

# Effects of spatial separation between stimuli in whole report from brief visual displays

SØREN KYLLINGSBÆK AND CHRISTIAN VALLA  
University of Copenhagen, Copenhagen, Denmark

JAN VANRIE  
University of Leuven, Leuven, Belgium

AND

CLAUS BUNDESEN  
University of Copenhagen, Copenhagen, Denmark

Direct measurements of the effects of spatial separation between stimuli in whole report from brief visual displays are reported. The stimuli were presented on the periphery of an imaginary circle centered on fixation. In Experiment 1, each display showed two capital letters (letter height  $\cong 1.3^\circ$ , width  $\cong 0.9^\circ$ , eccentricity  $\cong 5.5^\circ$ ). The proportion of correctly reported letters was a strictly increasing, decelerating function of the spatial separation between the letters for center-to-center separations ranging from less than  $2^\circ$  to more than  $10^\circ$  of visual angle. Experiment 2 yielded similar results with triples of letters. Experiment 3 showed that accuracy increased with spatial separation for report of two short words, and Experiment 4 showed the same result for words presented upside down. The results are explained by a model of lateral masking (crowding) based on competitive interactions within receptive fields of cortical neurons.

The perceptibility of a visual stimulus tends to be reduced when the stimulus is flanked by other stimuli. The phenomenon is known as *lateral masking*, or the *crowding effect*. The effect is weak in central vision (Flom, Weymouth, & Kahneman, 1963) but strong in the periphery of the visual field. In a classical study, Bouma (1970; see also Bouma, 1973; Bouwhuis & Bouma, 1979) investigated the effects of spatial separation between a (randomly chosen) target letter to be reported and two flanking letters to be ignored. Recognition scores improved as the spatial separation between the target and the flankers was increased. The effect of spatial separation varied with retinal eccentricity. For complete visual isolation, a spatial separation of about 40%–50% of target eccentricity seemed to be required.

The causes of the crowding effect and its relationship to visual attention are controversial. Thus, Pelli, Palomares, and Majaj (2004) argued that crowding limits all vision and has nothing to do with attention, whereas Strasburger (2005) argued that the impairment of character recognition by crowding is caused, to a large extent, by spatially imprecise focusing of attention (see also He, Cavanagh, & Intriligator, 1996; Intriligator & Cavanagh, 2001; Tripathy & Cavanagh, 2002). For example, in the flanker paradigm used by Bouma (1970) and Strasburger, among others, the observer must discriminate the target (center character) from the distractors (flanker characters). Strasburger found a strong tendency, in erroneous reports of the identity of the

target character, for the reported identity to be the identity of one of the flanking characters. From the high frequency of such *localization errors*, Strasburger argued that unfocused spatial attention underlay the crowding effect.

In the flanker paradigm used by Bouma (1970) and Strasburger (2005), it seems likely that the difficulty of determining which character is the target (the center character) decreases as the spatial separation between the target and the flanker is increased. Thus, to an unknown extent, the effect of spatial separation observed in the flanker paradigm may reflect variation in target–distractor discriminability (the difficulty of determining which character is the target) caused by the experimental variation of spatial separation. As will be argued below, the same seems true in many other paradigms that have been used for investigating lateral masking.

The flanker paradigm of Bouma (1970) and Strasburger (2005) is a kind of partial-report paradigm. A different kind of partial-report paradigm is exemplified by a study of Bahcall and Kowler (1999). In this case, as well, the observed effect of spatial separation may reflect variation in target–distractor discriminability (the difficulty of determining which stimuli are targets) caused by the variation in spatial separation.

Bahcall and Kowler (1999) studied the effects of spatial separation between two targets in partial report from briefly exposed circular arrays of the two targets among 22 distractors. The two target locations were chosen randomly

on each trial and were specified by color or symbolic cues before the beginning of the trial. Bahcall and Kowler had predicted that performance would be best when the pair of targets were “near enough to one another to fall within the same attentional ‘window’” (p. 72). However, performance degraded as the target separation was decreased, except that (for 2 of the 3 participants) an upturn was seen at the smallest separation, where the targets were adjacent to each other (see Müller, Mollenhauer, Rösler, & Kleinschmidt, 2005, for related results consistent with a Mexican hat distribution of attentional resources).

The improvement found by Bahcall and Kowler (1999) when the target locations were adjacent to each other is reminiscent of a study by Podgorny and Shepard (1983) in which participants first viewed a  $3 \times 3$  grid with four or five of the cells shaded; then a probe dot appeared, and the participants indicated as quickly as possible whether or not the dot appeared in the shaded area. The results showed that performance was better when the shaded area was “spatially compact” (adjacent cells), rather than dispersed. This finding suggests that the improvement found at the smallest target separation in the study by Bahcall and Kowler can be explained by assuming that the required discrimination between target and distractor locations is easier when the target locations are adjacent so that they form a single compact area. By this explanation, the improvement at the smallest separation reflects a variation in target–distractor discriminability caused by the variation of spatial separation.

Effects of spatial separation have recently been investigated in many paradigms other than partial report. For example, Cutzu and Tsotsos (2003) and McCarley, Mounts, and Kramer (2004) reported positive effects of separation (improved performance with increasing spatial separation), using a *same–different* procedure. Mounts (2000) and Theeuwes, Kramer, and Kingstone (2004) found positive separation effects with an attentional capture paradigm. Kristjánsson and Nakayama (2002) found positive separation effects with an attentional blink paradigm, and Mounts and Gavett (2004) found positive separation effects with a divided attention procedure. In each of these studies, the displays generally contained the same number of stimuli, at the same spatial locations. The only thing that varied was which locations were attended. For example, in Experiment 1 of Cutzu and Tsotsos, the two stimulus characters to be attended (compared with respect to shape) were distinguished by being red, whereas the remaining stimuli (distractor characters) were black. Note that each of the red targets might be expected to stand out more strongly from the background of black distractors when the target was the only red item within an extensive neighborhood of uniformly colored, black items than when the target was presented near the other target. Thus, the main finding from the experiment—that accuracy increased with the spatial separation between the two targets—may wholly or in part reflect variation in target–distractor discriminability (the difficulty of distinguishing the two targets from the distractors) caused by the variation in spatial separation.

As another example, in the study of Theeuwes et al. (2004), the presence of an irrelevant color singleton reduced

the detectability ( $d'$ ) of a target contained in a shape singleton, particularly when the irrelevant color singleton was close to the target singleton. As was noted by the authors themselves, the closer the singleton distractor was to the singleton target, the less salient the singleton target might become, so again the observed effect of spatial separation may reflect variation in target–distractor discriminability (the difficulty of distinguishing the target singleton from the distractors) caused by the variation in spatial separation.

Similar arguments can be made with respect to the results of Kristjánsson and Nakayama (2002), McCarley et al. (2004), Mounts (2000, Experiment 1; but see Experiment 4), and Mounts and Gavett (2004). The paradigms require that targets are discriminated from distractors, and the reported positive effects of spatial separation may reflect variation in target–distractor discriminability (the difficulty of isolating the targets from the distractors) caused by the variation in spatial separation.

Effects of target–distractor discriminability can be eliminated by using displays without distractors (target-only displays). Two common paradigms suggest themselves: (1) *same–different* comparison between the targets and (2) whole report. Effects of spatial separation between two patterns to be compared with respect to shape have been investigated with both successively and simultaneously presented patterns. Both positive and negative effects of spatial separation have been observed, and the results appear to depend on the procedure used for comparing the two patterns. For example, observers may use a procedure of encoding one of the patterns in a pair as a mental image, shifting the image by a process of mental translation (translational displacement) to the location of the other pattern and then testing for a match. When such a procedure is used, the time taken by the comparison operation seems to increase with the spatial separation between the patterns to be compared (see Larsen & Bundesen, 1998, for pertinent findings and a general review).

Whole report may be the best paradigm for investigating the effects of the spatial separation between stimuli on their perceptibility, uncontaminated not only by variations in target–distractor discriminability (the difficulty of determining which stimuli are targets), but also by variations in the difficulty of comparing two stimuli caused by variations in their spatial separation. Surprisingly, although the spatial separation between the items displayed in a whole-report task is a major display parameter, few studies have addressed its effects. There are data showing that when strings of unrelated letters are presented for whole report, the end letters are reported more often than their neighbors and that when spaces are inserted into strings, performance on adjacent letters is improved (e.g., Wolford & Hollingsworth, 1974). However, parametric investigations of the effects of spatial separation between stimuli presented for whole report seem not to have been previously reported.

In this article, we describe direct measurements of the effects of spatial separation between stimuli in whole report of letters and words from briefly exposed displays. In each experiment, stimulus eccentricity was kept constant. In Experiment 1, we measured the effect of spatial separa-

tion in whole report from displays of two capital letters. Experiment 2 extended the investigation from pairs to triples of capital letters. In Experiment 3, we measured the effect of spatial separation in whole report from displays of two short words. Experiment 4 was similar to Experiment 3, but one or both of the stimulus words could be presented upside down (i.e., rotated 180° in the picture plane). After the presentation of the four experiments, we will propose a model of lateral masking (crowding) to account for the results of the experiments. The model is based on competitive interactions within receptive fields (RFs) of cortical neurons.

## EXPERIMENT 1 Report of Two Letters

In Experiment 1, we explored the effect of spatial separation in whole report from displays consisting of two capital letters. The letters were presented on the periphery of an imaginary circle centered on fixation. Exposures were brief and postmasked.

### Method

**Participants.** Eight students (3 males, 5 females) from the University of Copenhagen participated in the experiment. The participants served for payment (about 80 DKK, or \$10/h) and were naive with respect to the purpose of the experiment. All the participants had normal or corrected-to-normal visual acuity.

**Stimuli.** The stimulus material consisted of all the capital letters of the English alphabet except W. The letters were white (76.4 cd/m<sup>2</sup>) on a black (0.0 cd/m<sup>2</sup>) background and were printed in Arial font type. The mean height and width of a letter were 18 and 13 mm, respectively. At the viewing distance of 80 cm, the mean height of a letter subtended 1.3° of visual angle (range, 1.2°–1.3°), and the mean width of a letter subtended 0.89° of visual angle (range, 0.11°–1.4°). A pattern mask was formed by superimposing all of the possible stimulus letters upon each other, so that a pixel was included in the mask if, and only if, the pixel was included in one or more of the possible stimulus letters.

Each stimulus display showed a pair of letters. One of the letters was located (i.e., centered both vertically and horizontally) at the *point of origin* of the pair. There were 18 possible points of origin, which were evenly spaced on the perimeter of an imaginary circle centered at a small white fixation cross. Measured in the positive (counterclockwise) direction from the 3 o'clock position, the possible points of origin were found at 0°, 20°, . . . , and 340°, respectively. The other letter was located on the periphery of the same circle at a (counterclockwise) directional center-to-center separation of 20°, 40°, 60°, 120°, or 180° from the former letter. The diameter of the imaginary circle was 150 mm (11° of visual angle). The fixation cross was continuously visible at the center of the screen during each trial (see Figure 1).

**Procedure.** The experiment was run on a CRT controlled by an IBM-compatible PC. The participants were seated at a viewing distance of 80 cm from the screen in a semidarkened room. A chinrest was used to keep viewing distance constant during the whole experiment. Responses were recorded via a computer keyboard placed in front of the participants.

Before the actual experiment, a short practice session was run with a stimulus exposure duration of 112 msec. On the basis of performance during the practice session, the experiment was calibrated to the participant by setting the exposure duration of the stimuli at a particular value between 41 and 69 msec. This ensured that the performance of the participant was neither at ceiling nor at floor.

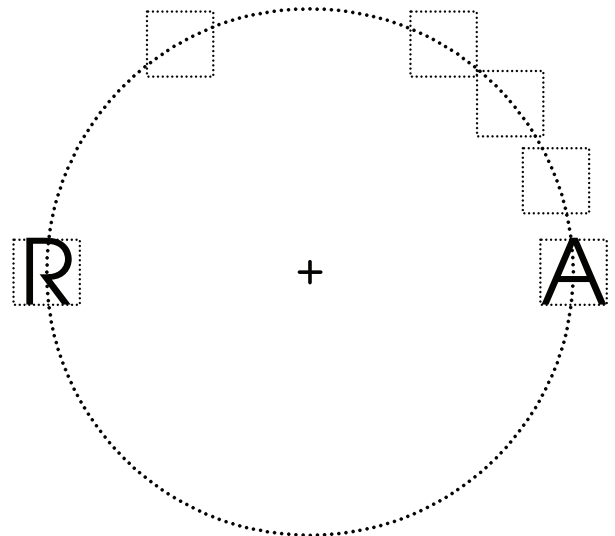
Each trial began by presentation of the fixation cross. When adequately fixated, the participant pressed a button to produce an immediate exposure of the two stimulus letters. The (41- to 69-msec) presentation of the two stimulus letters was immediately followed by an 800-msec exposure of two pattern masks, one at each of the two locations at which the letters had appeared.

After the masks had disappeared, two response windows (open squares) were presented, one at each of the locations at which a letter and a mask had appeared. The outline of one of the windows (chosen at random) was brighter than the outline of the other one. The participant first reported the letter that had appeared at the location of the brighter window by pressing the corresponding letter key on the keyboard, guessing whenever unsure (forced choice). The letter that was typed appeared in the response window. When the letter was typed, the response window became darker in outline, the other response window became brighter, and the participant reported the other letter by a similar procedure. The typing could be corrected by retyping, using the space bar to toggle between the two response windows. No time pressure was imposed. The participant terminated the trial by pressing the return key of the keyboard. Short (2- to 3-min) breaks were allowed between blocks.

**Design.** For each trial, the two stimulus letters were drawn at random, with replacement, from the set of 25 capital letters. The center-to-center distance between the two letters was systematically varied using the five different angles of directional separation (20°, 40°, 60°, 120°, and 180°), which corresponded to linear center-to-center distances of 1.9°, 3.7°, 5.3°, 9.3°, and 10.7° of visual angle.

Trials were presented in three different types of blocks: *near* blocks (N) using only directional separations of 20°, 40°, and 60°; *far* blocks (F) using only separations of 60°, 120°, and 180°; and *mixed* blocks (M) using all five directional separations.

In each block, each of the possible spatial separations was combined once with each of the possible points of origin, which yielded a total of 54 trials (3 separations × 18 points of origin) in each near or far block and 90 trials (5 separations × 18 points of origin) in each mixed block. The order of trials was random within blocks.



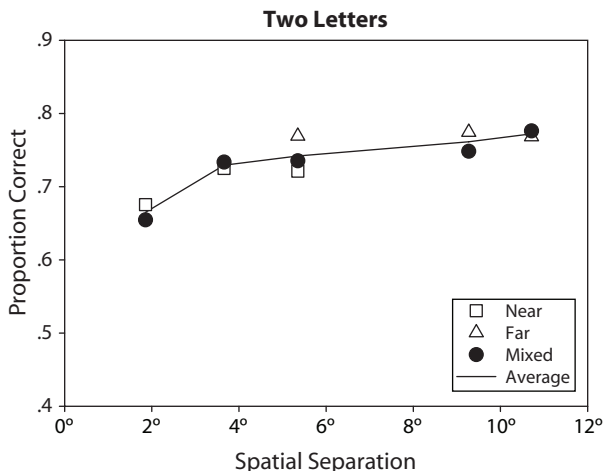
**Figure 1.** Sample stimulus display from Experiment 1. The dotted circle and rectangles were not visible in the display. The rectangles around letters R and A show the maximum vertical and horizontal extents of the letters at the indicated locations. The four empty rectangles show other locations where the R could have appeared given that (1) the A appeared at the 3 o'clock location and (2) the spatial separation was measured from A to R. Note that in the stimulus display, the letters were printed in white on a black background.

Each participant completed four blocks of each type. The order of the three different block types was counterbalanced across participants, using the following four sequences: NNNFFNNMMMM, FFNNNNFFMMMM, MMMMNFFNN, and MMMMFFNNNNFF. Each of the four sequences was used for 2 participants.

## RESULTS

The main results of Experiment 1 are shown in Figure 2. To test for possible effects of block type (near, far, or mixed), 2 two-way (block type  $\times$  spatial separation) repeated measures ANOVAs were done.<sup>1</sup> For the comparison between near and mixed blocks (for the three smaller spatial separations), the main effect of block type was not significant ( $F < 1$ ), nor was the interaction between block type and spatial separation significant ( $F < 1$ ). Similarly, for the comparison between far and mixed blocks (for the three larger spatial separations), the main effect of block type was not significant ( $F < 1$ ), nor was the interaction between block type and spatial separation significant [ $F(2,14) = 3.14, p = .08$ ]. Accordingly, the results from the three types of blocks were combined.

Averaged across the three types of blocks, the proportion of correct responses was a monotonically increasing, decelerating function of the spatial separation (visual angle) between the two stimulus letters (see Figure 2). By a one-way repeated measures ANOVA, the effect of spatial separation was highly significant [ $F(4,28) = 13.30, p < .001$ ]. The linear trend for the effect of spatial separation was also significant [ $F(1,7) = 33.13, p < .005$ ]. By Bonferroni-corrected pairwise comparisons, just two pairwise differences reached the .05 level of significance: the difference between the smallest and the largest spatial separation and the difference between the smallest and the second-largest spatial separation. However, the linear effect shown by the trend analysis remained significant when the data for the smallest spatial separation were excluded



**Figure 2.** Proportions of correctly reported letters as a function of the spatial separation (degrees of visual angle) between the stimuli in Experiment 1. Results from near, far, and mixed blocks are shown by squares, triangles, and circles, respectively. The overall means are indicated by unmarked points connected with straight lines.

[ $F(1,7) = 11.75, p < .05$ ] and, again, when both the data for the smallest and the data for the second-smallest separations were excluded [ $F(1,7) = 6.70, p < .05$ ].

## EXPERIMENT 2 Report of Three Letters

In Experiment 2, we explored the effect of spatial separation in whole report from brief displays consisting of three capital letters (a center letter and two flankers). The letters were presented on the periphery of an imaginary circle centered on fixation so that the flankers were located at the same distance but in opposite directions (clockwise vs. counterclockwise) from the center letter. The spatial separation associated with the triple was defined as the separation between the center letter and each of the flankers. The spatial separation was varied in smaller steps than in the previous experiment, and performance on the center letter was compared with performance on the flankers.

### Method

The method was the same as that used in Experiment 1, except when noted.

**Participants.** Six students (2 males, 4 females) from the University of Copenhagen participated in the experiment. One of the participants had previously participated in Experiment 1.

**Stimuli.** The stimulus material was the same as that in Experiment 1, and the stimuli were again presented at the perimeter of an imaginary circle, with a diameter of 150 mm (11°), centered on fixation. However, in Experiment 2, each stimulus display showed three letters instead of two.

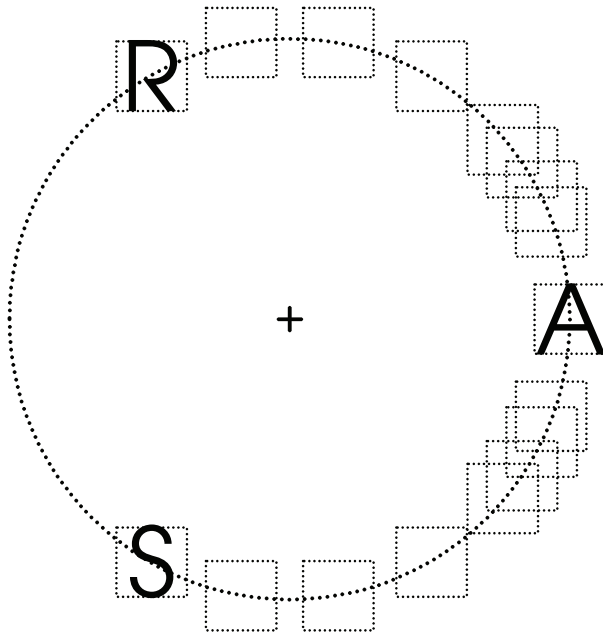
**Procedure.** The duration of stimulus exposure varied between 69 and 96 msec across participants. The presentation of the three stimulus letters was followed by a 700-msec exposure of three pattern masks, one at each of the three locations at which the letters had appeared. Following the pattern masks, three response windows (open squares) were presented in a row at the center of the screen. Contrary to the procedure used in Experiment 1, the participants were instructed to report the identities of the three letters without regard to their locations in the display. We adopted this method, instead of report by location as in Experiment 1, due to the increased difficulty of remembering three, rather than two, letters.

**Design.** For each trial, the 3 stimulus letters were drawn at random, without replacement, from the set of 25 capital letters. One member of the letter triple was designated the *center* letter and was located at one of the 18 possible points of origin on the perimeter of the imaginary circle centered on fixation. The 2 remaining letters (*flankers*) were located at the same distance but in opposite directions (clockwise vs. counterclockwise) from the center letter, and the directional separation associated with the triple was defined as the positive-valued directional separation between a center letter and a flanker. Eight directional center-to-center separations were used: 20°, 26°, 34°, 40°, 60°, 80°, 100°, and 120°, corresponding to linear center-to-center distances of 1.9°, 2.4°, 3.1°, 3.7°, 5.3°, 6.9°, 8.2°, and 9.3° of visual angle (see Figure 3).

As in Experiment 1, the eight different separations were factorially combined with the 18 possible points of origin, yielding 144 trials per block. Each participant participated in four experimental blocks.

### Results

The main results of Experiment 2 are shown in Figure 4. For each level of spatial separation between center letter and flankers, the proportion of correct responses was lower for the center letter than for the flankers; but



**Figure 3.** Sample stimulus display from Experiment 2. The dotted circle and rectangles were not visible in the display. The rectangles around letters R, A, and S show the maximum vertical and horizontal extents of letters at the indicated locations. The 14 empty rectangles show other locations where the letters R and S could have appeared, given that the center letter, A, appeared at the 3 o'clock location.

for both the center letter and the flankers, the proportion of correct responses increased with the spatial separation. A two-way relative position (central vs. flanking)  $\times$  spatial separation repeated measures ANOVA substantiated these conclusions by showing a significant main effect of relative position [ $F(1,5) = 64.93, p < .001$ ] and a significant main effect of spatial separation [ $F(6,30) = 11.93, p < .001$ ], in addition to a significant interaction between relative position and spatial separation [ $F(6,30) = 5.96, p < .001$ ]. The linear trend of the effect of spatial separation also reached significance [ $F(1,5) = 15.43, p < .05$ ]. Data for the largest spatial separation were excluded from the analysis because, in this condition, the spatial separation between the two flankers was no larger than the spatial separation between the center letter and the flankers (i.e., there was no real difference between center letter and flankers). By separate one-way ANOVAs of the data for the center letter and the flankers, respectively, the effect of spatial separation remained significant [for the center letter,  $F(6,30) = 11.03, p < .001$ ; for the flankers,  $F(6,30) = 4.53, p < .005$ ]. The linear trend in the effect of spatial separation also remained significant for the center letter [ $F(1,5) = 27.51, p < .005$ ], although not for the flankers [ $F(1,5) = 3.65, p = .11$ ].

### Discussion of Experiments 1 and 2

In Experiment 1, we measured the effect of spatial separation in whole report from brief displays consisting of two capital letters. The proportion of correct responses (the probability that a letter was correctly reported) was

found to be a strictly increasing, decelerating function of the spatial separation (visual angle) between the two letters. In Experiment 2, we investigated whole report from displays consisting of three capital letters (a center letter and two flankers). As was expected from previous studies of lateral masking in whole report (e.g., Wolford & Hollingsworth, 1974), the decrement in performance found at small spatial separations was stronger for the center letter than for the flankers. However, for both the center letter and the flankers, the proportions of correct responses appeared to be increasing, decelerating functions of spatial separation.

Experiments 3 and 4 were done to test the generality across stimulus materials of the effects of spatial separation found in Experiments 1 and 2. Visual recognition of words is more complex than recognition of single letters, and recognition of disoriented words is particularly difficult. Experiment 3 was done with words presented in their normal, upright position, and Experiment 4 with a mixture of upright words and words presented upside down (i.e., rotated 180° in the picture plane).

## EXPERIMENT 3 Report of Two Words

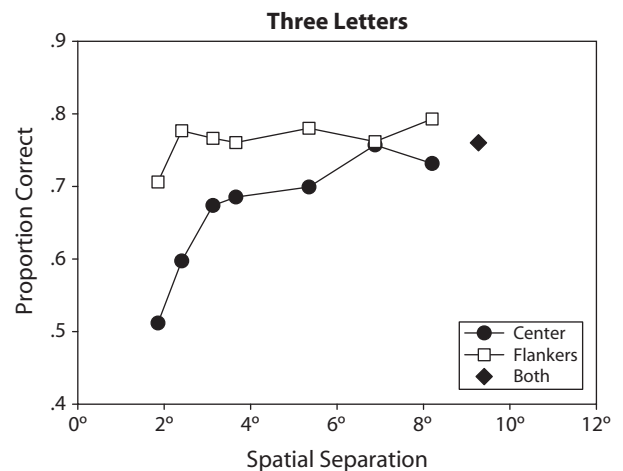
Experiment 3 was similar to Experiment 1, except that each stimulus display showed a pair of common three-letter words, instead of a pair of letters.

### Method

The method was the same as that used in Experiment 1, except when noted.

**Participants.** Seven students (2 males, 5 females) from the University of Copenhagen participated in the experiment. Two of the participants had previously participated in Experiment 2.

**Stimuli.** Each stimulus display showed two lowercase three-letter words presented in white on a black background. The two words



**Figure 4.** Proportions of correctly reported letters as a function of the spatial separation (degrees of visual angle) between the stimuli in Experiment 2. Results for the center letter and the flankers are shown by circles and squares, respectively. The average proportion correct at the largest spatial separation is shown by a diamond shape.

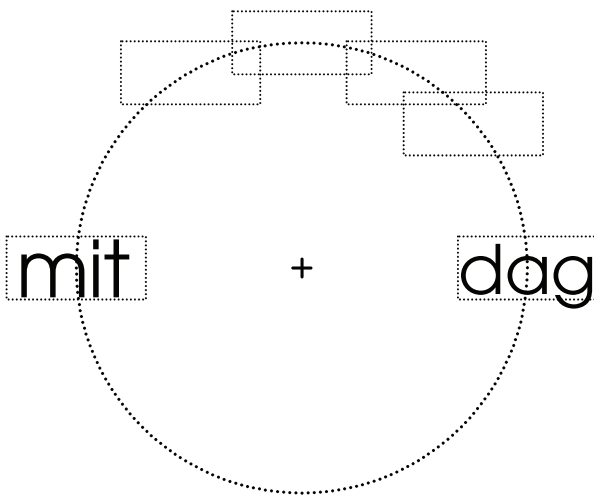
were drawn at random, without replacement, from a set of 25 common Danish words: *alt, bag, dag, dig, din, dog, får, før, god, gør, går, her, hos, liv, mit, mor, nok, nye, par, sad, ser, sit, tid, tog, and tre*. The mean height and width of a stimulus word were 18 mm (1.3° of visual angle; range, 0.92°–1.5°) and 36 mm (2.6°; range, 1.5°–3.3°), respectively. A pattern mask was constructed by superimposing all of the 25 stimulus words upon each other.

**Procedure.** The duration of stimulus exposure varied between 150 and 206 msec across participants. The presentation of the two stimulus words was followed by a 1,200-msec exposure of two pattern masks, one at each of the two locations at which the words had appeared. Following the pattern masks, two response windows (open rectangles) were presented, one at each of the locations at which a word and a mask had appeared. As in Experiment 1, the two words were reported by typing each word in the response box presented at the location at which that word had appeared. As before, the participants were instructed to guess whenever unsure (forced choice).

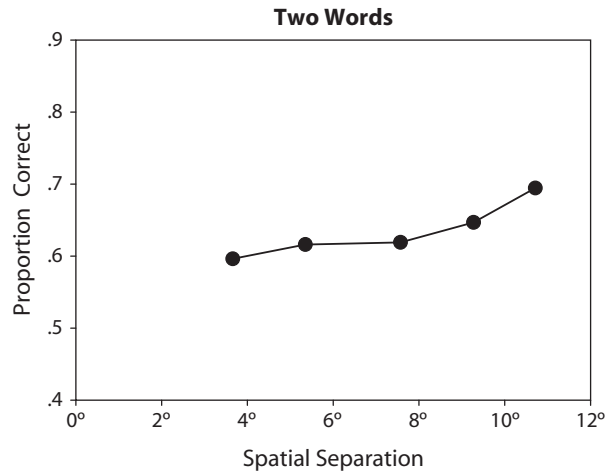
**Design.** In each stimulus pair, one of the stimulus words was located at 1 of the 18 possible points of origin on the perimeter of the imaginary circle centered on fixation. The other word was separated from the former by a directional center-to-center distance of 40°, 60°, 90°, 120°, or 180°, corresponding to a linear distance of 3.7°, 5.3°, 7.6°, 9.3°, or 10.7° of visual angle, respectively (see Figure 5). The five separations were factorially combined with the 18 points of origin, yielding a total of 90 trials per experimental block. Each participant participated in four blocks.

## Results

As in the prior experiments, the proportion of correct responses across the two stimulus words was an increasing function of the spatial separation (visual angle) between the two stimuli (see Figure 6; note that the smallest center-to-center distance represented in Figure 6 is comparable to the second-smallest center-to-center distance represented in Figure 2 and the fourth-smallest distance represented in Figure 4). By a one-way repeated measures ANOVA, the effect of spatial separation was significant [ $F(4,24) = 6.17, p < .005$ ]. The linear trend for the effect



**Figure 5.** Sample stimulus display from Experiment 3. The dotted circle and rectangles were not visible in the display. The rectangles around the stimulus words *mit* and *dag* show the maximum vertical and horizontal extents of words at the indicated locations. The four empty rectangles show other locations at which *mit* could have appeared, given that (1) *dag* appeared at the 3 o'clock location and (2) the spatial separation was measured from *dag* to *mit*.



**Figure 6.** Proportions of correctly reported words as a function of the spatial separation (degrees of visual angle) between the stimuli in Experiment 3.

of spatial separation also reached significance [ $F(1,6) = 10.14, p < .05$ ].

## EXPERIMENT 4 Report of Rotated Words

Experiment 4 was similar to Experiment 3, except that one or both of the stimulus words could be presented upside down.

### Method

The method was the same as that used in Experiment 3, except when noted.

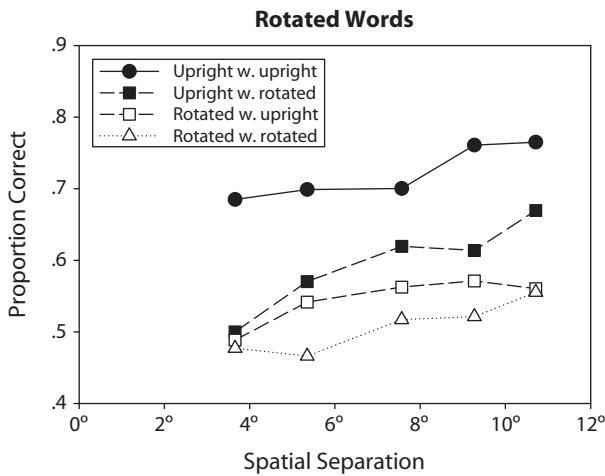
**Participants.** Five students (2 males, 3 females) from the University of Copenhagen participated in the experiment. All of the participants had previously participated in Experiment 3.

**Stimuli.** The stimulus displays were the same as those in Experiment 3, but one or both of the stimulus words could be presented upside down (i.e., rotated 180° in the picture plane).

**Design.** Each of the four experimental blocks run with each participant comprised a total of 270 trials, representing the factorial combinations of 18 points of origin  $\times$  5 spatial separations  $\times$  3 rotation conditions (viz., both words upright, one word rotated and the other upright, and both words rotated).

### Results

The results of Experiment 4 are shown in Figure 7. For each level of spatial separation between the stimulus words, the proportion of correct responses was (1) relatively high for an upright word that was accompanied by an upright word, (2) lower for an upright word accompanied by a rotated word, (3) still lower for a rotated word accompanied by an upright word, and (4) lowest for a rotated word accompanied by another rotated word. But for each of these four conditions, the proportion of correct responses increased with the spatial separation. A two-way (orientation condition  $\times$  spatial separation) repeated measures ANOVA substantiated these conclusions by showing both a significant main effect of orientation condition [ $F(3,12) = 18.47, p < .001$ ] and a significant main effect of spatial separation [ $F(4,16) = 6.04, p < .005$ ].



**Figure 7.** Proportions of correctly reported words as a function of the spatial separation (degrees of visual angle) between the stimuli in Experiment 4. The results are separately shown for upright words accompanied by upright words (circles), upright words accompanied by rotated words (filled squares), rotated words accompanied by upright words (open squares), and rotated words accompanied by rotated words (triangles).

The linear trend for the effect of spatial separation also was significant [ $F(1,4) = 11.88, p < .05$ ].

### Discussion of Experiments 3 and 4

Like Experiment 1 on whole report of letters, Experiments 3 and 4 showed that the wider the spatial separation between two words to be reported, the more accurate was the report. This was true whether the words were presented in their normal, upright position (Experiment 3) or one or both of the words were presented upside down (Experiment 4).

### A MODEL OF LATERAL MASKING (CROWDING)

We propose a model of lateral masking (crowding) to account for the results of our experiments. In addition to providing qualitative explanations for the observed effects of spatial separation in Experiments 1–4, the model provides quantitative predictions for the effects of spatial separation in whole report of small objects such as the stimulus letters in Experiments 1 and 2. The model builds on the theory of visual attention (TVA) proposed by Bundesen (1990) and the neural interpretation of TVA (NTVA) proposed by Bundesen, Habekost, and Kyllingsbæk (2005). It also incorporates ideas from the biased competition model of Desimone and Duncan (1995) and the CODE theory of visual attention (CTVA) developed by Logan (1996; see also Bundesen, 1998a, 1998b; Logan & Bundesen, 1996).

In the model we propose, stimuli compete for representation within the RFs of neurons in many cortical visual areas (cf. Desimone & Duncan, 1995; Moran & Desimone, 1985). Competition takes place when two or more stimuli are present in the RF of the same neuron, and it is a *winner-take-all* process. Thus, each neuron comes to

represent only one of the stimuli in its RF on any given trial (see Bundesen et al., 2005).

Lateral masking is generated by competition between stimuli for representation in cortical neurons with RFs that are limited in size. In cortical areas with neurons whose RFs individually cover the entire visual field, the competition between stimuli to become represented by the neurons are independent of the spatial separation between the stimuli. However, in cortical areas with neurons whose RFs are limited in size, stimuli close to each other will induce competition in a larger number of neurons than will stimuli that are far apart. The effect of the competition will be especially strong for stimuli with many close neighbors. Conversely, stimuli without close neighbors get little or no competition from other stimuli. Thus, in cortical areas with neurons whose RFs are limited in size, the number of neurons in which a stimulus becomes represented increases with increasing spatial separation of the stimulus from other stimuli (see Figure 8).

### Formulas

Consider the probability  $P$  that a stimulus  $x$  is correctly reported as an object of category  $i$  (e.g., a token of letter type  $i$ ) in whole report from a briefly exposed display of a set of stimuli,  $S$ . Let the number of stimuli in  $S$  be less than the capacity of visual short-term memory and assume that false identifications (misperceptions) of  $x$  are so rare that they can be neglected. In TVA, then, probability  $P$  is given by

$$P = 1 - \exp \left[ -\eta(x, i) \beta_i \frac{w_x}{\sum_{z \in S} w_z} \tau \right], \quad (1)$$

where  $\eta(x, i)$  is the strength of the sensory evidence that stimulus  $x$  belongs to category  $i$ ,  $\beta_i$  is a decision bias associated with category  $i$ ,  $w_x$  is the attentional weight of stimulus  $x$ , and  $\tau$  is the effective exposure duration.

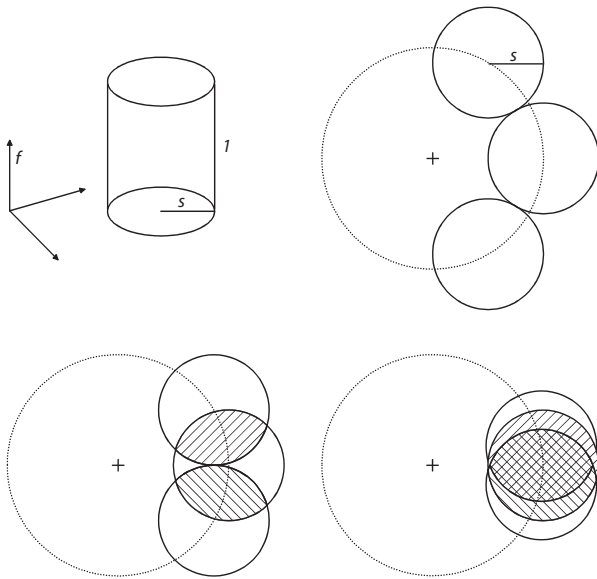
In NTVA, the value of the expression

$$\eta(x, i) \beta_i \frac{w_x}{\sum_{z \in S} w_z}$$

is proportional to the total activation of neurons representing the categorization that stimulus  $x$  belongs to category  $i$  at a level of processing at which the classical RFs of neurons are so large that each one covers the entire stimulus display. Now, to account for lateral masking, we assume that the rate of activation of each individual neuron representing that stimulus  $x$  belongs to category  $i$  is an increasing function of the number of those neurons in a certain lower level cortical area from which the given neuron receives critical information about  $x$ . Formally, we assume that

$$\eta(x, i) = \eta_{\max}(x, i) \frac{N_x^n}{a^n + N_x^n}, \quad (2)$$

where  $N_x$  is the number of neurons representing stimulus  $x$  in the given lower level cortical area,  $a > 0$ , and  $n > 0$ . [By Equation 2,  $\eta(x, i)$  is an increasing function of  $N_x$ , whose shape is determined by parameters  $a$  and  $n$ . As  $N_x$  increases from 0 toward infinity,  $\eta(x, i)$  increases from 0 toward a



**Figure 8.** Graphical illustrations of the model of lateral masking. The top-left panel of the figure shows a cylindrical receptive field (RF) function  $f$  with radius  $s$  and a height of 1. The other three panels are illustrations of experimental conditions with spatial separations between the stimuli that are large (top-right panel), medium (bottom-left panel), and small (bottom-right panel). The stimuli are depicted as points positioned on the periphery of an imaginary circle (dotted circle) centered at fixation. The circular area around a stimulus shows the set of possible locations of RF centers of neurons that have the given stimulus within their RF. Thus, the intersection between two overlapping circles—one for each of two stimuli—shows the set of possible locations of RF centers of neurons with both stimuli within their RFs, and the intersection between three overlapping circles—one for each of three stimuli—shows the set of possible locations of RF centers of neurons with all three stimuli within their RFs. Areas within which three stimuli compete for representation in each neuron are cross-hatched, areas where two stimuli compete for representation in each neuron are diagonally hatched, and areas where no competition occurs are white.

value of  $\eta_{\max}(x, i)$ . The function reaches a value of  $\eta_{\max}(x, i)/2$  at  $N_x = a$ . For  $n > 1$ , the function is sigmoid.]

Because neurons in the lower level area have RFs that are limited in size,  $N_x$  depends on both the shape of the RFs and the spatial configuration of the stimuli in the display. The RFs of neurons in a given cortical area are assumed to be circular with the same radius  $s$  and with centers distributed across the visual field with a certain density  $\rho$ .<sup>2</sup> The same RFs can be represented by cylindrical indicator functions with radius  $s$  and a height of 1 (see Figure 8). In the Appendix, we derive a general formula (Equation A1) for the expected value of  $N_x$  for a stimulus  $x$  in a display with a certain configuration. The formula presupposes that  $x$  is a small object (e.g., an individual letter), so that, to a reasonable approximation,  $x$  can be regarded as being inside the RF of a given neuron if, and only if, the centroid of  $x$  is located inside the RF.

**Define  $N'_x = N_x \rho$  and  $a' = a \rho$ .** For a display with a given configuration of stimulus objects, Equation A1 gives  $N'_x$  as a function of (1) a set of indicator functions ( $f_x, f_y, f_z$ ) corresponding to RFs and (2) the attentional weights ( $w_x, w_u$ ) of the stimuli. Given that the indicator

functions are cylindrical with the same radius  $s$ , and assuming that attentional weights are the same for all objects in the set of stimuli  $S$ , the equation gives  $N'_x$  as a function of  $s$ . Equation 2 can be rewritten as

$$\eta(x, i) = \eta_{\max}(x, i) \frac{(N'_x)^n}{(a')^n + (N'_x)^n},$$

so jointly, Equations 2 and A1 yield  $\eta(x, i)$  as a function of  $\eta_{\max}(x, i)$ ,  $a'$ ,  $n$ , and  $s$ . Given that attentional weights are the same for all objects in  $S$ , the ratio

$$\frac{w_x}{\sum_{z \in S} w_z}$$

on the right-hand side of Equation 1 equals 1/2 for the displays used in Experiment 1 and 1/3 for the displays used in Experiment 2. Accordingly, for either experiment, Equations 1, 2, and A1 give the probability  $P$  that a stimulus letter  $x$  is correctly reported as a function of curve-shape parameters  $a'$  and  $n$ , RF radius  $s$ , and a parameter given by the product  $\eta_{\max}(x, i) \beta_i \tau$ .

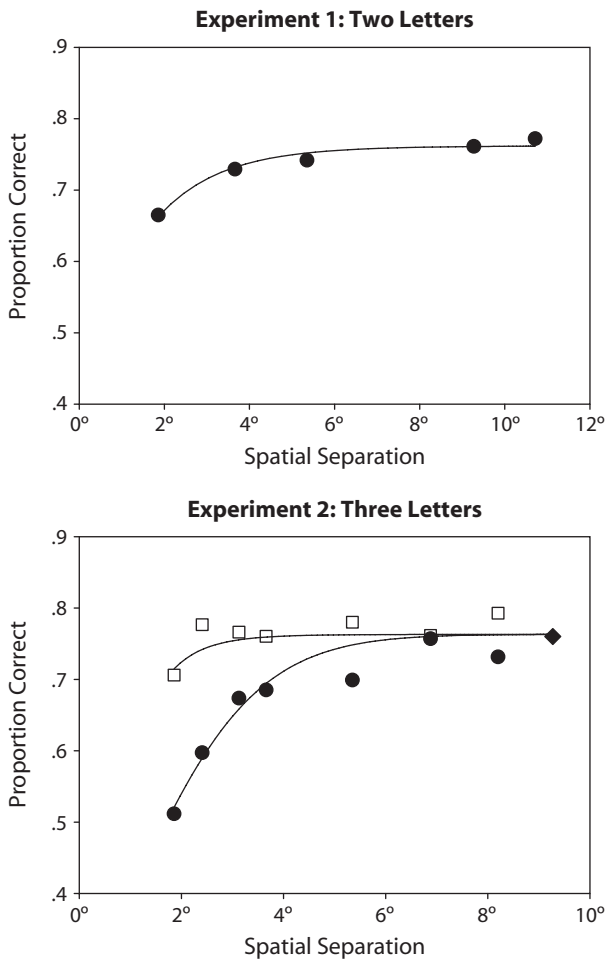
**Fits**

Figure 9 shows fits of the model to the group data from Experiments 1 and 2. Although the groups of participants differed between the two experiments, both RF radius  $s$  and curve-shape parameter  $n$  were constrained to be the same. The fits were obtained with radius  $s$  at a value of 29° of visual angle, parameter  $n$  at 18.5, parameter  $a'$  at a value of 493 in Experiment 1 and 699 in Experiment 2, and the product  $\eta_{\max}(x, i) \beta_i \tau$  at a value of 2.87 in Experiment 1 and 4.32 in Experiment 2. As can be seen, the model predicts both the general pattern of results in Experiment 1 and the dramatic drop in performance for the center, as compared with the flanking letters at small spatial separations, in Experiment 2. The estimated value of the RF radius  $s$  is well within the range of RF sizes reported for neurons in the inferotemporal cortical visual area of macaque monkeys (see, e.g., Desimone & Gross, 1979; Desimone & Ungerleider, 1989; Kastner et al., 2001). This area contains neurons that appear to respond preferentially to visual forms of approximately the same complexity as individual letters (see, e.g., Tanaka, Saito, Fukada, & Moriya, 1991).

**GENERAL DISCUSSION**

**Achievements**

The way the perceptibility of a visual stimulus is reduced when the stimulus is flanked by other stimuli (lateral masking, crowding) has previously been investigated with a number of techniques, including partial report and comparison between simultaneously presented stimuli. As was pointed out in the introduction, the effects of spatial separation between stimuli shown by these techniques may reflect not only variations in perceptibility of the target stimuli, but also correlated variations in such factors as the difficulty in determining which stimuli are targets in partial report and the difficulty of the comparison operation implicated in stimulus comparison tasks. To eliminate



**Figure 9.** Fits by the model of lateral masking to the results of Experiments 1 and 2. The observed data (circles and squares) are replotted from Figures 2 and 4. The predictions of the model are shown by smooth curves (solid lines).

both the effects of variations in target–distractor discriminability and the effects of variations in the difficulty of comparison operations, we measured the effect of spatial separation between stimuli in a whole-report paradigm. Although whole report is a major paradigm in the study of visual cognition (see, e.g., Bundesen, 1996; Shibuya & Bundesen, 1988; Sperling, 1960, 1963, 1967; Townsend & Ashby, 1983, chap. 11; see also Cattell, 1885; Erdmann & Dodge, 1898), parametric investigations of the effect of spatial separation between stimuli presented for whole report seem not to have been previously reported.

In this article, we have presented four experiments providing direct measurements of the effects of spatial separation between stimuli in whole report of letters and words from briefly exposed displays. In each experiment, stimulus eccentricity was kept constant (at about 5°–6°). In all four experiments and in every experimental condition, accuracy was found to be an increasing function of spatial separation. The increase was weak for items separated by more than a few degrees, but the lateral interference was long-ranging.

In Experiment 1, the displays contained just two capital letters. In degrees of visual angle, the mean height of a letter was 1.3°, and the mean width was 0.9°. The proportion of correctly reported letters was found to be a strictly increasing, decelerating function of the spatial separation (visual angle) between the stimulus letters for center-to-center separations ranging from less than 2° to more than 10°.

In Experiment 2, the displays contained three capital letters (a center letter and two flankers). The decrement in performance found at small spatial separations was stronger for the center letter than for the flankers (cf. Wolford & Hollingsworth, 1974). However, for both the center letter and the flankers, the proportions of correct responses again appeared to be strictly increasing, decelerating functions of spatial separation.

Experiments 3 and 4 tested the generality of our findings with other types of stimulus materials. Experiment 3 was similar to Experiment 1, except that each stimulus display showed a pair of common, lowercase three-letter words, instead of a pair of capital letters. The mean height of a word was 1.3°, and the mean width was 2.6°. As in the prior experiments, accuracy appeared to be a strictly increasing function of the spatial separation (visual angle) between the stimulus words as the center-to-center separation was varied from 3.7° to more than 10°.

Experiment 4 was similar to Experiment 3, except that one or both of the stimulus words could be presented upside down (i.e., rotated 180° in the picture plane). Accuracy was high for an upright word accompanied by an upright word, lower for an upright word accompanied by a rotated word, still lower for a rotated word accompanied by an upright word, and lowest for a rotated word accompanied by a rotated word. However, in each orientation condition, accuracy appeared to be a strictly increasing function of the spatial separation. Thus, for each type of stimulus material we tested, accuracy of whole report increased monotonically with the spatial separation between the items to be reported.

We proposed a model of lateral masking (crowding) based on NTVA (Bundesen et al., 2005) to account for the results of our experiments. In this model, lateral masking is generated by competition between stimuli for representation in neurons with relatively small RFs. Whole-report performance is based on processing in high-level neurons whose RFs individually cover the entire visual field. To explain how lateral masking affects whole-report performance, the model assumes that the rate of activation of a high-level neuron representing that stimulus  $x$  belongs to category  $i$  is an increasing, S-shaped function of the number of those neurons in a certain lower level cortical area from which the high-level neuron receives critical information about  $x$ . Because neurons in the lower level area have RFs that are limited in size, the number of lower level neurons representing stimulus  $x$  increases with the spatial separation of stimulus  $x$  from other, competing stimuli.

In addition to providing qualitative explanations for the observed effects of spatial separation in Experiments 1–4, the model provided quantitative predictions for the effects of spatial separation in whole report of small objects, such as the stimulus letters in Experiments 1 and 2 (objects so

small that, to a reasonable approximation, they could be regarded as being inside the RF of a given neuron if, and only if, their centroids were located inside the RF). With four free parameters, the model predicted both the general pattern of results in Experiment 1 and the dramatic drop in performance for the center, as compared with the flanking letters at small spatial separations, in Experiment 2.

In all four experiments, performance was limited by the duration of stimulus exposure (i.e., *resource limited*, in the sense of Norman & Bobrow, 1975). Informal experimentation showed that with long exposure durations, performance was essentially perfect. Our model is consistent with this observation. By Equation 1, the probability  $P$  that a stimulus  $x$  is correctly reported as an object of category  $i$  (e.g., a token of letter type  $i$ ) tends toward a value of 1 as the effective exposure duration,  $\tau$ , tends toward infinity. Since lateral masking (crowding) is explained by sharing of resources, the model is attentional. At the same time, as Pelli et al. (2004) would have it, the spatial extent of lateral masking (the *critical spacing* of crowding) is “a preattentive limitation, unaffected by attentional manipulations” (p. 1160).

### Limitations

Our present model assumes no limitations with respect to the spatial resolution of attentional weighting (cf. Intriligator & Cavanagh, 2001) or feature integration (cf. Pelli et al., 2004). Presumably, however, there are both (1) limitations on how fine-grained in space attentional weighting can be (e.g., limitations on the spatial resolution of a map of attentional weights; Bundesen et al., 2005) and (2) lower bounds on the spatial extent of those areas (*integration fields*; Pelli et al., 2004) over which features are integrated (e.g., by being embedded in feedback loops gated by object units in a visual short-term memory map; Bundesen et al., 2005). Such limitations may explain cases in which lateral masking (crowding) is observed even when viewing time is unlimited (*data-limited* performance, in the sense of Norman & Bobrow, 1975; see Townsend, Taylor, & Brown, 1971, for an early demonstration of lateral masking for letters with unlimited viewing time; see Pelli et al., 2004, for a recent review).

### AUTHOR NOTE

This work was supported by grants from the Carlsberg Foundation and the Danish Council for Strategic Research to the first author, an Erasmus/Socrates student travel grant from the European Union to J.V., and a grant from the Danish Research Council for the Humanities to C.B. Correspondence concerning this article should be addressed to S. Kyllingsbæk, Center for Visual Cognition, Department of Psychology, University of Copenhagen, Linnésgade 22, DK-1361 Copenhagen K, Denmark (e-mail: sk@psy.ku.dk).

### REFERENCES

- BAHCALL, D. O., & KOWLER, E. (1999). Attentional interference at small spatial separations. *Vision Research*, **39**, 71-86.
- BOUMA, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, **226**, 177-178.
- BOUMA, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, **13**, 767-782.
- BOUWHUIS, D., & BOUMA, H. (1979). Visual word recognition of three-letter words as derived from the recognition of the constituent letters. *Perception & Psychophysics*, **25**, 12-22.
- BUNDESEN, C. (1990). A theory of visual attention. *Psychological Review*, **97**, 523-547.
- BUNDESEN, C. (1996). Formal models of visual attention: A tutorial review. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 1-43). Washington, DC: American Psychological Association.
- BUNDESEN, C. (1998a). A computational theory of visual attention. *Philosophical Transactions of the Royal Society of London: Series B*, **353**, 1271-1281.
- BUNDESEN, C. (1998b). Visual selective attention: Outlines of a choice model, a race model and a computational theory. *Visual Cognition*, **5**, 287-309.
- BUNDESEN, C., HABEKOST, T., & KYLLINGSBÆK, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, **112**, 291-328.
- CATTELL, J. M. (1885). Über die Zeit der Erkennung und Benennung Schriftzeichen, Bildern und Farben. *Philosophische Studien*, **2**, 635-650.
- CUTZU, F., & TSOTSOS, J. K. (2003). The selective tuning model of attention: Psychophysical evidence for a suppressive annulus around an attended item. *Vision Research*, **43**, 205-219.
- DESIMONE, R., & DUNCAN, J. (1995). Neural mechanisms of selective visual attention. *Annual Review in Neuroscience*, **18**, 193-222.
- DESIMONE, R., & GROSS, C. G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Research*, **178**, 363-380.
- DESIMONE, R., & UNGERLEIDER, L. G. (1989). Neural mechanisms of visual perception in monkeys. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 2, pp. 267-299). Amsterdam: Elsevier.
- ERDMANN, B., & DODGE, R. (1898). *Psychologische Untersuchungen über das Lesen*. Halle, Germany: Niemeyer.
- FLOM, M. C., WEYMOUTH, F. W., & KAHNEMAN, D. (1963). Visual resolution and contour interaction. *Journal of the Optical Society of America*, **53**, 1026-1032.
- HE, S., CAVANAGH, P., & INTRILIGATOR, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, **383**, 334-337.
- INTRILIGATOR, J., & CAVANAGH, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, **43**, 171-216.
- KASTNER, S., DE WEERD, P., PINSK, M. A., ELIZONDO, M. I., DESIMONE, R., & UNGERLEIDER, G. (2001). Modulation in sensory suppression: Implications for receptive field sizes in the human visual cortex. *Journal of Neurophysiology*, **86**, 1398-1411.
- KRISTJÁNSSON, A., & NAKAYAMA, K. (2002). The attentional blink in space and time. *Vision Research*, **42**, 2039-2050.
- LARSEN, A., & BUNDESEN, C. (1998). Effects of spatial separation in visual pattern matching: Evidence on the role of mental translation. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 719-731.
- LOGAN, G. D. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, **103**, 603-649.
- LOGAN, G. D., & BUNDESEN, C. (1996). Spatial effects in the partial report paradigm: A challenge for theories of visual spatial attention. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 35, pp. 243-282). San Diego: Academic Press.
- MCCARLEY, J. S., MOUNTS, J. R. W., & KRAMER, A. F. (2004). Age-related differences in localized attentional interference. *Psychology & Aging*, **19**, 203-210.
- MORAN, J., & DESIMONE, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, **229**, 782-784.
- MOUNTS, J. R. W. (2000). Evidence for suppressive mechanisms in attentional selection: Feature singletons produce inhibitory surrounds. *Perception & Psychophysics*, **62**, 969-983.
- MOUNTS, J. R. W., & GAVETT, B. E. (2004). The role of salience in localized attentional interference. *Vision Research*, **44**, 1575-1588.
- MÜLLER, N. G., MOLLENHAUER, M., RÖSLER, A., & KLEINSCHMIDT, A. (2005). The attentional field has a Mexican hat distribution. *Vision Research*, **45**, 1129-1137.
- NORMAN, D. A., & BOBROW, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, **7**, 44-64.
- PELLI, D. G., PALOMARES, M., & MAJAJ, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, **4**, 1136-1169.
- PODGORNY, P., & SHEPARD, R. N. (1983). Distribution of visual attention

- over space. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 380-393.
- SHIBUYA, H., & BUNDESEN, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 591-600.
- SPELTING, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, **74**(11, Whole No. 498).
- SPELTING, G. (1963). A model for visual memory task. *Human Factors*, **5**, 19-31.
- SPELTING, G. (1967). Successive approximations to a model for short-term memory. *Acta Psychologica*, **27**, 285-292.
- STRASBURGER, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision*, **5**, 1024-1037.
- TANAKA, K., SAITO, H., FUKADA, Y., & MORIYA, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology*, **66**, 170-189.
- THEEUWES, J., KRAMER, A. F., & KINGSTONE, A. (2004). Attentional capture modulates perceptual sensitivity. *Psychonomic Bulletin & Review*, **11**, 551-554.
- TOET, A., & LEVI, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, **32**, 1349-1357.
- TOWNSEND, J. T., & ASHBY, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.
- TOWNSEND, J. T., TAYLOR, S. G., & BROWN, D. R. (1971). Lateral masking for letters with unlimited viewing time. *Perception & Psychophysics*, **10**, 375-378.
- TRIPATHY, S. P., & CAVANAGH, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Research*, **42**, 2357-2369.
- WINER, B. J., BROWN, D. R., & MICHELS, K. M. (1991). *Statistical principles in experimental design*. New York: McGraw-Hill.
- WOLFORD, G., & HOLLINGSWORTH, S. (1974). Lateral masking in visual information processing. *Perception & Psychophysics*, **16**, 315-320.

## NOTES

1. All accuracy data (proportions of correct responses) were arcsin transformed to stabilize the variances before the ANOVAs were performed (see Winer, Brown, & Michels, 1991, pp. 354-358).
2. These assumptions are only rough approximations. For instance, peripheral vision shows strong anisotropies. Typical RFs of cells in the striate cortex are elongated and have a radial orientation bias. In line with this, Toet and Levi (1992) found that the spatial extent of lateral masking was two to three times larger for items arrayed radially (relative to fixation), as compared with items arrayed with the same eccentricity.

## APPENDIX Computation of $N_x$

Here, we derive a general formula for the expected number of neurons representing a small object  $x$  (e.g., an individual letter) in a given cortical visual area. Consider the array of RF centers of all neurons in the given cortical area. This array forms a distribution of points across the 2-D visual field. To a first approximation, the density of the points (i.e., the number of points per unit solid angle) is assumed to be constant across the visual field. For specificity, we assume that the points are distributed in accordance with a 2-D Poisson process with density  $\rho$ . This assumption implies that (1) the numbers of points in nonoverlapping regions are independent random variables and (2) the number of points in a region with area  $A$  is Poisson distributed with a mean of  $\rho A$ .

The expected number of neurons with RF centers within a small piece of area  $dA$  equals  $\rho dA$ . The probability that a neuron whose RF contains object  $x$  represents object  $x$ , rather than representing any other objects in its RF, equals

$$\frac{w_x}{\sum_{u \in J \cup \{x\}} w_u},$$

where  $J$  is the set of objects other than  $x$  in the RF (see Bundesen et al., 2005). The expected number of neurons that have RF centers in  $dA$  and represent object  $x$  is given by the product

$$\sum_{u \in J \cup \{x\}} \frac{w_x}{w_u} \rho dA.$$

The expected number of neurons representing object  $x$ ,  $N_x$ , is found by integrating this product across the relevant part of the visual field, keeping track of the way in which  $J$  varies across the field. The details are as follows.

For any small object  $x$  and any point  $Q$  in the visual field, let the indicator function  $f_x$  take on a value of 1 at the point  $Q$  if object  $x$  lies within the RF of a neuron with RF center at  $Q$ , and let  $f_x = 0$  at  $Q$  if object  $x$  lies outside the RF of a neuron whose RF is centered at  $Q$ . The integration shall be done across that part of the visual field in which  $f_x = 1$ . The value of  $J$  to be used at a particular location is that value for which

$$\prod_{y \in J} f_y \prod_{z \in S - J - \{x\}} (1 - f_z) = 1,$$

where  $S$  is the set of all stimuli in the display. The possible values of  $J$  are all possible subsets of  $S - \{x\}$  (i.e., the elements in the power set of  $S - \{x\}$ ). Letting  $P(S - \{x\})$  stand for the power set of  $S - \{x\}$ , we get the general formula

$$N_x = \iint \sum_{f_x=1} \sum_{J \in P(S - \{x\})} \prod_{y \in J} f_y \prod_{z \in S - J - \{x\}} (1 - f_z) \frac{w_x}{\sum_{u \in J \cup \{x\}} w_u} \rho dA \quad (\text{A1})$$

for the expected number of neurons representing object  $x$  in the given cortical visual area.