

# How color, regularity, and good Gestalt determine backward masking

**Bilge Sayim**

Laboratory of Psychophysics, Brain Mind Institute,  
Ecole Polytechnique Fédérale de Lausanne (EPFL),  
Lausanne, Switzerland  
Laboratory of Experimental Psychology,  
University of Leuven, Leuven, Belgium



**Mauro Manassi**

Laboratory of Psychophysics, Brain Mind Institute,  
Ecole Polytechnique Fédérale de Lausanne (EPFL),  
Lausanne, Switzerland



**Michael Herzog**

Laboratory of Psychophysics, Brain Mind Institute,  
Ecole Polytechnique Fédérale de Lausanne (EPFL),  
Lausanne, Switzerland



The strength of visual backward masking depends on the stimulus onset asynchrony (SOA) between target and mask. Recently, it was shown that the conjoint spatial layout of target and mask is as crucial as SOA. Particularly, masking strength depends on whether target and mask group with each other. The same is true in crowding where the global spatial layout of the flankers and target-flanker grouping determine crowding strength. Here, we presented a vernier target followed by different flanker configurations at varying SOAs. Similar to crowding, masking of a red vernier target was strongly reduced for arrays of 10 green compared with 10 red flanking lines. Unlike crowding, single green lines flanking the red vernier showed strong masking. Irregularly arranged flanking lines yielded stronger masking than did regularly arranged lines, again similar to crowding. While cuboid flankers reduced crowding compared with single lines, this was not the case in masking. We propose that, first, masking is reduced when the flankers are part of a larger spatial structure. Second, spatial factors counteract color differences between the target and the flankers. Third, complex Gestalts, such as cuboids, seem to need longer processing times to show ungrouping effects as observed in crowding. Strong parallels between masking and crowding suggest similar underlying mechanism; however, temporal factors in masking additionally modulate performance, acting as an additional grouping cue.

## Introduction

Visual masking is a standard tool in visual psychophysics. In masking, a target is preceded, followed, or presented simultaneously with a mask that renders the target harder or impossible to perceive. Simultaneous presentation of a target and a nonoverlapping mask (flankers) is often referred to as crowding, in particular when target and mask are presented in the visual periphery (for recent reviews, Levi, 2008; Whitney & Levi, 2011). Bachmann (1994) estimated that 14% of all articles in vision science and psychology use masking. Enns and Di Lollo (2000) came to a similar estimate. Masking is often used as a tool to limit the processing time of a target stimulus, for example, to investigate unconscious processing and priming or to investigate the characteristics of masking itself to shed light on the dynamics of the visual brain. Even though masking is a frequently used tool in many fields of visual perception, its mechanisms are far from well understood.

In visual backward masking, a target is followed by a mask at varying interstimulus intervals (ISI). The mask impairs perception of the target depending on the ISI. There are two main types of masking, A-type and B-type masking (Kolars, 1962). In A-type masking, strongest masking is obtained when target and mask are presented simultaneously. Performance improves monotonically when the ISI increases. A-

Citation: Sayim, B., Manassi, M., & Herzog, M. (2014). How color, regularity, and good Gestalt determine backward masking. *Journal of Vision*, 14(7):8, 1–11, <http://www.journalofvision.org/content/14/7/8>, doi:10.1167/14.7.8.

doi: 10.1167/14.7.8

Received October 30, 2013; published June 18, 2014

ISSN 1534-7362 © 2014 ARVO

type masking is usually explained by temporal integration of target and mask.

In B-type masking, performance deficits are strongest for intermediate ISIs, often around 50 ms. Early theories of B-type masking proposed that the distance between the inner contour of the mask and the outer contour of the target determines B-type masking (Werner, 1935). This approach was replaced by the dual channel model by Breitmeyer and Ganz (1976), which served as a standard explanation for two decades. In this model, each stimulus—target and mask—is processed in a fast transient channel, often attributed to the magnocellular system (M-system), and a slower sustained channel, often attributed to the parvocellular system (P-system). B-type masking occurs when the fast magnocellular mask signals inhibit the slower parvocellular target signals. Evidence for this proposal comes from studies showing that feature discrimination but not detection is influenced by the mask (Fehrer & Raab, 1962), in accordance with the idea that detection of an element is mainly due to the M-system whereas its feature processing is mainly due to the P-system. While there is indeed convincing evidence for the dual channel approach (see monograph by Breitmeyer & Öğmen, 2006), other studies have shown limitations, and alternative explanations were proposed (e.g., Enns & Di Lollo, 2000; Francis, 1997; Hermens, Luksys, Gerstner, Herzog, & Ernst, 2008; Herzog & Fahle, 2002; Herzog & Koch, 2001). For example, Bachmann (1984) proposed a dual channel retouch theory based on an interaction of cortical and thalamic processing. Francis (1997) showed mathematically that B-type masking should primarily occur when masks are “weaker” than the target, which may well occur within one channel (but see Francis & Herzog, 2004). Enns and Di Lollo (1997) proposed that the mask substitutes the target during recurrent processing.

In all of these approaches, the spatial layout of the target and the mask does not play a crucial role. However, the spatial layout is of major importance. For example, in an earlier study, a vernier that consisted of two vertical lines slightly offset in the horizontal direction was presented as target (Duangudom, Francis, & Herzog, 2007). Observers indicated the offset direction (left vs. right). When two lines flanked the vernier (i.e., a typical metacontrast mask), B-type masking occurred in accordance with most previous findings and theories (Figure 1A; Duangudom et al., 2007). However, when arrays of lines, including the single flanking lines, were presented, strong A-type masking occurred (Figure 1A). Importantly, the inner contours of the masks (i.e., the innermost flanking lines) were identical in both conditions. Most theories of masking cannot easily

explain these results. Duangudom et al. (2007) proposed that target–mask grouping plays a crucial role. When target and mask are part of one group, masking is strong. When target and mask ungroup, masking is weak (e.g., Duangudom et al., 2007; Herzog & Fahle, 2002). Furthermore, it was proposed that the very same mechanism also holds true for other visual phenomena such as crowding (Malania, Herzog, & Westheimer, 2007; Manassi, Sayim, & Herzog, 2012; Saarela, Sayim, Westheimer, & Herzog, 2009; Sayim, Westheimer, & Herzog, 2008, 2010, 2011), surround suppression (Saarela & Herzog, 2008), and motion integration (see also nonretinotopic processing; Boi, Öğmen, Krummenacher, Otto, & Herzog, 2009; Öğmen, Otto, & Herzog, 2006).

There are many parallels between crowding and masking. In crowding, a target that can be easily discriminated in isolation is harder (or impossible) to discern when flanked by close-by items. Thus, crowding can be described as a form of masking. There are also differences between crowding and masking. For example, masking often strongly deteriorates target detection, while crowding is usually assumed not to influence detection (but see Allard & Cavanagh, 2011). As with masking, the underlying mechanisms of crowding are still largely unknown. A number of recent crowding studies showed that, on a level of perceptual organization, grouping between the target and the mask is a good predictor of performance: The stronger the grouping, the weaker the performance (e.g., Manassi, Sayim, & Herzog, 2013; Sayim et al., 2010). For example, it was shown that (un)grouping by color (Sayim et al., 2008; see also Kooi, Toet, Tripathy, & Levi, 1994) and good Gestalt (Sayim et al., 2010) plays a crucial role in crowding (Figure 1B).

Here, we investigate backward masking and compare our results with previous results in crowding. In particular, we ask how far the two processes are similar by investigating whether masking is subject to similar modulation by grouping cues as crowding. We show that color and Gestalt grouping leads to very different results in visual masking than it does in crowding. We argue that spatial factors override color differences between the target and the flankers in masking and that the short presentation times in masking are not sufficient to compute good Gestalt. Our results show that despite many similarities between crowding and masking there are clear differences in the pattern of results, suggesting that stimulus onset asynchronies (SOA) are of major importance in masking. We propose that SOA is a strong grouping factor that competes with (or reinforces) spatial and surface-based grouping cues.

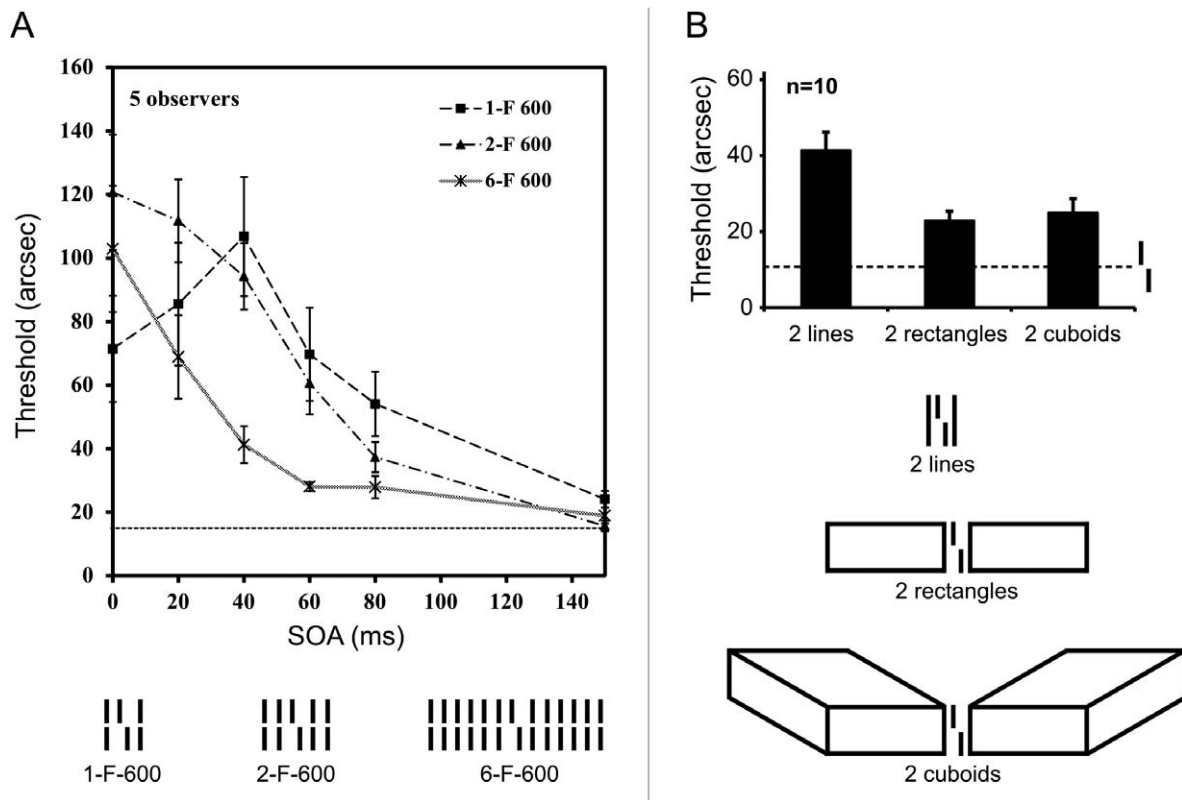


Figure 1. (A) Masking and the spatial layout of target and mask. When two lines flank a vernier target (condition 1-F 600, B-type masking occurs. When arrays of lines flank the vernier (2-F 600, 6-F 600), A-type masking occurs (adapted from Duangudom et al., 2007). (B) Crowding and spatial layout. A vernier flanked by two lines yields higher thresholds than does a vernier flanked by two rectangles or two cuboids that include the two lines (adapted from Sayim et al., 2010).

## General methods

### Participants

Observers were students of the Ecole Polytechnique Fédérale de Lausanne (EPFL) or the Université de Lausanne (UNIL). Five observers participated in Experiment 1 (one female); four observers of Experiment 1 and two additional observers participated in Experiment 2 (one female); three observers who participated in Experiments 1 and 2 and two additional observers participated in Experiment 3 (two females). Observers were between 20 and 35 years of age. All observers were naive to the purpose of the experiment. Before the experiments, observers were informed about the general purpose of the experiment and gave their written consent. Participants were told that they could quit the experiment at any time. Experiments were approved by the local ethics committee.

The Freiburg visual acuity test was used to determine observers' visual acuity (Bach, 1996). To participate in the experiments, observers had to reach a value of 1.0 (corresponding to 20/20) for at least one eye. All

subjects had normal or corrected-to-normal visual acuity. The Ishihara pseudoisochromatic color plates were used to test for red–green deficiencies; no color deficiencies were observed.

### Apparatus

In Experiment 1, stimuli were presented on a Philips 201B4 CRT monitor (Philips, Amsterdam, the Netherlands) driven by a standard accelerated graphics card with a screen resolution of 1024 by 768 pixels. The monitor's white point was adjusted to D65. Color space was computationally linearized by applying individual gamma corrections to each color channel. A Minolta CA-210 display color analyzer (Minolta, Osaka, Japan) was used for calibration measurements. Observers viewed the screen from a distance of 5 m. In Experiments 2 and 3, stimuli were presented on an X-Y display (HP-1332A with a P31 phosphor; Hewlett-Packard, Palo Alto, CA) controlled by a PC via fast 16-bit D/A converters. Observers viewed the screen from a distance of 2 m. The experimental room was dimly illuminated (0.5 lx).

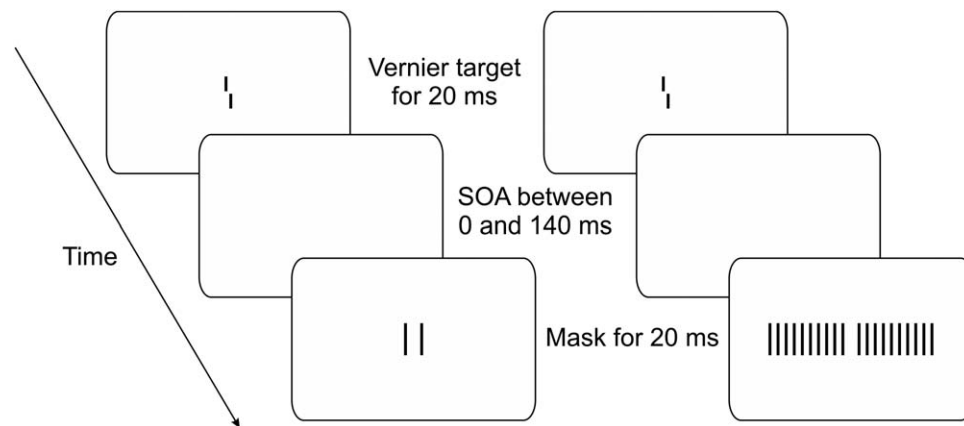


Figure 2. Stimuli. A vernier target was presented for 20 ms, followed by an ISI and a mask. Left panel: A standard metacontrast mask consisting of two lines. Right panel: A mask consisting of 20 lines (10 on each side).

## Stimuli and procedure

A foveally presented vernier target was flanked by different mask configurations (Figures 2 through 5). Verniers consisted of two vertical 10' long lines separated by a vertical gap of 1'. In each trial, the vernier was randomly offset either to the left or to the right. Observers indicated the offset direction. In all mask configurations, two lines were presented next to the vernier at a distance of 1.33'. When arrays of lines were presented, the spacing between each two lines on each side of the vernier was 1.33' (except for the “irregular” condition in Experiment 2). The lines and the vernier were horizontally centered in all conditions. The vernier was presented for 20 ms. The mask was presented for 20 ms at a variable SOA of 0, 20, 40, 60, and 80 ms (Figure 2; in Experiment 1 additionally 100, 120, and 140 ms in two conditions—see Figure 3). In all experiments, thresholds for verniers without masks were measured as a baseline.

All mask and SOA conditions were presented in separate blocks. In each block, left and right offsets were presented equally often in different pseudorandom sequences. Maximally four subsequent offsets in one direction were presented. After each trial, the screen remained blank for maximal 3 s, during which the observer was required to make a response. After the response, the screen remained blank for 500 ms until the next trial. Errors and omissions were indicated by auditory feedback.

An adaptive staircase procedure was used to determine the threshold for which an observer reached 75% correct responses (parameter estimation by sequential testing [PEST]; Taylor & Creelman, 1967). The threshold and slope of the psychometric function (cumulative Gaussian) were estimated by means of a maximum likelihood analysis, taking all trials into account (Wichmann & Hill, 2001). The guessing rate was set to 50%, the lapse rate to 3%. Starting offset was

150". The order of conditions was randomized individually for each observer. Observers performed two blocks of 80 trials in each condition. To compensate for possible learning effects, the order of conditions was reversed after each condition had been measured once.

## Experiment 1: Color

### Materials and methods

Red vernier targets were presented on a black background flanked by isoluminant green or red flanking lines (13.5 cd/m<sup>2</sup>) with the same length as the vernier (21'). Four different mask conditions were used. In the first condition, single red lines to the left and right of the vernier were presented at a distance of 1.33' (two-red condition). In the second condition, the two masking lines were green (two-green condition). In the third condition, 10 red lines—including the two lines of the two-red condition—were presented on each side of the vernier (20-red condition). In the fourth condition, 10 green lines—including the green lines of the two-green condition—were presented on each side of the vernier (20-green condition). As baseline condition, performance for red verniers without masks was measured.

### Results

In the two-red condition, B-type masking occurred as expected (Figure 3). Thresholds peaked at an SOA of 60 ms. As there was still strong masking at the SOA of 80 ms, we added SOAs of 100, 120, and 140 ms in this condition (and in the two-green condition). As expected, masking strength declined, reaching the baseline



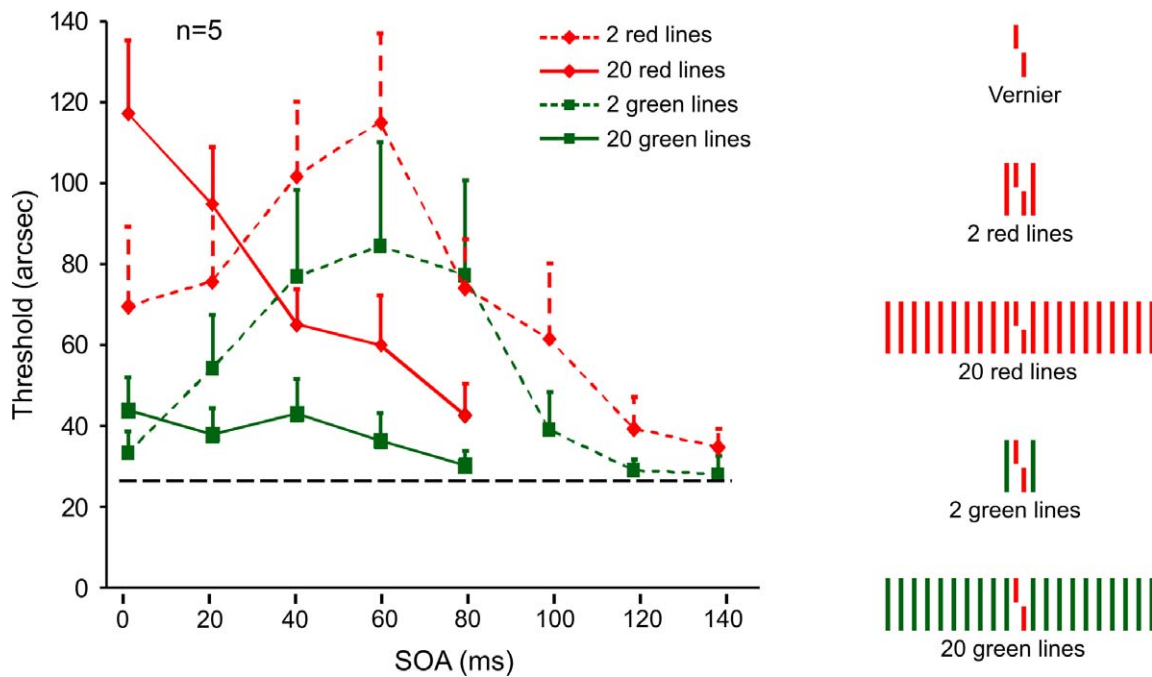


Figure 3. Results of Experiment 1. The two-red and two-green conditions yielded B-type masking. The 20-red condition yielded A-type masking. In the 20-green condition, there was only little masking compared with the other conditions. The dashed black line shows the unmasked vernier threshold (baseline level). Error bars indicate standard errors of the mean of five observers (only positive error bars are shown).

threshold at an SOA of around 140 ms. Also in the two-green condition, B-type masking occurred. As in the two-red condition, thresholds peaked at the SOA of 60 ms and declined with shorter and longer SOAs. Almost no masking occurred at the maximum SOA of 140 ms.

To compare performance in the two two-line conditions, we performed a repeated measures analysis of variance (ANOVA) with the factors color (two-red and two-green) and SOA (eight levels from 0 to 140 ms). We found a main effect of SOA,  $F(7, 28) = 5.585$ ;  $p < 0.001$ , no effect of color,  $F(1, 4) = 5.960$ ;  $p = 0.071$ , and no interaction,  $F(7, 28) = 0.821$ ;  $p = 0.578$ .

The 20-red condition yielded A-type masking. Thresholds were highest when vernier and mask were presented simultaneously (SOA = 0 ms) and declined with increasing SOAs, reaching the baseline level at an SOA of 80 ms. The masking function in the 20-green condition was nearly flat with a shallow slope descending to baseline level.

We again compared the two color conditions. A repeated-measures ANOVA with the two factors color (20-red and 20-green) and SOA (five levels from 0 to 80 ms) revealed a main effect of color,  $F(1, 4) = 27.162$ ;  $p < 0.01$ , a main effect of SOA,  $F(4, 16) = 8.551$ ;  $p < 0.001$ , and an interaction,  $F(4, 16) = 3.967$ ;  $p < 0.05$ .

We did not statistically test whether the resulting masking functions were A-type or B-type functions as

we were not interested in the function types per se but rather the differences between the conditions.

## Discussion

The two conditions with red flanking lines replicate previous results where two white lines flanking a white vernier yielded strong B-type masking and 20 white lines flanking a white vernier yielded A-type masking (Duangudom et al., 2007). The 20-green condition leads to weak A-type masking as did long flanking lines in a previous study (Duangudom et al., 2007). We interpret these results in terms of ungrouping: Masking is weak because long and green lines ungroup from the vernier because of length or color dissimilarity. In this respect, the result of the two-green condition is surprising. We had expected weak A-type masking, similar to the 20-green condition, because of ungrouping by color differences. However, while there was a clear trend toward better performance in the two-green compared with the two-red condition, masking was strong and of B-type. This result is contrary to crowding studies where color differences between the target and the flankers yielded only very weak interference for both two and 20 flankers (Manassi et al., 2012; Sayim et al., 2008). It seems that color

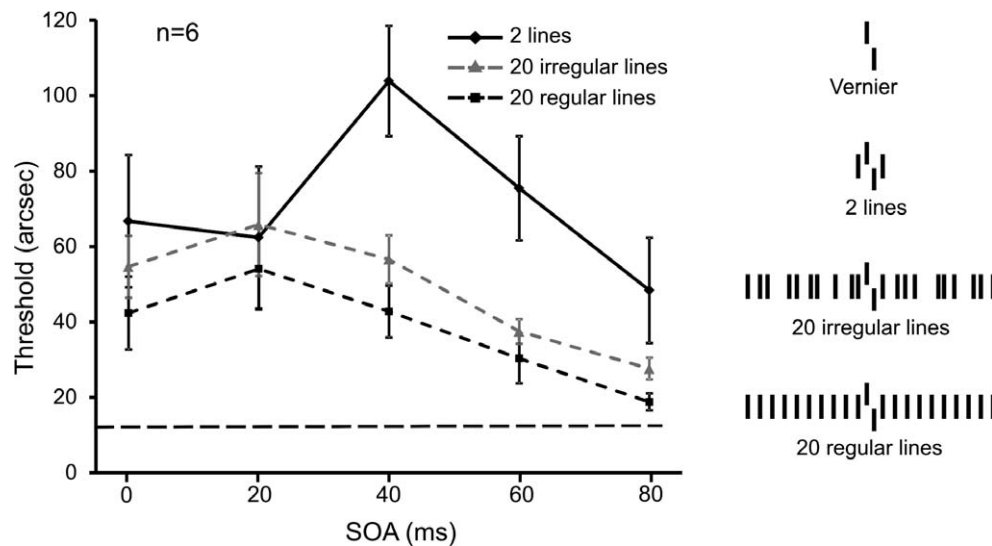


Figure 4. Results of Experiment 2. The two-lines condition yielded B-type masking. Masking strength was weaker in the 20-regular compared with the 20-irregular condition and weaker in both 20-lines conditions compared with the two-lines condition. The dashed line indicates the unmasked vernier threshold. Error bars indicate standard errors of the mean for six observers.

differences between the target and the flankers alone are not decisive in backward masking.

## Experiment 2: Regularity

### Materials and methods

Verniers were flanked by 10.5' lines (i.e., half the length of the vernier, including the vertical gap). In the first condition, single lines with a distance of 1.33' to the center of the vernier were presented to the left and right (two-lines condition). In the second condition, nine further lines were added to the single lines with a distance of 1.33' between each two lines (20-regular condition). In the third condition, the horizontal positions of the eight inner lines of the two arrays were randomly varied in each trial (20-irregular condition). The margin of each line's position was set to 45% (0.60' to the left and right) the distance between each two lines (1.33') to ensure that lines never overlapped. As a baseline condition, the vernier without mask was presented.

### Results

In the two-lines condition, B-type masking occurred (Figure 4). The threshold peaked at an SOA of 40 ms. When the vernier was masked by 20 regularly spaced lines on each side (20-regular condition), masking strength was weaker than in the two-lines condition for

all SOAs. Thresholds decreased from the peak threshold (at 20 ms) with increasing SOAs. The type of the masking function (A- vs. B-type) is unclear.

The masking function of the 20-irregular condition was almost perfectly parallel to the function of the 20-regular condition. Thresholds in the 20-irregular condition were higher than in the 20-regular condition for all SOAs.

A repeated-measures ANOVA with the two factors mask type (two-lines, 20-regular, 20-irregular) and SOA (five levels from 0 to 80 ms) revealed a main effect of mask type,  $F(2, 10) = 29.641$ ;  $p < 0.001$ , a main effect of SOA,  $F(4, 20) = 3.87$ ;  $p < 0.05$ , and no interaction,  $F(8, 40) = 1.174$ ;  $p = 0.338$ . Next, we asked whether the main effect of mask type was mainly driven by the high thresholds in the two-lines condition or whether there was a difference between the two 20-lines conditions. A planned comparison revealed a difference between the 20-regular and 20-irregular conditions,  $F(1, 5) = 6.580$ ,  $p = 0.05$  (note that the rounded  $p$  value equals 0.05).

### Discussion

As expected, we found B-type masking in the two-lines condition, replicating our results from Experiment 1 and earlier studies (Duangudom et al., 2007). Masking was weaker in the 20-regular condition compared with the two-lines condition. We, again, interpret this result in terms of grouping. When the innermost lines are ungrouped from the vernier through grouping with the additional lines in the 20-regular condition, there is less masking. When the

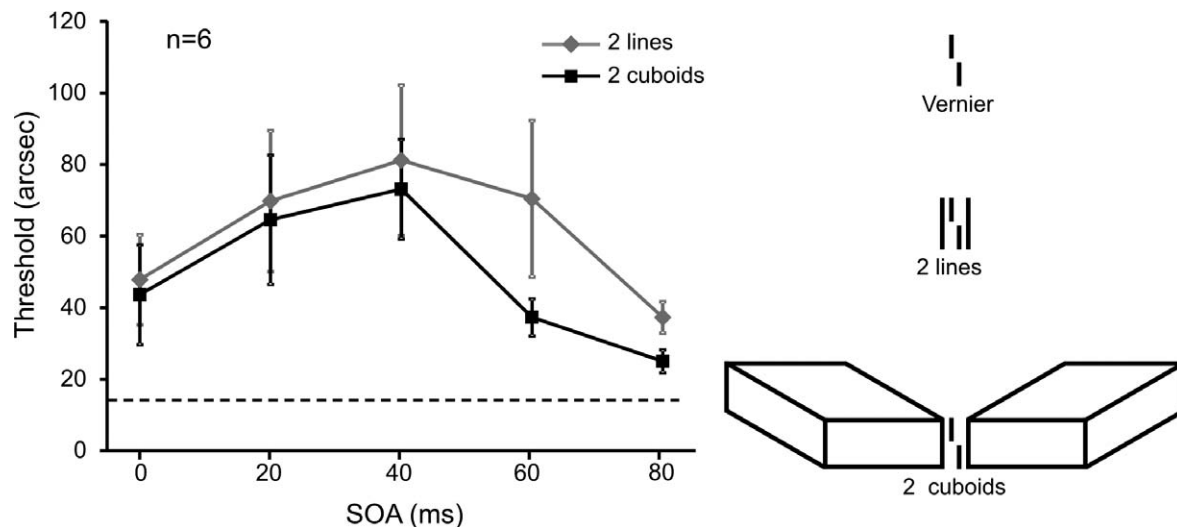


Figure 5. Results of Experiment 3. Both the two-lines condition and the two-cuboids condition yielded B-type masking. There is a trend for less masking in the two-cuboids condition for SOAs longer than 40 ms. The dashed line indicates the unmasked vernier threshold. Error bars indicate standard errors of the mean for six observers.

spacing between the lines was irregular (20-irregular condition), masking was stronger than in the regular condition. Again, this is in line with an interpretation in terms of grouping. Grouping of the lines in the arrays is reduced when the spacing is irregular. The results are similar to earlier results in crowding where horizontal or vertical irregular spacing between flankers increased interference compared with regular spacing (Manassi et al., 2012).

## Experiment 3: Good Gestalt

### Materials and methods

In the first condition, verniers were flanked by two lines with the same length (21') as the vernier. In the second condition, these lines were part of two cuboids (two-cuboids condition; Figure 5). The rectangles making up the front side of the cuboids had a width of 46.7'. The entire cuboids were each 70.2' wide and 44.7' high.

### Results

In the two-lines and two-cuboids condition, the masking function is of type-B with a peak threshold at an SOA of 40 ms (Figure 5). There is a trend toward lower masking in the two-cuboids condition than in the two-lines condition for longer SOAs.

A repeated-measures ANOVA with the two factors mask type (two-lines and two-cuboids) and SOA (five levels from 0 to 80 ms) revealed no main effect of mask type,  $F(1, 4) = 3.870$ ;  $p = 0.121$ , a main effect of SOA  $F(4, 16) = 3.375$ ;  $p < 0.05$ , and no interaction,  $F(4, 16) = 0.899$ ;  $p = 0.488$ .

### Discussion

As in Experiments 1 and 2, we expected and found B-type masking in the two-lines condition. In the two-cuboids condition, however, we expected weak A-type masking similar to the 20-green condition (Experiment 1) because of ungrouping by shape dissimilarity and good Gestalt. This expectation was based on earlier results in crowding where a foveal vernier target that was flanked by single lines yielded worse performance compared with flanking cuboids that contained the single lines (Sayim et al., 2010). The same result was observed when comparing rectangles and single lines in peripheral crowding (Manassi et al., 2012).

However, at larger SOAs, there was a trend for better performance in the two-cuboids compared with the two-lines condition (see Figure 5), indicating that the ungrouping by good Gestalt requires longer SOAs than other ungroupings—for example, in the 20-lines conditions where relative improvement compared with the two-lines conditions occurred at shorter SOAs. Alternatively, the integration of information from a larger area in the two-cuboids condition could require longer processing times.

## General discussion

As in previous studies, we found that visual masking strongly depends on the spatial configuration of the target and the mask (e.g., Duangudom et al., 2007). The results are not easily explained by models of masking, which are based only on local spatial interactions and temporal delays (Breitmeyer & Ganz, 1976; Werner, 1935). These models are reminiscent of many other research topics in vision science where local spatial approaches prevail despite accumulating evidence for the important influence of global configurations.

For example, in crowding, performance on a target deteriorates when flankers are next to the target, similar to visual masking (but elements are presented simultaneously and usually for much longer durations [ $>100$  ms] than in visual masking). As in masking, the deleterious influence of flankers is often explained by local interactions between the target and the flankers, for example, by spatial pooling (e.g., Parkes et al., 2001; Pelli et al., 2004; Wilkinson, Wilson, & Ellemberg, 1997). In crowding, we showed that single flanking lines presented on both sides of a vernier target deteriorate performance strongly. However, crowding almost vanishes when the lines are part of a good Gestalt, such as a rectangle or cuboid (Manassi et al., 2012; Sayim et al., 2010). Simple pooling models cannot account for these effects of global layout because the flankers next to the target are identical in the local and the global configurations. Moreover, we showed that the usual reduction of crowding by local color differences between the target and the flankers (as in the two-green and 20-green conditions; Figure 3; see also Kooi et al., 1994) does not occur when the target and the flankers are grouped within a texture-like pattern of alternating red and green lines (Manassi et al., 2012; Sayim et al., 2008). Based on these results, we proposed that crowding is strong when target and flankers group and weak when they ungroup. We found very similar effects for foveal (e.g., Sayim et al., 2010) and peripheral (e.g., Manassi et al., 2012) crowding, psychophysically and electrophysiologically (Chicherov, Plomp, & Herzog, 2012). Here, we asked whether similar effects also hold true for visual backward masking.

In Experiment 1, we presented a red vernier target that was flanked by either two or 20 red lines, yielding strong A-type masking in the 20-red condition and strong B-type masking in the two-red condition. These results replicate previous findings where A-type masking occurred when 20 lines flanked the vernier and B-type masking occurred when two lines flanked the vernier (Duangudom et al., 2007). Next, we changed the color of the masks to (isoluminant) green. In the 20-green condition, we obtained much weaker A-type masking than in the 20-red condition. This was expected because of target–mask dissimilarity, in

accordance with most results in masking and crowding (e.g., masking: Bevan, Jonides, & Collyer, 1970; Foster, 1976; Reeves, 1981; Yellott & Wandell, 1976; but see Bowen, Pokorny, & Cacciat, 1977; crowding: Kooi et al., 1994; Sayim et al., 2008). Surprisingly, we found strong B-type masking in the two-green condition, similar to the two-red condition. In crowding, by contrast, the number of green lines flanking a red vernier did not influence performance (Sayim et al., 2008). Differences between presentation times in masking (20 ms) and crowding (150 ms) can be ruled out as an explanation as thresholds in Experiment 1 were equally low in the two-green and the 20-green conditions when target and flankers were presented simultaneously (SOA = 0; Figure 3).

We propose an explanation in terms of spatiotemporal grouping. Simultaneous presentation of target and flankers yielded strong grouping (and strong masking) when the target and the flankers had the same color. When the flankers were green and the target was red, target and flankers ungrouped and masking was weak. Hence, when target and flankers were presented simultaneously, the results resemble those in crowding (Sayim et al., 2008). For nonzero SOAs, however, grouping occurs not only by color cues but also by the temporal delays. In the 20-lines conditions, the simultaneous presentation of the flanking line supports within-flanker grouping. In the two-lines conditions, the temporal delay induces an apparent motion cue (i.e., binding across space and time; see, e.g., Otto, Öğmen, & Herzog, 2006). We argue that in the two-green condition, this apparent motion cue overrides the color difference between the green flankers and the red vernier. This is in accordance with results showing that apparent motion is not highly sensitive to feature differences such as color (e.g., Navon, 1976).

We propose that maximum interference at SOAs around 60 ms reflects optimal grouping of target and flankers caused by apparent motion. Shorter and longer SOAs reduce apparent motion. In the 20-lines conditions, on the other hand, spatial grouping of the innermost lines with the additional lines trumps temporal grouping with the target and reduces apparent motion. Hence, spatial and temporal factors are strongly interlinked (for the relation of temporal integration and spatial grouping, see, e.g., Hermens, Scharnowski, & Herzog, 2009).

Interestingly, a study with amblyopes showed that performance in a purely temporal crowding paradigm correlated with performance in spatial crowding (Bonneh, Sagi, & Polat, 2007), indicating that these two forms of interference might be related. In particular, the authors proposed that strong sustained inhibition in strabismic amblyopes may underlie both increased temporal and spatial crowding. On a level of perceptual organization, we suggest that their results show that the



capability to isolate a target from distracters may be a general capacity of the visual system (temporal as well as spatial) to group and ungroup parts and wholes.

The results of Experiment 2 resembled earlier findings in crowding where irregular flank configurations yielded stronger interference than did regular configurations (e.g., Saarela, Westheimer, & Herzog, 2010). These results can be easily accounted for by grouping. Masking was weaker when the spacing between the flanking lines was regular compared with irregular. It seems that grouping between the flankers was reduced when the spacing between the flankers was irregular compared with regular.

However, the results of Experiment 3 seem to be at odds with an explanation based on grouping. We found that performance with two flanking lines was not worse compared with two cuboids containing the two lines. This result is different from crowding studies where we found better performance when flanking lines were embedded in Gestalts, such as cuboids or rectangles (Manassi et al., 2012; Sayim et al., 2010). While the masking functions of the two conditions are roughly on top of each other up to 40 ms, there is a trend toward better performance in the cuboids condition with increasing SOA. This trend may indicate that ungrouping of the target and the cuboids requires more time than the ungrouping of a target and simple structures, such as arrays of flanking lines, because of additional processing time required by complex Gestalt formation (see also Feldman, 2007; Kimchi, 1998).

Overall, our results show that B-type functions peaked around 40 to 60 ms whereas performance in A-type functions was already close to baseline at these SOAs. We propose that shallow A-type functions occur when target and flankers are not strongly grouped. Steep A-type functions occur when grouping by spatial and surface-based factors is strong but is quickly reduced by increasing temporal delays. B-type functions, on the other hand, occur when increasing SOAs initially increase target–flanker grouping. While ungrouping can occur early in visual processing (e.g., 20-green condition in Experiment 1; see also Sayim et al., 2011), feedback processes may play a role to either increase or decrease target–flanker grouping. For example, B-type functions could primarily be the result of higher level feedback that increases the integration of target and flankers because of sufficient spatial similarity. On a level of perceptual organization, this integration can be described as grouping over time, integrating target and flankers as belonging to a single object. On the other hand, when similarity between flanking elements is strong—for example, when presenting arrays of lines as flankers—feedback mechanism only strengthen the early, possibly purely feed forward, organization of target and flankers into different objects (groups).

In this framework, the shape of a masking function depends not only on target–flanker similarity but also on the efficiency of grouping between flanking elements. When grouping between flanking elements is weak, flanker integration takes longer processing time and, at the same time, feedback processes that integrate target and flankers take effect for longer periods. For example, if strong grouping among the cuboid elements was not established early, feedback could favor integration of target and local flanker features.

All in all, the pattern of results in backward masking is more complex than in crowding. Masking strength depended on a combination of spatial and temporal factors and color cues. Single lines always yielded B-type masking, and larger configurations—except the cuboids—always yielded A-type masking. Ungrouping as determined in crowding experiments (e.g., Sayim et al., 2008) did not always predict masking. Future studies will show how interactions between spatial and temporal factors, basic features, and grouping determine backward masking.

*Keywords: backward masking, grouping, Gestalt, crowding, perceptual organization*

## Acknowledgments

We thank Marc Repnow for technical support. This work was supported by the Swiss National Science Foundation (SNF) project “Basics of visual processing: what crowds in crowding?”. Bilge Sayim is currently funded by a Pegasus Marie Curie fellowship by the FWO.

Commercial relationships: none.

Corresponding author: Bilge Sayim.

Email: bilge.sayim@ppw.kuleuven.be.

Address: Laboratory of Experimental Psychology, University of Leuven, Leuven, Belgium.

## References

- Allard, R., & Cavanagh, P. (2011). Crowding in a detection task: External noise triggers change in processing strategy. *Vision Research*, 51(4), 408–416.
- Bach, M. (1996). The Freiburg Visual Acuity test—Automatic measurement of visual acuity. *Optometry and Vision Science*, 73, 49–53.
- Bachmann, T. (1984). The process of perceptual retouch: Nonspecific afferent activation dynamics

- in explaining visual masking. *Perception and Psychophysics*, 35, 69–84.
- Bachmann, T. (1994). *Psychophysiology of visual masking: The fine structure of conscious experience*. Commack, NY: NOVA Science Publishers.
- Bevan, W., Jonides, J., & Collyer, S. C. (1970). Chromatic relationships in metacontrast suppression. *Psychonomic Science*, 19, 367–368.
- Boi, M., Öğmen, H., Krummenacher, J., Otto, T. U., & Herzog, M. H. (2009). A (fascinating) litmus test for human retino- vs. non-retinotopic processing. *Journal of Vision*, 9(13):5, 1–11, <http://www.journalofvision.org/content/9/13/5>, doi:10.1167/9.13.5. [PubMed] [Article]
- Bonneh, Y. S., Sagi, D., & Polat, U. (2007). Spatial and temporal crowding in amblyopia. *Vision Research*, 47(14), 1950–1962.
- Bowen, R. W., Pokorny, J., & Cacciat, D. (1977). Metacontrast masking depends on luminance transients. *Vision Research*, 17, 971–975.
- Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, 83, 1–36.
- Breitmeyer, B. G., & Öğmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision*. New York: Oxford University Press.
- Chichero, V., Plomp, G., & Herzog, M. H. (2012). The N1 wave amplitude reflects perceptual grouping and correlates with crowding. *Journal of Vision*, 12(9):1296, <http://www.journalofvision.org/content/12/9/1296>, doi:10.1167/12.9.1296. [Abstract]
- Duangudom, V., Francis, G., & Herzog, M. H. (2007). What is the strength of a mask in visual metacontrast masking? *Journal of Vision*, 7(1):7, 1–10, <http://www.journalofvision.org/content/7/1/7>, doi:10.1167/7.1.7. [PubMed] [Article]
- Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science*, 8, 135–139.
- Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, 4, 345–352.
- Fehrer, E., & Raab, D. (1962). Reaction time to stimuli masked by metacontrast. *Journal of Experimental Psychology*, 63, 143–147.
- Feldman, J. (2007). Formation of visual “objects” in the early computation of spatial relations. *Perception and Psychophysics*, 69(5), 816–827.
- Foster, D. H. (1976). Rod-cone interaction in the after-flash effect. *Vision Research*, 16, 393–396.
- Francis, G. (1997). Cortical dynamics of lateral inhibition: Metacontrast masking. *Psychological Review*, 104, 572–594.
- Francis, G., & Herzog, M. H. (2004). Testing quantitative models of backward masking. *Psychonomic Bulletin and Review*, 11(1), 104–112.
- Hermens, F., Luksys, G., Gerstner, W., Herzog, M. H., & Ernst, U. (2008). Modeling spatial and temporal aspects of visual backward masking. *Psychological Review*, 115(1), 83–100.
- Hermens, F., Scharnowski, F., & Herzog, M. H. (2009). Spatial grouping determines temporal integration. *Journal of Experimental Psychology: Human Perception and Performance*, 35(3), 595–610.
- Herzog, M. H., & Fahle, M. (2002). Effects of grouping in contextual modulation. *Nature*, 415, 433–436.
- Herzog, M. H., & Koch, C. (2001). Seeing properties of an invisible object: Feature inheritance and shine-through. *Proceedings of the National Academy of Sciences, USA*, 98, 4271–4275.
- Kimchi, R. (1998). Uniform connectedness and grouping in the perceptual organization of hierarchical patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1105–1118.
- Kolers, P. A. (1962). Intensity and contour effects in visual masking. *Vision Research*, 2, 277–294.
- Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial inter-action in peripheral vision. *Spatial Vision*, 8, 255–279.
- Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, 48(5), 635–654.
- Malania, M., Herzog, M. H., & Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds. *Journal of Vision*, 7(2):1, 1–7, <http://www.journalofvision.org/content/7/2/1>, doi:10.1167/7.2.1. [PubMed] [Article]
- Manassi, M., Sayim, B., & Herzog, M. H. (2012). Grouping, pooling, and when bigger is better in visual crowding. *Journal of Vision*, 12(10):13, 1–14, <http://www.journalofvision.org/content/12/10/13>, doi:10.1167/12.10.13. [PubMed] [Article]
- Manassi, M., Sayim, B., & Herzog, M. H. (2013). When crowding of crowding leads to uncrowding. *Journal of Vision*, 13(13):10, 1–10, <http://www.journalofvision.org/content/13/13/10>, doi:10.1167/13.13.10. [PubMed] [Article]
- Navon, D. (1976). Irrelevance of figural identity for resolving ambiguities in apparent motion. *Journal*

- of Experimental Psychology: Human Perception and Performance*, 2(1), 130–138.
- Öğmen, H., Otto, T. U., & Herzog, M. H. (2006). Perceptual grouping induces non-retinotopic feature attribution in human vision. *Vision Research*, 46(19), 3234–3242.
- Otto, T. U., Öğmen, H., & Herzog, M. H. (2006). The flight path of the phoenix – the visible trace of invisible elements in human vision. *Journal of Vision*, 6(10):7, 1079–1086, <http://www.journalofvision.org/content/6/10/7>, doi:10.1167/6.10.7. [PubMed] [Article]
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, 4, 739–744.
- 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://www.journalofvision.org/content/4/12/12>, doi:10.1167/4.12.12. [PubMed] [Article]
- Reeves, A. (1981). Metacontrast in hue substitution. *Vision Research*, 21(6), 907–912.
- Saarela, T. P., & Herzog, M. H. (2008). Time-course and surround modulation of contrast masking in human vision. *Journal of Vision*, 8(3):23, 1–10, <http://www.journalofvision.org/content/8/3/23>, doi:10.1167/8.3.23. [PubMed] [Article]
- Saarela, T. P., Sayim, B., Westheimer, G., & Herzog, M. H. (2009). Global stimulus configuration modulates crowding. *Journal of Vision*, 9(2):5, 1–11, <http://www.journalofvision.org/content/9/2/5>, doi:10.1167/9.2.5. [PubMed] [Article]
- Saarela, T. P., Westheimer, G., & Herzog, M. H. (2010). The effect of spacing regularity on visual crowding. *Journal of Vision*, 10(10):17, 1–7, <http://www.journalofvision.org/content/10/10/17>, doi:10.1167/10.10.17. [PubMed] [Article]
- Sayim, B., Westheimer, G., & Herzog, M. H. (2008). Contrast polarity, chromaticity, and stereoscopic depth modulate contextual interactions in vernier acuity. *Journal of Vision*, 8(8):12, 1–9, <http://www.journalofvision.org/content/8/8/12>, doi:10.1167/8.8.12. [PubMed] [Article]
- Sayim, B., Westheimer, G., & Herzog, M. H. (2010). Gestalt factors modulate basic spatial vision. *Psychological Science*, 21(5), 641–644.
- Sayim, B., Westheimer, G., & Herzog, M. H. (2011). Quantifying target conspicuity in contextual modulation by visual search. *Journal of Vision*, 11(1):6, 1–11, <http://www.journalofvision.org/content/11/1/6>, doi:10.1167/11.1.6. [PubMed] [Article]
- Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America*, 41, 782–787.
- Werner, H. (1935). Studies on contour: I. Qualitative analysis. *American Journal of Psychology*, 47, 40–64.
- Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, 15(4), 160–168.
- Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception and Psychophysics*, 63(8), 1293–1313.
- Wilkinson, F., Wilson, H. R., & Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture arrays. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, 14, 2057–2068.
- Yellott, J. I., Jr., & Wandell, B. A. (1976). Color properties of the contrast flash effect: Monoptic vs. dichoptic comparisons. *Vision Research*, 16(11), 1275–1280.