Metacontrast masking within and between visual channels: Effects of orientation and spatial frequency contrasts

Maximilian Bruchmann  
Institute for Biomagnetism and Biosignal Analysis, University of Muenster, Germany

Bruno G. Breitmeyer  
Department of Psychology, University of Houston, Houston, TX, USA, & Center for Neuro-Engineering and Cognitive Science, University of Houston, Houston, TX, USA

Christo Pantev  
Institute for Biomagnetism and Biosignal Analysis, University of Muenster, Germany

We measured the strength and optimal target–mask onset asynchrony (SOA\textsubscript{max}) of metacontrast masking using Gabor patches as targets and sinusoidal rings with Gaussian envelopes as masks. We varied spatial frequencies (f) between 0.5 and 8 cpd to manipulate the degree to which spatial frequency channels in the visual system are triggered. By varying spatial frequencies as well as spatial frequency contrast (Δf) between target and mask, we measured the properties of inter- as well as intra-channel inhibition. We found that an increase of the mask’s spatial frequency decreased its effectiveness but did not change its SOA\textsubscript{max}. When orientation contrast was introduced between targets and masks with the same spatial frequency, SOA\textsubscript{max} increased with orientation contrast. An effect of orientation contrast was not observed with low spatial frequency-on-high spatial frequency masking, indicating that orientation selectivity is a unique feature of within-channel masking. Spatial frequency contrast affects SOA\textsubscript{max} and effectiveness in an asymmetric fashion: low-on-high masking is strong and yields a longer SOA\textsubscript{max}, compared to low-on-low and high-on-high masking; high-on-low masking is ineffective.

Keywords: backward masking, surround suppression, sustained and transient channels


**Introduction**

Metacontrast masking refers to the phenomenon that the visibility of a briefly presented target stimulus can be reduced by the presentation of a second stimulus, the mask, if it (a) follows the target with a stimulus onset asynchrony around 30–100 ms, (b) closely surrounds the target, and (c) shares some critical visual features with the target.

Several theories and models have aimed at explaining this phenomenon. For a review, see Breitmeyer and Ögmen (2006). Two features in which these models can be differentiated are whether they assume neural feedback to play a crucial role in metacontrast (Breitmeyer & Ögmen, 2006; Fahrenfort, Scholte, & Lamme, 2007; Haynes, Driver, & Rees, 2005; Ögmen, Breitmeyer, & Melvin, 2003) or not (Macknik & Livingstone, 1998; Macknik & Martinez-Conde, 2007), and whether they assume target–mask interactions to take place within a single neural pathway (Macknik & Livingstone, 1998; Macknik & Martinez-Conde, 2007) or more (Bachmann, 1997; Breitmeyer, 1984; Breitmeyer & Ganz, 1976; Breitmeyer & Ögmen, 2006; Breitmeyer, Rudd, & Dunn, 1981; Ögmen et al., 2003; Rogowitz, 1983).

The empirical evidence pro and contra feedback theories is mainly fed by electrophysiological studies and is not the subject of the present article. We focus on the dissociation of one- and multi-pathway models by means of psychophysical experiments.

Several types of multi-process models have evolved in the past to explain metacontrast masking. One such model, the Sustained-Transient model (ST model), was first described by Breitmeyer and Ganz (1976) and explains masking by means of inhibitory interactions between and within the transient and sustained neural pathways or channels, which are defined by their respective spatial and temporal response profiles. Especially the predictions of the temporal response profile have been supported by many psychophysical experiments (Breitmeyer & Ganz, 1977; Breitmeyer, Levi, & Harwerth, 1981; Lupp, Hauske, & Wolf, 1976; Tolhurst, 1975a; Macknik & Martinez-Conde, 2007) or more (Bachmann, 1997; Breitmeyer, 1984; Breitmeyer & Ganz, 1976; Breitmeyer & Ögmen, 2006; Breitmeyer, Rudd, & Dunn, 1981; Ögmen et al., 2003; Rogowitz, 1983).

doi: 10.1167/10.6.12  
Received December 11, 2009; published June 16, 2010  
ISSN 1534-7362 © ARVO
confirming that the transient channel responds quickly and is short-lived to the onset of a stimulus, while the sustained channel shows a delayed and more prolonged response. As a consequence in order to optimize inter-channel, transient-on-sustained inhibition, the mask would have to be delayed by some tens of milliseconds for its transient-channel activation to have a maximal suppressive effect on the target’s sustained-channel activation, thus producing an SOA$_{\text{max}} > 0$ ms.

On the other hand, we know from psychophysical studies that transient channels respond preferably to low spatial frequencies ($f$) and sustained channels to high spatial frequencies (Legge, 1978; Parker, Salzen, & Lishman, 1982; Wilson, 1980; Wilson & Bergen, 1979). However, the link between the spatial response characteristics of the two channels and the role of spatial frequency in metacontrast masking is still unclear.

Breitmeyer and Ögmen (2006) propose that increasing the spatial frequency of a stimulus delays the transient as well as the sustained component. Thus, if the spatial frequency of the target and mask are increased in tandem, we should not expect significant changes of the optimal SOA (SOA$_{\text{max}}$). Furthermore, they propose that the magnitude of the transient component should decrease with increasing spatial frequency. This should lead to a decrease of masking effectiveness as the spatial frequency shared by the target and mask increases.

A part of these assumptions was tested by Rogowitz (1983). In contrast to the assumption cited above, the conditions in which target and mask share increasing frequencies show that SOA$_{\text{max}}$ does shift but in a non-systematic fashion. In the same conditions, her data also do not show the proposed relationship between spatial frequency and masking effectiveness. However, if the target’s $f$ is fixed while the mask’s $f$ increases, Rogowitz clearly shows that SOA$_{\text{max}}$ decreases. It may also be noted that only two subjects took part in the experiment with their results indicating considerable differences.

However, Rogowitz’s study and results are not necessarily tied to the sustained-transient approach but do bear more generally on spatiotemporal-specific channel models in which several temporal and spatial channels are assumed (Graham, 1972; Nilsson, Richmond, & Nelson, 1975; Pantle, 1971; Thompson, 1984; Wilson & Bergen, 1979; Wilson, McFarlane, & Phillips, 1983; Yang & Makous, 1994). For instance, it is well known that lower spatial frequency (LSF) channels respond more rapidly than higher spatial frequency (HSF) ones (Breitmeyer, 1975; Lupp et al., 1976; Vassilev & Mitov, 1976). Although a negative result has been reported (Klein & Stromeyer, 1980), other results also support the existence of inhibitory interactions between spatial frequency-specific channels (Sagi & Hochstein, 1985; Tolhurst, 1972). Such inhibitory interactions combined with the response-latency differences between LSF and HSF channels could also model metacontrast masking. In particular, assume that the mask is composed of low spatial frequencies and the target of high spatial frequencies. Then, on the one hand, given that an LSF channel can suppress an HSF channel, it is readily seen that the mask, activating a faster responding LSF channel, must be delayed by a few tens of milliseconds in order for it to optimally suppress the target’s visibility, since the target activates a slower responding HSF channel. Consequently, a typical U-shaped masking function relating target visibility to SOA should be obtained. The SOA at which target visibility is lowest is taken as the SOA$_{\text{max}}$. On the other hand, if the target and the mask share the same spatial frequency, the response latencies of their, respectively, activated spatial frequency channels ought to be nearly identical. Hence, the maximal inhibitory interaction between target and mask ought to occur when the onsets of the two stimuli are simultaneous.

The sustained-transient channel model, a version of this more general spatiotemporal frequency approach, is characterized by several properties missing from other versions. It assumes (a) that the visibility of the target, when presented alone, is directly proportional to the level of sustained-channel activity generated by the target, whereas (b) the suppressive effect of the mask is directly proportional to the level of transient activity generated by the mask. Thus, given that the latency of the transient channels is shorter than that of the sustained channels (Breitmeyer & Ögmen, 2006), (c) the inter-channel transient-on-sustained inhibition, which suppresses the visibility of the target, will be optimal at a positive SOA max of a few tens of milliseconds, thus yielding a non-monotonic/U-shaped metacontrast masking function with increases of SOA. Moreover, prior psychophysical studies (Breitmeyer & Julesz, 1975; Kulikowski & Tolhurst, 1973; Tolhurst, 1973, 1975b) have shown that while transient channels can be characterized as low-pass spatial filters, with a response optimal at low spatial frequencies and decreasing as spatial frequency increases beyond about 5 cpd, the sustained channels can be characterized by band-pass filters with optimal response at intermediate spatial frequencies ranging from 4 to 10 cpd. Thus, according to the sustained-transient channel model, metacontrast masking functions should therefore be determined jointly by the spatial frequency contents of the target and of the mask.

To test the role of spatial frequency channels in metacontrast, Ishikawa, Shimegi, and Sato (2006) presented a novel stimulus, which combined the classical disk-annulus arrangement with a sinusoidal grating. The authors tested the effects of target–mask differences in orientation, spatial frequency, and luminance contrast with various stimulus onset asynchronies (SOAs) and found sharp orientation and spatial frequency tuning curves for masking with short SOAs and broader tuning curves at long SOAs. The authors concluded that a slow target–response component conveys highly feature-specific information, which is suppressed by a similarly slow mask–response component at short SOAs; in contrast, a
fast mask–response component conveys less feature specificity, hence the broad spatial frequency and orientation tuning curves of the mask at long SOAs. This trend can be seen in the data of Ishikawa et al. in their Figures 2a and 2b. For spatial frequency, however, this trend is not as evident as it is for orientation and therefore deserves more critical empirical testing. Given the psychophysical finding (a) that frequencies below 2–4 cycles per degree (cpd) activate primarily the transient channel and while higher ones activate primarily the sustained channel and (b) that Ishikawa et al. tested the orientation effect only with 2 cpd stimuli and the spatial frequency effect with targets held constant at 2 cpd, their findings cannot be adequately interpreted in the light of the multi-channel and in particular the sustained-transient framework of visual masking.

With the first experiment, we aim at a replication of the orientation contrast effect observed by Ishikawa et al. but extend the observations to a wider range of spatial frequencies. A special focus will be placed on the temporal characteristics of the obtained masking functions in relation to orientation contrast and its modulation by spatial frequency. The second experiment is also an extension of the studies by Ishikawa et al. in the way that we systematically pair various mask frequencies with target frequencies. By this means, we aim at testing an asymmetric hypothesis about HSF-on-LSF vs. LSF-on-HSF masking, predicted by the sustained-transient theory. The third experiment will contrast the spatial selectivity of within- and between-channel masking.

### Experiment 1: Effect of orientation contrast and spatial frequency

#### Subjects

Nine subjects (6 females) took part in the experiment. All had normal or corrected-to-normal vision and no history of neurological or psychiatric diseases. Age ranged from 23 to 29, mean = 27 years. The subjects gave their informed consent and were paid €9 per hour.

All subjects gave informed written consent to participate in the experiment and received monetary compensation. All procedures were carried out according to the declaration of Helsinki and were approved by the ethical committee of the medical faculty of the University of Muenster.

#### Apparatus and stimuli

The experiments were run using MATLAB and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Stimuli were presented on calibrated ViewSonic G90fB CRT monitor at 100 Hz and a resolution of 1024 × 768 pixels at a viewing distance of 80 cm. The mean brightness of the monitor was set to approximately 50 cd/m² ($I_{\text{min}} = 0.442$ cd/m², $I_{\text{max}} = 100.145$ cd/m²). Participants gave their responses by pressing one of four buttons on an external response box.

All stimuli were always presented at the maximum Michelson contrast of ($I_{\text{max}} - I_{\text{min}})/(I_{\text{max}} + I_{\text{min}}) = 0.991$. As targets, we used Gabor patches with a diameter of 2 degrees of visual angle (deg; measured from $\pm 2.5$ to $\pm 2.5$ SD of the Gaussian envelope). As masks, we used a grating annulus with a Gaussian envelope. The diameter (at maximum height of the Gaussian) of the annulus was 3 deg; the width of the Gaussian envelope was 2 deg. Targets and masks were presented randomly 3.5 deg to the left or right of a central fixation mark with spatial frequencies $f = 0.5, 1, 2, 4$, and 8 cpd at orientations $\phi = 0, 30, 60, 90, 120$, and 150°. The orientation contrast between targets and masks could be $\Delta \phi = 0, 30, 60,$ or 90°. Figure 1 depicts the spatial and temporal arrangements of targets and masks.

#### Procedure

The participants were instructed to focus on a central fixation mark. After a randomized interval between 400 and 600 ms, a target–mask sequence appeared randomly to the left and right of the fixation mark. Target and mask
duration was always 30 ms. The stimulus onset asynchrony (SOA) between target and mask was either 0, 20, 30, 40, 50, 60, 70, 80, 100, or 120 ms. Stimulus orientation, spatial frequency, and SOA were randomized. Besides the target–mask sequences, we occasionally presented reference trials containing only the target or only the mask at each spatial frequency and orientation. In total, there were 210 experimental conditions (4 orientation contrasts × 5 spatial frequencies × 10 SOAs + 10 reference conditions). Each condition was repeated 30 times. The resulting 6300 trials were distributed over 5 sessions of 1 h. Before starting the main experiment, the participants had 1 h of training to get familiar with the task.

After each trial, the participants were asked to rate the visibility of the target with 4 buttons. They were instructed to press button 1 if the target was not visible at all, and button 4 if it was very visible, and to use buttons 2 and 3 for intermediate degrees of visibility. The participants were asked to try to use the full rating scale and to use the training session to establish an individual rating scheme. The participants were informed that maintaining the rating scheme over the sessions would increase data quality, which was quantified by a goodness-of-fit index of a masking function described below. To enhance the participants’ motivation, cinema tickets were offered as a reward for the highest goodness-of-fit.

Data normalization and function fitting

Subjects gave ratings based on their subjective rating scheme. Furthermore, stimulus visibility depends on several factors, including spatial frequency. In order to make masking effects comparable between subjects and spatial frequencies, we normalized the rating data by means of the ratings obtained from the reference trials in the following way:

\[
R_{\text{norm}} = \frac{R - R_{\text{maskonly}}}{R_{\text{targetonly}} - R_{\text{maskonly}}},
\]

The normalized rating \(R_{\text{norm}}\) of target visibility for a given SOA, spatial frequency, and orientation contrast is obtained for each subject, by subtracting the subject’s mean rating over the 30 trial repetitions per condition for an invisible (i.e., not presented) target \(R_{\text{maskonly}}\) (where the mask has the given spatial frequency) and dividing the difference by the maximum rating range defined as the difference between the mean rating of an unmasked target \(R_{\text{targetonly}}\) and \(R_{\text{maskonly}}\) with the given spatial frequency. Thus, \(R_{\text{norm}} = 1\) means “as visible as an unmasked target” and \(R_{\text{norm}} = 0\) means “invisible”. After normalization, the data were fit by the following function:

\[
\hat{R}(t)_{\text{norm}} = \alpha - \beta \cdot e^{-((\log(t) - t_0)/\sigma)^2}.
\]

The function is based on an inverted log-normal curve, where \(\alpha\) represents a baseline rating, \(\beta\) controls the depth of the function, \(t_0\) determines the temporal position of the minimum, and \(\sigma\) is a variance parameter and thus reflects the temporal specificity of the masking effect. A log-normal curve is regarded as appropriate for fitting metacontrast data since they typically show a relatively steep decrease of target visibility followed by a shallower increase with increasing SOA.

SOA\(_{\text{max}}\) is defined as the temporal position of the minimum of the fitted function. The effectiveness of masking is defined as the difference between \(\alpha\) and \(\hat{R}(t)_{\text{norm}}\) at the location of SOA\(_{\text{max}}\).

The individually fitted parameters were used for further analyses. A major advantage of fitting a function to the data over merely localizing the lowest data point is that the information of all data points is integrated to form a more stable estimate of the location and size of the true masking effect. By this means, one can also observe temporal shifts of masking functions below the sampling range of SOAs.

Results

An analysis of the control ratings for targets and masks presented alone showed that subjects made use of more or less the full range of ratings. However, Figure 2f shows that the ratings, nearly constant across spatial frequencies, are not greatly affected by the spatial contrast sensitivity function (CSF) as might be expected, indicating that subjects intuitively compensated for the CSF when giving subjective visibility ratings. The reason for this may be the fact that subjectively perceived contrast of gratings remains constant over a sufficiently wide range of spatial frequencies (e.g., see Peli, Arend, & Labianca, 1996) when, as in our experiments, a background luminance of 50 cd/m\(^2\) and high-contrast target and mask stimuli are used.

Figure 2f further shows that target visibility ratings in trials where only the mask was presented is on average slightly above the minimum rating of 1 and appears to vary non-monotonically with spatial frequency. A similar finding has been reported by Genter and Weisstein (1981) who reported that phantom gratings can appear in an empty region between two gratings if the display is flickered on and off. The visibility of these gratings was also shown to vary non-monotonically with spatial frequency.

Despite possible minor deviations from perfect contrast constancy, the ratings were normalized as described above.

The masking functions of one subject did not exhibit a U-shaped pattern. An interview with that subject revealed a misunderstanding of the task. The subject was therefore excluded from further analysis.

As noted above, target orientation (\(\phi\)) was randomized between 0° and 150°. This variation was not included in
the analysis as an experimental factor, as there would have been too few trial repetitions for a robust analysis. However, we checked by eye inspection for interactions of $\phi$ with the experimental factors orientation contrast ($\Delta\phi$), SOA, and spatial frequency ($f$). We did not observe any main effect of target orientation on visibility or any interaction with the experimental factors (see Supplementary Figure 1).

Figures 2a–2e show the averaged normalized visibility ratings of the remaining nine subjects. In all conditions, type B masking functions can be observed. $R_{\text{norm}}$ at SOA = 0 ms is always in the range of 1, indicating that there was no common onset masking. At SOA = 120 ms, visibility is on average not fully recovered. However, clear minima can be identified in all cases. The goodness-of-fit index (a possible maximum of 1.0) of all individual masking functions was 0.914.

For each subject and each condition, we calculated masking latency (SOA$_{\text{max}}$) and masking effectiveness ($E$) as explained above. Figure 3 contains the mean values. These measures were analyzed by means of a $4 \times 5$ analysis of variance (ANOVA) for repeated measures with the factors orientation contrast ($\Delta\phi$) and spatial frequency ($f$). A Mauchly test revealed that the assumption of sphericity was violated for the factor $f$. We will therefore report Greenhouse–Geisser corrected $p$-values for the main effect $f$ and the interaction $f \times \Delta\phi$. ANOVA results are summarized in Table 1.

Planned linear contrasts revealed that there was a significant linear decline of $E$ with increasing $\Delta\phi$, $F(1,8) = 13.515$, $p < 0.01$. Post hoc Tukey tests for the effect of
on SOAmax revealed that all levels differed significantly from each other (all $p < 0.001$), except the means for $\Delta\phi = 0^\circ$ and $\Delta\phi = 30^\circ$ ($p = 0.056$). Further post hoc Tukey tests were used to analyze the effect of $f$ on $E$: significant differences were found between $f = 0.5$ cpd and $f = 2$ cpd, $f = 0.5$ cpd and $f = 4$ cpd, $f = 0.5$ cpd and $f = 8$ cpd ($all \ p < 0.05$).

We replotted the data to analyze orientation tuning, i.e., the degree to which mask effectiveness is affected by orientation. Figure 4 shows in each plot mean visibility ratings as a function of spatial frequency, with separate lines for each orientation contrast. It can be seen that orientation contrast has little effect at SOA = 0 ms, increases up to about SOA = 40–50 ms and then decreases again. At SOA = 120 ms, orientation tuning is flat.

**Discussion**

The first hypothesis concerned a decrease of masking effectiveness with increasing orientation contrast. This has been shown by Ishikawa et al. (2006) and is in line with a general positive relationship between target–mask similarity and masking effectiveness. We showed that this effect is robust over a wide range of spatial frequencies and seems to be monotonically related to variations of orientation contrast.

As can be seen from Figure 3e, the interaction between $f$ and $\Delta\phi$ is significant because the effect of $\Delta\phi$ is less pronounced at $f = 0.5$. Otherwise, the effects seem additive. The observation that the orientation contrast effect is less pronounced at $f = 0.5$ may be explained by

<table>
<thead>
<tr>
<th>Measure</th>
<th>Effect</th>
<th>MSS</th>
<th>df</th>
<th>dfe</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA$_{\text{max}}$</td>
<td>$f$</td>
<td>284.45</td>
<td>4</td>
<td>32</td>
<td>0.845</td>
<td>0.454</td>
</tr>
<tr>
<td></td>
<td>$\Delta\phi$</td>
<td>6783.27</td>
<td>3</td>
<td>24</td>
<td>35.121</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>$f \times \Delta\phi$</td>
<td>775.136</td>
<td>12</td>
<td>96</td>
<td>2.933</td>
<td>0.077</td>
</tr>
<tr>
<td>$E$</td>
<td>$f$</td>
<td>0.54</td>
<td>4</td>
<td>32</td>
<td>5.606</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td></td>
<td>$\Delta\phi$</td>
<td>1.15</td>
<td>3</td>
<td>24</td>
<td>12.123</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>$f \times \Delta\phi$</td>
<td>0.57</td>
<td>12</td>
<td>96</td>
<td>2.064</td>
<td>0.128</td>
</tr>
</tbody>
</table>

Table 1. Summary of 4 (orientation contrast) $\times$ 5 (spatial frequency) ANOVA with repeated measures. The symbols $f$ and $\Delta\phi$ denote the main effects of spatial frequency and orientation contrast, respectively.

Figure 4. The same data as depicted in Figure 2 are replotted to illustrate orientation tuning of masking at each SOA. Each plot shows the averaged and normalized visibility ratings ($R_{\text{norm}}$) as a function of spatial frequency ($X$-axis) and orientation contrast (separate lines). The more the lines are separated the more is masking affected by orientation contrast. It can be seen that orientation tuning is low at SOA = 0, increases up to about 40 to 50 ms, and then decreases again.

Downloaded From: http://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933480/ on 06/30/2018
the broad orientation information provided by only about one visible cycle.

Unreported so far is the effect that orientation contrast has on SOA$_{\text{max}}$. Our results showed that the SOA of maximal masking is increased with orientation contrast. This is also evident in the results from Ishikawa et al. (2006, p. 2133, Figure 2a) but is not discussed by the authors. Our data indicate that this relationship is constant for all spatial frequencies except $f = 0.5$ cpd, where there is only a slight increase. On average, the temporal difference between $\Delta \phi = 0^\circ$ and $\Delta \phi = 90^\circ$ is about 15 ms, which is in the same order of magnitude as the temporal difference in Ishikawa et al.’s data.

In general, the temporal shift of SOA$_{\text{max}}$ with increasing orientation contrast can be explained in two ways: either orientation contrast has an effect on the latency of some masking-related process or it affects the relative strength of at least two differently timed processes involved in masking. The former explanation would imply that the effect of orientation contrast is to speed up the response to the mask with increasing orientation contrast, as orientation contrast is by definition a property that can only emerge as a consequence of the presentation of the second stimulus, i.e., the mask. It has been shown that V1 cells show stronger but not earlier responses to oriented lines if they “pop-out” from an orthogonal oriented surround compared to a parallel oriented surround (Knierim & Van Essen, 1992). We conclude that it is unlikely that masking functions shift due to a negative relationship between response latency and orientation contrast.

The latter explanation assumes that two differently timed processes contribute to metacorrelation masking. Orientation contrast may affect the relative strength of these components. As shown in Figure 4 and also by Ishikawa et al. (2006) masking functions exhibit a high orientation specificity at lower SOAs as compared to higher SOAs, indicating that only the earlier process is substantially attenuated by orientation contrast. For that reason, the attenuation of masking strength produced by orientation contrast at the shorter SOAs results in an apparent shift of the optimal masking SOA to higher values as orientation contrast increases.

We did not find any effect of spatial frequency on SOA$_{\text{max}}$. This contradicts the results from Rogowitz (1983), who found non-systematic changes of SOA$_{\text{max}}$ with $f$ when target and mask shared the same $f$. On the other hand, we found a systematic effect of $f$ on the mask’s effectiveness in the form of a progressive decline of effectiveness with increasing $f$. In this case, the data of Rogowitz (1983) also showed some variation of mask effectiveness but also in a non-systematic fashion. The fact that spatial frequency affects the latency of neural response to gratings (Breitmeyer, 1975; Lupp et al., 1976; Vassilev & Mitov, 1976) is well supported by electrophysiological findings (Jones & Keck, 1978). However, it is an open issue whether spatial frequency affects the latency of sustained and transient responses equally, or as assumed in Breitmeyer and Ganz (1976), the latency of the sustained response more than that of the transient response. If the latter were the case, we would have expected an increase of SOA$_{\text{max}}$ with increasing $f$ because the sustained response to the target component would be increasingly delayed with respect to the transient response to the mask.

### Experiment 2: Effect of spatial frequency contrast

We did not observe such a shift in Experiment 1, where the spatial frequency shared by target and mask was varied. Two explanations may account for this failure to observe such a shift of SOA$_{\text{max}}$. One possibility is that there is no effect of variations of spatial frequency on the latency of any signal component relevant to masking. A second possibility is that variations of spatial frequency shared by the target and mask produce equal shifts of the latency of both the response to the mask responsible for suppressing the target visibility and the response to the target responsible for its visibility. In either case, SOA$_{\text{max}}$ would not change. In the present experiment, we manipulated the spatial frequency contrast by varying the spatial frequency of the target and mask independently of each other. Given that increases of spatial frequency produce increases in response latency (Breitmeyer, 1975; Lupp et al., 1976; Vassilev & Mitov, 1976), one might expect that introducing differences of spatial frequency between the target and mask may also vary SOA$_{\text{max}}$.

### Subjects

Nine subjects (5 females) took part in the experiment. All had normal or corrected-to-normal vision and no history of neurological or psychiatric diseases. Age ranged from 25 to 29, mean = 27.22 years. The subjects gave their informed consent and were paid 9 € per hour. All procedures were carried out according to the declaration of Helsinki and were approved by the ethical committee of the medical faculty of the University of Muenster.

### Apparatus and stimuli

The same apparatus and stimuli as in the first experiment were used. Stimuli were presented at orientations $\phi = 0, 30, 60, 90, 120,$ and $150^\circ$.

The orientation contrast between targets and masks was always $0^\circ$. Spatial frequencies of targets $f_T$ and masks $f_M$ could be any combination of 0.5, 1, 4, and 8 cpd. Figure 5 shows all possible combinations.
Procedure

The same procedure as in Experiment 1 was used, with the exception that SOAs could be 0, 10, 30, 50, 60, 70, 80, 100, 120, or 150 ms and measurements were repeated 20 times per condition. The total experiment duration was about 5 h per subject.

Results

Similar to Experiment 1, the data of one subject could not be fitted by the masking function. The subject was therefore excluded from analysis. One further subject was excluded because his/her rating data showed extremely high ratings (≈2.7 standard deviations above the mean) at short SOAs, whereas all other subjects’ data showed rather low inter-subject variability. Figure 6 shows the averaged normalized visibility ratings of the remaining seven subjects. In all conditions, type B masking functions can be observed.

As in the first experiment, SOA_{max} is defined as the temporal position of the minimum of the fitted function. Again, the effectiveness of masking, E, is defined as the difference between \( a \) and \( R(t)_{\text{norm}} \) at the location of SOA_{max}.

ANOVA results are summarized in Table 2. Mean values are depicted in Figure 7. As can be seen from
Table 2, there were significant main effects on SOA$_{\text{max}}$ from $f_T$ and $f_M$ as well as a significant interaction of $f_T/C_2$ $f_M$. (see Figures 7a, 7c, and 7e, respectively). For $E$, we found a significant main effect of $f_M$, as well as a significant interaction of $f_T/C_2$ $f_M$ (see Figures 7b and 7f, respectively).

**Discussion**

The analysis of SOA$_{\text{max}}$ and mask effectiveness yielded a dichotomous pattern for high and low spatial frequencies. With low-frequency targets, there does not seem to be a systematic relationship between mask frequency and SOA$_{\text{max}}$ (see the results for $f_T = 0.5$ cpd in Figure 7e). However, with high-frequency targets, it can be clearly seen that low-frequency masks produce later masking maxima than high-frequency targets (see the results for $f_T = 8$ cpd in Figure 7e). A converse pattern can be observed for mask effectiveness: with low-frequency targets (see the results for $f_T = 0.5$ cpd in Figure 7f), there are substantial differences in effectiveness between low- and high-frequency masks, with low-frequency masks producing almost full masking and high-frequency masks producing almost no masking. On the other hand, with high-frequency targets (see the results for $f_T = 8$ cpd in Figure 7f), we observe roughly the same intermediate masking strength with all mask frequencies.

We can further see in Figure 7f that a mask with a given spatial frequency is always most effective when the target is identical in frequency. If we translate high and low spatial frequencies (HSFs and LSFs) as preferential activators of sustained and transient channels, respectively, and refer to the terminology of inter- and intra-channels

<table>
<thead>
<tr>
<th>Measure</th>
<th>Effect</th>
<th>MSS</th>
<th>df</th>
<th>dfe</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA$_{\text{max}}$</td>
<td>$f_T$</td>
<td>1966.84</td>
<td>3</td>
<td>21</td>
<td>9.888</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>$f_M$</td>
<td>2556.85</td>
<td>3</td>
<td>21</td>
<td>5.537</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td></td>
<td>$f_T/C_2$ $f_M$</td>
<td>1223.41</td>
<td>9</td>
<td>63</td>
<td>4.257</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>$E$</td>
<td>$f_T$</td>
<td>0.010</td>
<td>3</td>
<td>21</td>
<td>0.226</td>
<td>0.877</td>
</tr>
<tr>
<td></td>
<td>$f_M$</td>
<td>1.760</td>
<td>3</td>
<td>21</td>
<td>18.477</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>$f_T/C_2$ $f_M$</td>
<td>0.374</td>
<td>9</td>
<td>63</td>
<td>10.979</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 2. Summary of 4 (target spatial frequency) × 4 (mask spatial frequency) ANOVA with repeated measures in Experiment 2 on the mean masking latency (SOA$_{\text{max}}$) and masking effectiveness ($E$). The symbols $f_T$ and $f_M$ denote the main effects of target and mask spatial frequencies, respectively.

Figure 7. (Left column) Means in SOA$_{\text{max}}$ and (right column) masking effectiveness ($E$) corresponding to the main effects for (top row) target spatial frequency, (middle row) mask spatial frequency, and (bottom row) the interactions of target and mask spatial frequencies.
inhibition (Breitmeyer, 1984; Breitmeyer & Ganz, 1976; Breitmeyer & Öğmen, 2006), we can summarize the underlying processes involved in metacontrast masking as follows: intra-channel inhibition is very strong and acts early in the LSF transient channel and is moderately strong and comparably early in the HSF sustained channel. In contrast, inter-channel inhibition is not observed in the form of HSF sustained-on-LSF transient channels; however, LSF transient-on-HSF sustained inhibition is moderately strong but about 30 ms later than the intra-channel inhibition (see and compare at $f_T = 8$ cpd in Figure 6e the SOA max values for $f_M = 0.5$ or 1 cpd vs. $f_M = 8$ cpd). This difference of 30 ms accords well with similar differences between reaction times of about 205 ms to the onset of a grating of 0.5 or 1 cpd and about 240 ms to the onset of a grating of 7 cpd reported by Breitmeyer (1975).

These results show that SOA max and mask effectiveness do not necessarily correlate, as was observed in the first experiment (compare Figures 4c and 4d where SOA max and $E$ appear to be negatively correlated, to Figures 7c and 7d where there appears to be no correlation between SOA max and $E$). Furthermore, they substantiate the assumption of two spatiotemporally different channels in the visual system on each of which stimuli with low spatial frequencies of 0.5 and 1 cpd have different effects than stimuli with higher spatial frequencies of 4 and 8 cpd.

The results of the second experiment strengthen the interpretation offered in Experiment 1 that the SOA max shift by increasing orientation contrast is due to a decrease in suppression within either transient or sustained channels while leaving the strength of transient-on-sustained suppression unaffected. This implies that between-channel masking is not orientation selective. This hypothesis is tested in the third experiment.

### Experiment 3: Effect of orientation contrast with low spatial frequency masks and high spatial frequency targets

#### Subjects

One female and two male subjects took part in the experiment. All had normal or corrected-to-normal vision and no history of neurological or psychiatric diseases. Age ranged from 20 to 31 years. All three subjects were naive with respect to the hypothesis of the experiment. Subject PH was highly trained from participating in similar metacontrast experiments. All procedures were carried out according to the declaration of Helsinki and were approved by the ethical committee of the medical faculty of the University of Muenster.

#### Apparatus and stimuli

The same apparatus and stimuli as in the first two experiments were used. Stimuli were again presented at orientations $\phi = 0$, 30, 60, 90, 120, and 150°. The orientation contrast between targets and masks was always $\Delta \phi = 0°$, 45°, or 90°. The targets’ spatial frequency was fixed at $f_T = 8$ cpd and the mask’s at $f_M = 1$ cpd.

#### Procedure

The same procedure as in Experiments 1 and 2 was used. SOAs could be 0, 10, 30, 50, 60, 70, 80, 100, 120, or 150 ms and measurements were repeated in randomized fashion 50 times per condition.

#### Results

As can be seen from Figure 8, all three subjects show U-shaped masking functions, although subject PH’s functions did not yield a very pronounced U-shape. In all cases, the data indicate that at SOAs between 0 and 40 ms the target was rated as more visible as an unmasked target. This replicates a part of the results from Experiment 2 (see light blue squares in Figure 6d). The enhancement of target visibility at the shortest SOAs may be related to similar findings at short SOAs observed in paracontrast masking (where the onset of target follows that of the mask; Breitmeyer et al., 2006; Kafaligönül, Breitmeyer, & Öğmen, 2009). Since Gabor patches, defined by both luminance increments and decrements relative to an intermediate-luminance background, were used in the present experiment, it is possible that the paracontrast target-enhancement effects that shift toward a paracontrast SOA of 0 ms (Kafaligönül et al., 2009) also carry over into the lowest metacontrast SOAs.

All SOA max values were equal to 80 ms, except for subject PH at $\Delta \phi = 45°$ (SOA max = 70 ms) and $\Delta \phi = 90°$ (SOA max = 90 ms). Most importantly, however, none of the three subjects’ data indicate an effect of orientation contrast on mask effectiveness.

#### General discussion

The first experiment showed an effect of target–mask similarity, here defined as orientation similarity, on the effectiveness of masking. Additionally, it was shown that increasing orientation contrast increased SOA max. Thus, SOA max and masking effectiveness appeared as two correlated consequences of target–mask similarity.

One could try to explain the SOA max shift with increasing orientation contrast by invoking a simple
orientation-specific mechanism: if we assume that masking is the result of neural inhibition among orientation-selective cell populations, such as V1 and V2 cells, it follows from their cortical organization (Hubel & Wiesel, 1962) that cells activated by similarly oriented targets and masks are in closer cortical proximity than cells activated by orthogonally oriented targets and masks. Is it therefore possible that the SOA max shift reflects the cortical distance traveled by inhibitory signals? We assume that this is unlikely based on propagation speed estimates in V1 and V2 obtained from a filling-in paradigm (Davey, Maddess, & Srinivasan, 1998). Speed estimates in V1 and V2 are 155 and 205 mm/s (+20), respectively. A very similar estimate (224 mm/s) was also obtained in a binocular rivalry study (Wilson, Blake, & Lee, 2001). Given the temporal delay of 15 ms between orientation contrasts of 0° and 90°, we would expect that these cells are between 2.025 and 2.625 mm apart in V1 or between 2.775 and 3.375 mm apart in V2. Hubel and Wiesel’s data show that orientation changes by 400° every millimeter. That equals 0.225 mm for 90°, which is far below the above given estimates.

We assume that the SOA max shift is better explained by the relative contributions of two differently timed masking processes: early within-channel masking and late between-channel masking. Within-channel masking is strong when targets and masks share critical features such as spatial frequency and orientation. Orthogonal stimuli that are identical in spatial frequency are mainly masked by the faster acting between-channel suppression, in particular, transient-on-sustained suppression, peaking at an SOA around 80 ms (SOA max = 80 ms). Parallel and near parallel stimuli (see results of the 0° and 90° orientation contrasts in Figure 2) produce significantly stronger masking at the shorter SOAs where within-channel masking dominates. The increased contribution of this short-SOA masking shifts the SOA max to around 50 ms. Such a process would be consistent with the masking effects existing within a single channel, as proposed by (Macknik & Livingstone, 1998; Macknik & Martinez-Conde, 2007).

In Experiment 2, we showed that target–mask similarity can be extended to spatial frequency similarity. However, this general similarity effect is strongly modulated by the absolute frequency of the target and mask: the data show that the strongest masking is observed with low-frequency masks and high-frequency targets. As noted in the Introduction section, a general spatial frequency approach to masking could explain these results on the reasonably based assumptions that the LSF channels respond more quickly than HSF channels (Breitmeyer, 1975; Lupp et al., 1976; Vassilev & Mitov, 1976) and that different spatial frequency channels can inhibit each other (Sagi & Hochstein, 1985; Tolhurst, 1972). More specifically, given that LSF stimuli trigger preferably transient channels and HSF stimuli trigger sustained channels the results confirm that the strongest source to metacontrast masking at longer SOAs is inter-channel inhibition, especially in the form of transient-on-sustained masking (Breitmeyer, 1984; Breitmeyer & Ganz, 1976; Breitmeyer & Ögmen, 2006).

Finally, the results of Experiment 3 (illustrated in Figure 8) show that between-channel (LSF-on-HSF channel or transient-on-sustained channel) masking can be realized without contributions from the more feature-specific within-channel masking as is evident from the lack of masking effects that are indifferent to orientation contrast.

Figure 8. Individual masking data of three subjects. The target spatial frequency was always 8 cpd; the mask spatial frequency was always 1 cpd. Three orientation contrasts were used: 0° (green lines and symbols), 45° (gray lines and symbols), and 90° (red lines and symbols).
These results bear not only on models of metacontrast masking but also on prior findings of surround suppression in which the visibility of a central grating is modulated by a simultaneous presentation of a sinusoidal grating annulus surrounding the target (Petrov, Carandini, & McKee, 2005; Petrov & McKee, 2009). Several similarities between our findings and those reported for surround suppression exist. First, the orientation contrast effects found in Experiment 1 were also found for the surrounding-suppression effects reported by Petrov et al. (2005). Second, the results of Experiment 2 (illustrated in Figure 6) indicate that masking magnitude was typically stronger when the spatial frequency of the mask matched or was close to that of the target. Petrov et al. (2005) also reported that surround suppression was strongest when the spatial frequency of the surrounding mask matched or was close to that of the central target. It is therefore possible that the metacontrast- and surround-suppression effects share common underlying neural processes.

However, despite these similarities, Petrov and McKee’s (2009) study showed that surround suppression was strongest at target–surround onset simultaneity, whereas metacontrast suppression occurs when SOA$_{max} > 0$ ms, i.e., when the onset of the mask lags that of the target by several tens of milliseconds. Moreover, Petrov et al. (2005) found that surround suppression was absent when the stimuli were centered at the fovea. In contrast, although metacortex is generally weaker in the fovea than in the perifovea or periphery, it is found to occur in the fovea (Bridgeman & Leff, 1979; Lyon, Matteson, & Marx, 1981; Saunders, 1977). For these reasons—and for additional ones discussed more extensively by Breitmeyer and Ögmen (2006)—we opt for the transient-on-sustained suppression mechanism rather than surround-suppression mechanism to explain metacontrast. Moreover, for the following reasons, we also opt for the specific transient-on-sustained suppression mechanism rather than for a spatial frequency-specific inhibition mechanism. In Experiment 1, it was found that the mask’s effectiveness decreases as the spatial frequency shared by the target and mask stimuli increases. Given, as the transient-sustained channel approach assumes, that transient channels prefer low spatial frequencies and sustained channels prefer high spatial frequencies (Breitmeyer & Ögmen, 2006), this would be consistent with either a weaker activation of LSF transient channels or a stronger activation of HSF sustained channels as spatial frequency increases. Either or both of these would result in an overall weaker transient-on-sustained inhibition and thus in a weaker mask effectiveness as spatial frequency increases. Currently, we know of no other plausible reason why weaker metacontrast masking as the spatial frequency shared by the target and mask increases ought to be found.

Although the reported experiments are not meant to dissociate feedback from feedforward models, we interpret the temporal difference between within- and between-channel masking as caused by the relative contributions of small local feedback loops and larger feedback loops across different brain regions. The small feedback loops may take place within orientation-selective visual areas such as V1 and V2 while the large cross regional feedback loops may also incorporate attentional mechanisms in larger networks.

Acknowledgments

This work was funded by the German Research Foundation (DFG), BR3832/1-1.

Commercial relationships: none.
Corresponding author: Maximilian Bruchmann.
Email: Maximilian.Bruchmann@uni-muenster.de.
Address: Malmedyweg 15, Muenster 48149, Germany.

References


Petrov, Y., & McKee, S. P. (2009). The time course of contrast masking reveals two distinct mechanisms...


