task was truly independent of attention, the same pattern of results should be replicated here: a significant spacing effect for the face discrimination, but no such effect for the bisected disks (the logic for this experimental manipulation was partly inspired by a related study from Zelinsky, 1999).

## Methods

### Subjects

Two naïve subjects and the two authors participated in this experiment. Subjects performed one session, consisting of 12 face blocks interleaved with 12 disk blocks. Each block had 32 trials, of which 50% were target-present trials.

## Procedure

We intended that the experimental design should remain as close as possible to the protocol of our main visual search experiment ([Experiment I](#)) while, at the same time, introducing a key feature of classical crowding studies, namely, that the location of the target stimulus is fixed on all trials but that the spacing of distractors surrounding the target varies from trial to trial. Hence, at the beginning of each block, we cued the location of the target by a white cross (which appeared for 2000 ms). Then, for the entire length of the block, the target—when present (50% of all trials)—always appeared at this location (with a distractor on either side); otherwise, a distractor appeared at this location (again, with a distractor on each side). Distractors were placed at varying spacings around the target.

The first trial in [Figure 6](#) represents one target-present trial. From block to block, the position of the target was chosen at random so that eventually all 12 “clock positions” around the circle were used as target locations. As in [Experiment I](#), SOAs were adjusted individually for each subject to prevent performance from saturating; thus, each subject performed a number of training blocks to determine their SOAs on each task. However, in contrast to [Experiment I](#), where any given subject could have different SOAs on different tasks, each subject had the same SOA on both tasks (average SOA across subjects, 93.7 ms; range, 40–139 ms).

For the sake of simplicity, we limited the set size to three on all trials; for this set size, four interstimulus spacings were possible ( $0.8^\circ$ ,  $1.5^\circ$ ,  $2.2^\circ$ , and  $2.6^\circ$ ). Note that, except for this restriction on set size and the absence of spatial ambiguity about target location (due to the cue), [Experiment II](#) was similar to the main visual search experiment in all respects (go/no-go protocol, backward masking, stimulus size, eccentricity, etc.).

## Results

### Main experiment

The RT data obtained for [Experiment II](#) (four subjects) were as follows: The average ( $\pm SEM$ ) for the face discrimination task was 480 ( $\pm 35$ ) ms, and for the bisected disks task, it was 515 ( $\pm 17$ ) ms. There was no significant difference between the two,  $t(3) = 1.8$ ,  $p > .05$ ; two-tailed  $t$  test.

[Figure 7](#) is a plot of performance (in terms of percentage correct) as a function of the four possible spacings. We generated psychometric fits for our data using the Weibull function. The function is defined in terms of two parameters:  $\alpha$  is the  $x$  value (in our case, the spacing) at which the fitted performance is 82% of the maximum, whereas  $\beta$  represents the slope of the psychometric curve. The “spatial extent” of crowding is often reported as a specific point between chance and maximum performance (Bouma, 1970; Intriligator & Cavanagh,

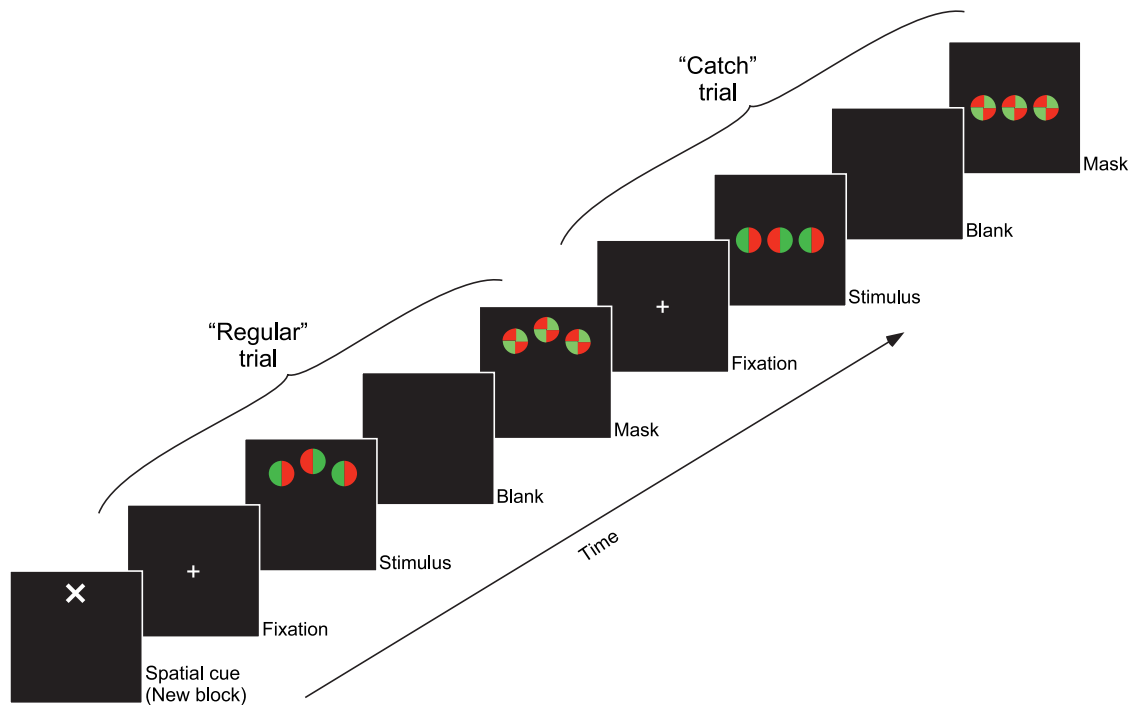


Figure 6. Schematic timeline for [Experiment II](#) (and the associated fixation control). In [Experiment II](#), a white “X” cued the location of the target at the beginning of each block. Then, for all succeeding trials within this block, the target—when present—appeared at this location, as illustrated in Trial 1; for target-absent trials, the cued location was occupied by a distractor. From block to block, the position of the cue changed randomly. The control experiment was identical to [Experiment II](#) in all respects, except for the addition of a few catch trials: On these trials, which occurred randomly within the blocks, the relevant stimulus appeared, not at the location of the cue, but, instead, at the location of the fixation cross. A distractor was placed on each side of this central stimulus, at a horizontal distance of  $0.8^\circ$ . As in other trials, the subject was instructed to determine whether this central stimulus was a target or distractor. If subjects were correctly fixating at the center (as instructed), performance on these catch trials should exceed performance on the regular trials (at the corresponding spacing of  $0.8^\circ$ ); otherwise, catch performance should be worse. In the figure, the task illustrated is the red–green versus green–red bisected disk discrimination, but the same protocol was employed for the upright versus inverted face discrimination task as well (which was the only other task tested here). Set size on all trials in these two experiments was restricted to three stimuli; however, as in [Experiment I](#), spacing varied from trial to trial (only for “regular” trials). All stimuli were masked individually following presentation. SOAs were adjusted individually for each subject; across subjects, SOAs ranged from 40 to 139 ms. Figure not to scale.

2001), here represented by  $\alpha$ . The average  $\alpha$  values ( $\pm SEM$ ) for four subjects were as follows:  $0.60^\circ (\pm 0.08^\circ)$  for the face task and  $0.77^\circ (\pm 0.26^\circ)$  for the bisected disks. There was, thus, a pronounced spacing effect in this experiment not only for the face task but also for the bisected disks discrimination. In fact, the spacing effect was larger (i.e., of larger spatial extent) for the bisected disks task, although the difference was not significant: A two-tailed  $t$  test revealed that there was no significant difference between  $\alpha$  values for the face and disk tasks,  $t(3) = 1.0, p > .05$ ; a  $t$  test on the  $\beta$  values for the two tasks similarly revealed that the slopes of the two psychometric fits were not significantly different,  $t(3) = 0.77, p > .05$ . Note, finally, that the observed values for the extent of this spacing effect are fairly close to the  $0.75^\circ$  value that would be expected from the “crowding extent  $\sim 1/2$  eccentricity” rule reported in most crowding experiments (Bouma, 1970; Intriligator & Cavanagh, 2001; Pelli et al., 2004; Toet & Levi, 1992).

To summarize, if we remove the ambiguity in target location (as is typically done in crowding experiments), then the limiting factor on performance in both search tasks appears to be the spatial resolution of a single mechanism—or the resolution of two mechanisms with comparable spatial properties (today’s literature is divided: resolution limit of the attentional focus or resolution limit of the so-called “integration fields.” This issue will be discussed further in the [Attention and crowding](#) section). However, when the target location is not cued (as is generally the case in visual search studies), spacing effects reveal different underlying mechanisms for the different search tasks (faces or bisected disks discrimination), with clearly distinct spatial properties. Locating an upright face target among inverted faces is facilitated at increasing stimulus spacings (as proposed by the “attention against competition” idea), whereas for the bisected disks, locating the target appears equally difficult at all spacings, presumably because every item in the

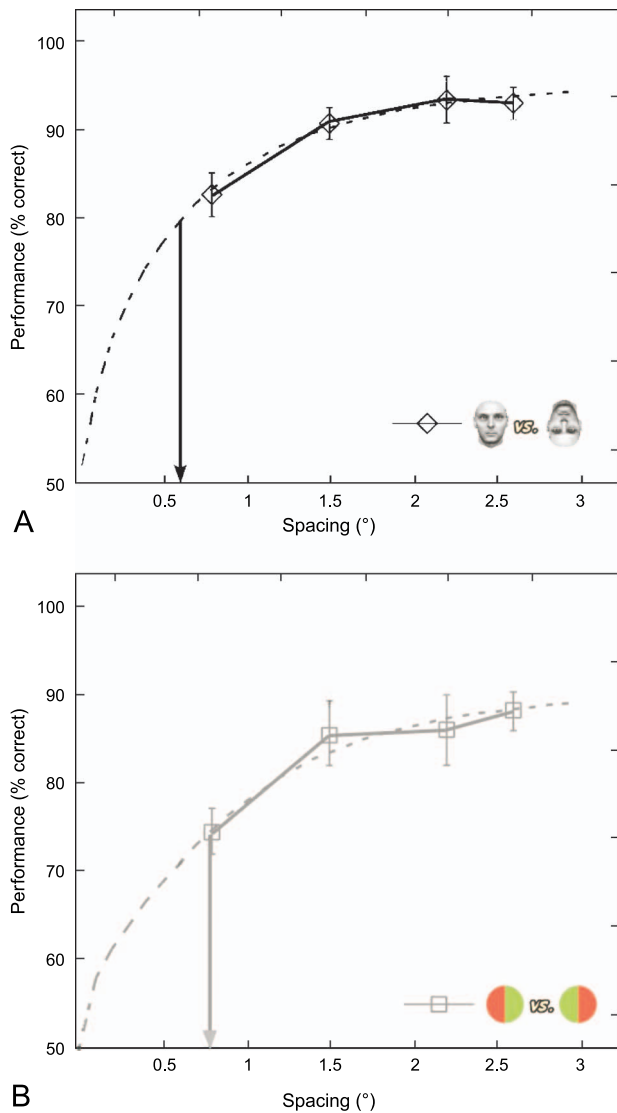


Figure 7. Spacing effects in precued search. Performance (in terms of percentage correct) as a function of spacing, for the face discrimination and the bisected disks tasks. Error bars represent standard error of the mean. The dashed lines represent psychometric fits of our data, obtained using Weibull functions. The spatial extent of crowding is typically reported as the abscissa value corresponding to a given performance level (e.g., 82% of maximum performance). From the psychometric fits, we obtained the spacing value corresponding to this level of performance for each task (indicated by the solid arrows), as well as the slope of the psychometric curves. The  $t$  tests revealed that the critical spacing values obtained from the Weibull functions for the face and bisected disks tasks were not significantly different from each other; neither were the slopes of the psychometric curves. Thus, it is evident that the marked qualitative difference in visual search performance between the two discrimination tasks (Figure 5) vanishes when attention is directed to the target location: Spacing effects now affect both tasks equally, a result that is compatible with classical crowding studies.

display needs attention (as would be expected from the “attention for recognition” idea).

### Controlling for fixation

In Experiment II, we did not control for eye movements; thus, although subjects were instructed to maintain fixation at the center of the screen for the length of the experiment, it is possible that during the 2,000-ms-long cue presentation, subjects moved their eyes to the cued location and maintained fixation there for the rest of the block. To check that subjects were correctly maintaining fixation at the center of the display as instructed, we repeated the main experiment described above (i.e., Experiment II), now incorporating an eye-movement control.

Three naïve subjects and one of the authors (L.R.) participated in this control. L.R. and one of the subjects had also participated in the main version of Experiment II. As in that experiment, after a brief training session (to determine SOAs), these subjects performed one test session, consisting of 12 face blocks randomly interleaved with 12 disk blocks. Each block had 32 regular trials, plus four “catch” trials (see below and Figure 6 for illustration).

The procedure for this control was similar to that of Experiment II in all respects, except for the addition of four catch trials on each block: In these trials, which occurred at random within the block, a target or distractor stimulus was presented—not at the cued location but, instead, at the center of the display. Two distractors were placed horizontally at a distance of  $0.8^\circ$  on either side of this central stimulus. As in the previous experiment, subjects had to report whether the central stimulus was a target (50% likelihood). Thus, if fixation was correctly maintained at the center, performance on these catch trials should exceed performance on the regular trials (at the corresponding spacing, i.e.,  $0.8^\circ$ ). However, if the subjects disregarded instructions and fixated the target location, catch performance should turn out to be worse than performance on regular trials. Figure 6 (Trial 2) illustrates a target-present catch trial.

Subjects’ average performance on the catch trials tended to be high: 98% ( $\pm 0.85$ ) in the face task and 92% ( $\pm 3.5$ ) in the disk task (no significant difference between the two performances:  $p > .05$ ). We noted that for both the face and the disks tasks, performance on the catch trials was significantly higher,  $t(3) > 4.9$ ,  $p < .01$ , than that on the corresponding regular trials, that is, those trials for which spacing was  $0.8^\circ$ ; this indicates that, in general, subjects were correctly fixating the center of the display.

The average  $\alpha$  values ( $\pm SEM$ ) for four subjects were as follows:  $0.55^\circ$  ( $\pm 0.08^\circ$ ) for the face task and  $0.73^\circ$  ( $\pm 0.21^\circ$ ) for the bisected disks (these values are derived from psychometric fits obtained with the regular, not the catch trials). These values are very similar to those obtained

in [Experiment II](#) ( $0.60^\circ \pm 0.08^\circ$  and  $0.77^\circ \pm 0.26^\circ$ , respectively), and as in that experiment, there is no significant difference between  $\alpha$  values for the face and disks tasks,  $t(3) = 2.08$ ,  $p > .05$ , which indicates that the spatial extent of the effect was similar for both tasks. Also, as in [Experiment II](#), a  $t$  test on the  $\beta$  values for the two tasks shows that the slopes of the psychometric fits are not significantly different,  $t(3) = 1.68$ ,  $p > .05$ . Thus, the main result of the previous experiment, replicated here under more controlled conditions, is not attributable to faulty fixation.

## Discussion

### “Attention for recognition” versus “attention against competition”

The dominant view of visual search suggests that serial search occurs because attention is needed to recognize each object in the display; thus, as the size of the display increases, RTs and error rates increase (Treisman & Gelade, 1980).

Alternatively, it has been shown that when target and distractors fall into a single receptive field, attention is necessary to resolve the neuronal competition that ensues (Desimone & Duncan, 1995). Presumably, this finding can be extended to a visual search situation: Attention would be required—and, thus, search should be serial—when target and distractors are simultaneously placed within the receptive field of neurons that can discriminate these stimuli.

In the current study, we set out to investigate this dichotomy: Do some serial search tasks fit the predictions made by the former theory (“attention for recognition”), whereas others behave according to the latter idea (“attention against competition”)? Particularly in the latter case, we would expect that search performance should improve as target–distractor spacing is increased.

A previous study did, in fact, investigate the effect of varying interstimulus spacing on visual search performance (Cohen & Ivry, 1991) to explain a body of conflicting evidence (Pashler, 1987; Wolfe et al., 1989). The authors found, for a classical example of serial search (a shape–color conjunction task), that search performance is typical of serial search when the target and distractors are “clumped” or crowded together (spacing of  $0.62^\circ$ ); when target and distractors are “spread out” ( $1^\circ$  or more apart), search performance improves significantly. In this study, we investigated the effects of interstimulus distance on search performance in a more extensive manner: We tested up to six different interstimulus spacings and two serial search tasks (and for completeness, we tested two parallel search tasks as

well). In keeping with the results of Cohen and Ivry (1991), we demonstrate for one typical serial search task (upright vs. inverted face discrimination) that search performance improves steadily as spacing is increased. However, we also show that performance on another serial search task (red–green vs. green–red bisected disk discrimination) is unaffected by changes in spacing. Why the discrepancy? Until more examples of serial search tasks that fit with either scheme can be provided (i.e., with or without an effect of spacing), any answer to this question must be limited to speculation.

One hypothesis, related to the “attention against competition”/“attention for recognition” conjecture that we put forward in the [Introduction](#) section, could be that the distinction is based on whether or not there exist neuronal selectivities that can differentiate between the target and distractor stimuli. When such populations exist, it is expected that target–distractor competition within a single receptive field should affect visual search performance and, furthermore, that attention will be required to resolve this competition (Desimone & Duncan, 1995). This might be the case for the face discrimination task. On the other hand, for tasks that cannot rely on preexisting neuronal selectivities to aid recognition, we expect that attention will be needed for each object in turn to bind individual features of the object into a coherent percept (Treisman & Gelade, 1980). In this case, increased competition will have minimal effects on visual search performance because attention is needed in any case. This might be the situation for the bisected disks task.

This hypothesis is similar to one we proposed recently (VanRullen et al., 2004). We predicted that visual discrimination tasks involving natural, familiar, and meaningful categories (e.g., faces), which presumably rely on existing neuronal selectivities in high-level visual areas, would only require focal attention in cluttered displays. This is compatible with electrophysiological findings that demonstrate that neurons in the inferotemporal cortex in monkeys and in the medial–temporal lobe in humans are selective to particular categories of objects such as animals and faces or even images of specific individuals (Gross, Rocha-Miranda, & Bender, 1972; Kreiman, Koch, & Fried, 2000; Oram & Perrett, 1992; Perrett, Rolls, & Caan, 1982; Quiroga, Reddy, Kreiman, Koch, & Fried, 2005). Furthermore, neuronal populations in IT have been found to respond more strongly and/or faster to upright faces than to inverted faces (Jeffreys, 1989; Kanwisher, Tong, & Nakayama, 1998; Perrett et al., 1988); thus, competition between upright and inverted faces in our visual search task might be expected to impair discrimination for these populations of cells (Reynolds, Chelazzi, & Desimone, 1999).

The results obtained in the case of the face discrimination task are compatible with the (at first seemingly contradictory) results obtained from a number of recent natural scene categorization experiments. First, we reported that a single natural scene can be processed



effortlessly (e.g., does it contain an animal or not?) even when attention is occupied elsewhere (dual-task situation; Li, VanRullen, Koch, & Perona, 2002). Rousselet, Fabre-Thorpe, and Thorpe (2002) demonstrated that for displays of only two scenes, the discrimination of animal versus nonanimal scenes can occur in parallel when the scenes (presented  $7.2^\circ$  apart) are in different hemifields. Subsequently, it was reported that this ability is significantly impaired when the display consists of four scenes, with one scene in each quadrant (Rousselet, Thorpe, & Fabre-Thorpe, 2004). Finally, we reported that when the stimuli (up to 16 scenes) are separated by about  $3^\circ$ , visual search is clearly a serial process (VanRullen et al., 2004). Presumably, when the scenes are separated by large-enough distances, they tend to fall into distinct receptive fields, which support a parallel recognition, whereas at smaller spacings, competition within a receptive field leads to an impairment in search performance (i.e., serial search). In a more recent study (VanRullen, Reddy, & Fei-Fei, 2005), we showed that two faces can be processed simultaneously (in an upright vs. inverted face discrimination task) when they are placed  $8^\circ$  apart, but not when they are placed  $3^\circ$  apart (keeping stimulus eccentricity constant throughout the various spacing conditions); the same result was shown to hold for the parallel categorization of two natural scenes (e.g., animal vs. nonanimal); however, in the same paradigm, two bisected colored disks could not be processed in parallel, regardless of the distance between them. This qualitative dissociation is reminiscent of the one obtained here in [Experiment I](#).

## Attention and crowding

The current literature on the “crowding effect” is divided between authors who attribute this decrease in stimulus discriminability at small spacings to a limitation of the spatial resolution of the attentional focus (He et al., 1996; Intriligator & Cavanagh, 2001) and authors who argue that the limit is the size of the relevant integration fields for the task at hand, independent of attentional resources (Pelli et al., 2004). How do the present results speak to this debate? The clear qualitative change between the patterns of results obtained in the main visual search experiment (spacing effect for face discrimination but not for bisected disk discrimination) versus the precued search (similar spacing effects for both tasks) suggests that attention does have something to do with at least some of the reported spacing effects.

The idea that multiple stimuli falling into a single integration field can lead to a decrease in discriminability at small spacings (Pelli et al., 2004) is, at first sight, very similar to our speculation described in the previous section. However, the integration field hypothesis of Pelli et al. (2004) raises two issues that appear difficult to address in their framework.

First, is it reasonable to assume that such integration fields could exist for any discrimination task involving

arbitrary combinations of simple features? While it seems that this idea could well be validated for familiar object categories that subjects experience on a regular basis (such as faces, scenes), it is more difficult to believe that all possible combinations of features in all possible spatial arrangements could be represented by dedicated integration fields—this would inevitably lead to combinatorial explosion of the required number of integration fields. It could be that the postulated integration fields are not constant entities; instead, there could be some specific process that has the ability to dynamically create, according to current task demands, the relevant integration field in which arbitrarily specified features for the task could be “bound.” Such a process, however, would look an awful lot like attention, as defined, for example, by the feature integration theory (Treisman & Gelade, 1980)—and the conceptual distinction between the spatial resolution of attention and that of the postulated integration field would, then, lose much of its substance.

Second, the idea that multiple stimuli at small spacings can prevent the selective activation of the relevant integration fields is fully compatible with neurophysiological observations of the detrimental effect of stimulus competition within neuronal receptive fields. Yet, the idea that this has nothing to do with attention goes against all the available physiological evidence: Directing attention to the “preferred stimulus” in a crowded receptive field *does* restore the activation of the neurons (e.g., Desimone & Duncan, 1995; Kastner et al., 2001; Motter, 1993; Motter, 2002; Reynolds et al., 1999). It is possible that the concept of integration field according to Pelli et al. is not directly related to high-level neuronal receptive fields, but in this case, the underlying neuronal substrate would need to be described more specifically.

The following proposed distinction between two *different* types of spacing effects, although highly speculative, is an attempt to reconcile these various theories with the available data.

**Proposition 1:** Some familiar object categories can be processed directly by dedicated neuronal populations (or “feature integrators” with well-defined integration fields in the terminology of Pelli et al., 2004).

Selective activation of these neurons can take place in the absence of attention, when stimuli are well isolated. When local competition increases, however (i.e., when spacing decreases), activation is impaired—but it can be restored by focusing attention on the target stimulus. We end up with a serial search situation, in which performance can nevertheless be enhanced by increasing spacing. This is the *first type* of spacing effect that we have encountered, in [Experiment I](#), which is present only for the face search but not for the bisected disk search. Similar to the hypothesis of Pelli et al. (2004), this type of spacing limitation probably reveals the spatial resolution of the available integration fields (or, simply, receptive fields), but contrary to their view, in our proposal, this spacing

effect critically depends on attention being “split” across the various stimuli in the search array; it will disappear (and/or be replaced by another type of spacing effect, as described below) when attention is focused directly on the target.

**Proposition 2:** Other complex stimulus categories (e.g., arbitrary conjunctions of simple features) need attention to be effectively discriminated because no dedicated population of neurons responds to the relevant distinction (i.e., no feature integrator exists).

Attention thus serves the purpose of the feature integrator, as proposed originally by Treisman and Gelade (1980). A prerequisite for any such discrimination to take place is, therefore, that attention should be focused on the target stimulus. Under these conditions (which were the conditions of our Experiment II, wherein we used a precued search task), a *second type* of spacing effect can still be observed. This spacing effect likely reflects the spatial resolution of the attention focus, as proposed originally He et al. (1996) and Intriligator and Cavanagh (2001). This type of spacing effect will be observed equally for all “complex” discrimination tasks (i.e., discriminations not relying on the presence of a single feature) because discrimination performance in these tasks is mainly limited by attention, whether it be for “feature integration” or for resolving competition between simultaneously presented stimuli. This explains why both tasks (face discrimination and bisected disks discrimination) underwent similar spacing effects in Experiment II.

In this postulated distinction, the first type of spacing effect could easily be called “competition,” based on the biased competition framework of Desimone and Duncan (1995), whereas the second type of spacing effect encompasses what most researchers would generally refer to as crowding. In our view, crowding does have something to do with attention.

### What happens at spacings beyond 3°?

This study examined the effect on visual search performance of varying spacing up to 3°. In a pilot experiment, we investigated the effect of varying spacing between 3° and 12°; we found no conclusive evidence for reliable competitive effects in this case. The major clutter effect therefore seems to occur at smaller spacings, on the order investigated here (<3°). Does this mean, in keeping with our receptive field hypothesis, that the size of the relevant receptive fields is 3° or less? This would be difficult to reconcile with the known size of receptive fields in higher level areas of the macaque monkey, such as IT (>10°; Desimone, Moran, & Spitzer, 1988; Kastner et al., 2001). However, a simple simulation reveals that even when receptive fields are 10° across, most of the reduction in competition (due to increasing spacing) is expected to occur between 0° and 3° (Figure 8). Thus, the

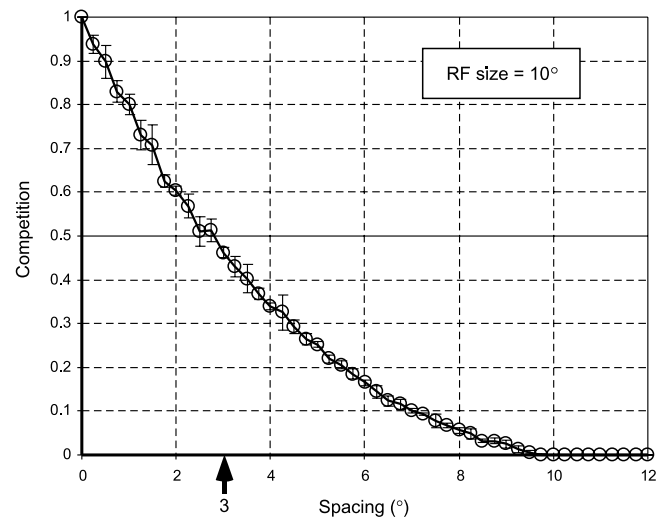


Figure 8. Competition as a function of spacing. We simulate a neuronal population whose idealized receptive field profiles are flat with a 10° diameter and whose centers are arranged, in a regular manner, on an infinite two-dimensional square grid of “tiling” distance  $t$  (i.e.,  $t$  is the distance between adjacent receptive field centers). Two stimuli, separated by a particular spacing, are placed somewhere in this 2D space. Let  $n_1$  be the number of neurons (in this 2D space) whose receptive fields hold only one stimulus. Let  $n_2$  be the number of neurons whose receptive fields hold both stimuli. We define competition as the ratio of the number of neurons in which the stimuli compete with each other ( $n_2$ ) to the total number of neurons that respond to these stimuli ( $n_1 + n_2$ ). That is, competition =  $n_2 / (n_1 + n_2)$ . Therefore, competition is 1 when every neuron responsive to either one of the stimuli responds to the other one as well (i.e., in effect,  $n_1 = 0$ ). Competition tends to zero as the number of neurons that respond unambiguously to one or the other stimulus is increased. The figure illustrates the change in competition with spacing for the given receptive field diameter (10°). The simulation was repeated with varying grid tiling distances (i.e.,  $t = 1^\circ, 2^\circ, 3^\circ, 4^\circ, \text{ or } 5^\circ$ , with 1,000 randomly placed stimulus pairs at each spacing). Error bars represent standard deviation across these five interneuronal distances. The parameter  $t$  does not seem to affect the amount of competition. Although this might seem counterintuitive, it is simply a consequence of the fact that (for values of  $t$  smaller than the receptive field radius) both  $n_1$  and  $n_2$  are inversely proportional to  $t$ . The key observation in these simulations is the concave shape of the competition versus spacing plot. Indeed, this implies that for the same increase in spacing (say, by 3°), a stronger reduction in competition is expected to occur at small spacings (e.g., between 0° and 3°) than at larger spacings (e.g., between 5° and 8°). In fact, most of the competition reduction occurs between 0° and 3°.

fact that the main effect of clutter is observed for spacing <3° in this experiment does not imply that the relevant receptive fields are smaller than 3°. Note that we are not directly measuring receptive field size but simply inferring it from our clutter manipulations.

## Can a serial visual search eventually become parallel?

In this study, we have found that for certain tasks, serial search might be a result of excessive competition (e.g., the face discrimination task). This leads us to ask: At large-enough spacings, can search eventually change from serial to parallel? Parallel search and serial search are generally defined in terms of the observed set size effects. However, in our experiment, not all set sizes exist at the largest spacings (in fact, at a spacing of  $3^\circ$ , only a set size of two is viable). Thus, in our protocol, it is not possible to determine whether serial search has, in fact, become parallel with increasing spacing; the significant increase in average search performance (Figure 5) only allows us to conclude that search is *more* parallel (or less serial) at large spacings compared with small spacings.

### Methodological considerations

Would the present result, obtained with masked displays, generalize to more classical, RT-based measures of search efficiency? This was not tested here, due in part to the costs involved in using self-terminating search paradigms: RTs in such tasks can be several times as long as in masked visual search, which would make the duration of data collection inappropriate for most experimental subjects. In addition, we believe that performance on masked search arrays better reflects the properties of the “feed-forward” visual system (Lamme & Roelfsema, 2000). By comparison, self-terminating search allows ample time for feedback interactions to occur, which might mask or decrease the RF effects postulated here.

It is also important to note that our circular search arrays do not necessarily optimize the amount of clutter. The spacing reported is the *minimal* distance between target and distractors; however, the *average* distance can be much larger. For example, consider a target-present trial with a spacing of  $1.5^\circ$  and a set size of four (one target, three distractors). This does not mean that the target stimulus is surrounded by three distractors, all  $1.5^\circ$  away from it. It simply means that there is *at least* one distractor  $1.5^\circ$  away from the target and *at most* two distractors  $1.5^\circ$  away from the target (as in Figure 1). The other distractors are positioned further away from the target so that the *average* distance to the target can be quite large. Thus, even at small spacings, clutter is not always optimized with respect to the reported spacing value, which means that there is little room for reducing it (and, consequently, improving search performance) at larger spacings. This is a potential limitation of the protocol used, which means that the spacing effects reported here might, in fact, be underestimated.

## Conclusion

In conclusion, we showed that competition (as indexed by interstimulus spacing) has a significant effect on serial search performance for a task that can rely on preexisting neuronal selectivities (face discrimination) but not for a task that depends on a more arbitrary combination of visual features (bisected disks discrimination). This may reveal different needs for attentional involvement in serial search, sometimes concerned with resolving local competition and sometimes serving a different purpose (binding/recognition?). Overall, it would seem that taking into account neuronal competition when interpreting visual search results might, in some cases, provide valuable insights into the role(s) of attention.

## Acknowledgments

This research was supported by grants from the NSF-sponsored Engineering Research Center at Caltech, the National Institutes of Health, the National Institutes of Mental Health, the Keck Foundation, and the McDonnell Foundation. We thank Heinrich Bülhoff for access to the face database and Christof Koch, Jeremy Wolfe, Patrick Cavanagh, and Leila Reddy for helpful comments on previous versions of the manuscript.

Commercial relationship: none.

Corresponding author: Rufin VanRullen.

Email: rufin@klab.caltech.edu.

Address: Centre de Recherche Cerveau et Cognition, CNRS, Université Paul Sabatier Toulouse 3, Faculté de Médecine Rangueil, 31062 Toulouse Cedex 9, France.

## References

- Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature*, *303*, 696–698. [PubMed]
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178. [PubMed]
- Brown, V., Huey, D., & Findlay, J. M. (1997). Face detection in peripheral vision: Do faces pop out? *Perception*, *26*, 1555–1570. [PubMed]
- Cohen, A., & Ivry, R. B. (1991). Density effects in conjunction search: Evidence for a coarse location mechanism of feature integration. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 891–901. [PubMed]



- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. [PubMed]
- Desimone, R., Moran, J., & Spitzer, H. (1988). Neural mechanisms of attention in extrastriate cortex of monkeys. In M. Arbib & S. Amari (Eds.), *Dynamic interactions in neural networks: Models and data* (pp. 169–182). New York: Springer-Verlag.
- Eckstein, M. (1998). The lower visual search efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological Science*, *9*, 111–118.
- Egeth, H. E. (1977). Attention and preattention. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 11, pp. 277–320). New York: Academic Press.
- Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology*, *79*, 1–16. [PubMed]
- Geisler, W. S., & Chou, K. L. (1995). Separation of low-level and high-level factors in complex tasks: visual search. *Psychological Review*, *102*, 356–378. [PubMed]
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology*, *35*, 96–111. [PubMed]
- He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*, 334–337. [PubMed]
- Hershler, O., & Hochstein, S. (2005). At first sight: A high-level pop out effect for faces. *Vision Research*, *45*, 1707–1724. [PubMed]
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, *43*, 171–216. [PubMed]
- Jeffreys, D. A. (1989). A face-responsive potential recorded from the human scalp. *Experimental Brain Research*, *78*, 193–202. [PubMed]
- Joseph, J. S., Chun, M. M., & Nakayama, K. (1997). Attentional requirements in a ‘preattentive’ feature search task. *Nature*, *387*, 805–807. [PubMed]
- Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, *68*, B1–B11. [PubMed]
- Kastner, S., De Weerd, P., Pinsk, M. A., Elizondo M. I., Desimone, R., & Ungerleider, L. G. (2001). Modulation of sensory suppression: Implications for receptive field sizes in the human visual cortex. *Journal of Neurophysiology*, *86*, 1398–411. [PubMed] [Article]
- Kreiman, G., Koch, C., & Fried, I. (2000). Category-specific visual responses of single neurons in the human medial temporal lobe. *Nature Neuroscience*, *3*, 946–953. [PubMed] [Article]
- Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*, 571–579. [PubMed]
- Li, F.-F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 9596–9601. [PubMed] [Article]
- McElree, B., & Carrasco, M. (1999). The temporal dynamics of visual search: Evidence for parallel processing in feature and conjunction searches. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1517–1539. [PubMed]
- Motter, B. C. (2002). Crowding and object integration within the receptive field of V4 neurons [Abstract]. *Journal of Vision*, *2*(7):274, 274a, <http://journalofvision.org/2/7/274/>, doi:10.1167/2.7.274.
- Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, *70*, 909–919. [PubMed]
- Nakayama, K., & Joseph, J. (1998). Attention, pattern recognition, and pop-out in visual search. In R. Parasuraman (Ed.), *The attentive brain* (pp. 279–298). Cambridge: MIT Press.
- Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*, 44–64.
- Nothdurft, H. C. (1993). Faces and facial expressions do not pop out. *Perception*, *22*, 1287–1298. [PubMed]
- Oram, M. W., & Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *Journal of Neurophysiology*, *68*, 70–84. [PubMed]
- 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. [PubMed]
- Pashler, H. (1987). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception & Psychophysics*, *41*, 191–201. [PubMed]
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*,



- 4(12):12, 1136–1169, <http://journalofvision.org/4/12/12/>, doi:10.1167/4.12.12. [PubMed] [Article]
- Perrett, D. I., Mistlin, A. J., Chitty, A. J., Smith, P. A., Potter, D. D., Broennimann, R., et al. (1988). Specialized face processing and hemispheric asymmetry in man and monkey: Evidence from single unit and reaction time studies. *Behavioural Brain Research*, 29, 245–258. [PubMed]
- Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, 47, 329–342. [PubMed]
- Purcell, D. G., Stewart, A. L., & Skov, R. B. (1996). It takes a confounded face to pop out of a crowd. *Perception*, 25, 1091–1108. [PubMed]
- Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, 435, 1102–1107. [PubMed]
- Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, 331, 163–166. [PubMed]
- Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *The Journal of Neuroscience*, 19, 1736–1753. [PubMed] [Article]
- Rousselet, G., Fabre-Thorpe, M., & Thorpe, S. (2002). Parallel processing in high-level categorization of natural images. *Nature Neuroscience*, 5, 629–630. [PubMed] [Article]
- Rousselet, G. A., Thorpe, S. J., & Fabre-Thorpe, M. (2004). Processing of one, two or four natural scenes in humans: The limits of parallelism. *Vision Research*, 44, 877–894. [PubMed]
- Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, 32, 1349–1357. [PubMed]
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136. [PubMed]
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95, 15–48. [PubMed]
- Troje, N., & Bühlhoff, H. (1996). Face-recognition under varying poses: The role of texture and shape. *Vision Research*, 36, 1761–1771. [PubMed]
- VanRullen, R. (2006). On second glance: Still no high-level pop-out effect for faces. *Vision Research*, 46, 3017–3027. [PubMed]
- VanRullen, R., Reddy, L., & Fei-Fei, L. (2005). Binding is a local problem for natural objects and scenes. *Vision Research*, 45, 3133–3144. [PubMed]
- VanRullen, R., Reddy, L., & Koch, C. (2004). Visual search and dual-tasks reveal two distinct attentional resources. *Journal of Cognitive Neuroscience*, 16, 4–14. [PubMed]
- Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, 31, 523–535. [PubMed] [Article]
- Verghese, P., & Nakayama, K. (1994). Stimulus discriminability in visual search. *Vision Research*, 34, 2453–2467. [PubMed]
- Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–73). London, UK: University College London Press.
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 419–433. [PubMed]
- Zelinsky, G. (1999). Precuing target location in a variable set size “nonsearch” task: Dissociating search-based and interference-based explanations for set size effects. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 875–903.