Article

Author(s)

Lakha, L
Wright, M J

Title

Capacity limitations of visual memory in two-interval comparison of Gabor arrays

Original Citation

Vision Research, Volume 44, Issue 14, June 2004, Pages 1707-1716

Capacity limitations of visual memory in two-interval comparison of Gabor arrays

Louise Lakha & Michael J. Wright

Department of Human Sciences, Brunel University, Uxbridge, UB8 3PH.

Abstract

The capacity of short-term visual memory (VSTM) was assessed in a two-interval spatial frequency (SF) discrimination task. The cued Gabor target in a multi-element array either increased or decreased in SF across a 2s interstimulus interval (ISI). Distracters as well as target were made to change across ISI so that memory of the individual SF of Gabor elements was required to solve the discrimination. The dynamics of the information loss from visual memory were analysed by manipulating the timing of spatial cues and masks. Cueing the target position before the first display gave thresholds comparable with those for a single Gabor patch. Cues placed after the first display gave higher thresholds indicating some loss of information. Within the ISI there was little increase in threshold or set size effect with cue delay. However there was a sharp rise in thresholds for cue positions after the second display. Gabor masks placed before a mid-ISI cue were more effective than noise masks or Gabor masks placed after the cue. With a cue placed late in the ISI, preceded by a Gabor mask, the masking effect decreased with increasing delay of the mask after the first display. This suggests a selective, dynamic but increasingly durable representation of the initial stimulus is built up in memory, and there is a graded form of "overwriting" of this representation by new stimuli.

Keywords: attention, psychophysics, visual memory, change blindness, spatial frequency.

1. Introduction.

There are limits to the ability to make visual comparisons between stimuli. This paper provides psychophysical evidence for successive limiting factors, the strongest being the overwriting of previous information in visual short-term memory (VSTM). We shall first of all review different explanations for performance limitations and set-size effects in experiments involving visual comparisons over time.

The failure to detect a changed target amongst multiple distracter stimuli is an example of "change blindness" (CB). More generally, CB can be described as a failure to detect changes in pictures, scenes or other visual stimuli, when local apparent motion cues (resulting from the change) are removed or are masked. CB phenomena occur not only in complex scenes, but also with arrays of simple stimuli such as Gabor patches (Scott-Brown and Orbach, 1998; Scott-Brown, Baker and Orbach, 2000; Wright, Green and Baker, 2000; Wright, Alston and Popple, 2001). These show increasingly large thresholds as the set size increases, and it has been argued that the set size slopes are too large to be accounted for by signal detection theory (SDT). Consistently with SDT, part of the large set size effect in these experiments may be due to distracter heterogeneity (Palmer,Verghese and Pavel, 2000). However, additional, limited capacity effects are evident where there is a memory component to the task (Scott-Brown, et al., 2000; Wright, et al., 2000; 2001), and this is the main focus of the present study.

Models of visual memory generally assume a distinction between a large capacity, rapidly decaying representation vulnerable to masking (iconic memory) and a relatively long-lasting, limited capacity, visual short term memory (VSTM) system. Gegenfurtner and Sperling (1993) investigated the links between iconic memory and more durable VSTM. They carried out a partial report task based on displays of letters, with cues and masks presented at various time delays following display onset. Early cues allowed efficient transfer from iconic memory to VSTM but performance dropped with increased cue delay. Masking before the cue decreased transfer. It is assumed that transfer prior to cueing is non-selective, since performance eventually reached an asymptote with increasing cue delay, which implied that some kind of capacity limitation was reached.

Becker, et al. (2000) looked specifically at the role of iconic memory in CB tasks using letter arrays. They found little evidence of the use of iconic representations even with ISI's as short as 82ms. One possible explanation is that "overwriting" by the second display prevented any use of an iconic representation, thus information had to be transferred into VSTM in order to avoid overwriting. This was confirmed when they placed a cue shortly after the offset of the first display, because the longer the delay before the second display, the greater the accuracy of detection and identification of the change. The delay is thought to allow fuller transfer from iconic memory into VSTM. A proportion of the CB effect may also result from interference between stimuli in VSTM (Hole, 1996; Tatler, 2001; Alston and Wright, 2002). We propose in this paper that VSTM, like iconic memory, may suffer from capacity limitations (e.g. overwriting) in simultaneous representations across time.

The timing of cues and masks can provide information on the encoding, storage and retrieval of items in iconic memory and VSTM. Our experiments are designed to analyse the stages where information is lost. Capacity limits are measured as SF discrimination thresholds and set size effects. Both cueing and masking have been used to determine the type of representation available at different points in time during discrimination. The results will be interpreted in terms of the operation of VSTM and the overwriting hypothesis.

2. General methods

2.1. Participants

The authors of the paper (LA, MW) were the principal participants on all experiments but data were also obtained from at least one of four naive observers (ST, AH, AM, SM). All had normal or corrected to normal acuity.

2.2. Apparatus

The stimuli were generated using a VSG 2/3 visual stimulus generator (Cambridge Research Systems, U.K and presented on an Eizo T662T Flexiscan display monitor using a frame rate of 100Hz, and calibrated to provide gamma correction.

2.3. Stimuli and Procedure

Gabor patches were generated by multiplying a sine-wave luminance grating by a circular Gaussian function. The standard deviation of the Gaussian envelope of the Gabor patches was 0.45 deg (1 to 4 Gabor array) or 0.25 deg (1 to 8 Gabor array), and the maximum contrast was 0.9. All stimuli were located on a circle of 1 deg radius centred on a fixation cross. There were 2 stimulus displays of 100-250 ms, with a 2000 ms inter-stimulus interval (ISI), set at mean luminance. We may define a stimulus element as a pair of Gabor patches appearing successively in a given position. The target element was always identified by a visual position cue. All visual cues were 0.03 deg black dots, 0.1 – 0.25 deg from the fixation cross and offset in the direction of the cued target and. The cue duration was 200 ms. Orientation of the grating was randomly set to 0, 45, 90 or 135 deg, and the phase was randomly varied between 0 and 90 deg. An example stimulus sequence (from Experiment 1) is shown in Figure 1. A 2AFC design was employed, in which the task was to indicate if the SF of the target element increased or decreased from stimulus display 1 to stimulus display 2. The probability of each type of change was 50%. Each data point was based on 4 – 8 blocks of trials with 54 trials in each block (9 constant stimuli x 6 repetitions). Cumulative Gaussian functions were fitted to block data and mean data to derive overall threshold means and standard errors. The threshold was the SF difference for the target element (between the first and second display) required to give correct detection of a change on 75% of trials. Error bars are ± 1 s.e.m. Some thresholds are normalised to a 4 c/deg baseline and expressed as log Weber fractions in order to facilitate comparisons with published data.

Figure 1. Near Here

2.4. Explanation of "all-change" design.

In the "all change" design (Wright, et al. 2000), the distracters change as well as the target, such that the target is only identifiable by cueing. The purpose is to ensure that the SF discrimination depends on a local comparison between corresponding elements in the first and second display, rather than a global SF comparison, or a criterion-setting effect (Lages and Treisman, 1998). SF difference of the cued stimulus element (between the first and second interval) was set on a given trial to one of 9 constant stimulus values. This target SF increment was divided equally between the first and second display thus if the increment on the first display was $+\delta f/2$, that on the second display was $-\delta f/2$. The increment was superimposed on a baseline spatial frequency, F, of 4 c/deg.. To ensure that the discrimination could be solved only by inter-display comparison and not by within-display judgements, a second, random SF increment, εf was added to both the target and the distracters. This increment, εf, was the same for the target element in both displays, whereas εf differed randomly for all the distracter elements between display 1 and display 2. Thus, the target SF in the first display was $F + \varepsilon_1 f + \delta f/2$, and that in the second display was $F + \varepsilon_1 f - \delta f/2$. The distracter SF's in the first display were $F + \varepsilon_{[2\ldots4]} f$ and in the second display they were $F + \varepsilon_{[6\ldots8]} f$. Whilst the orientation remained constant from the first stimulus display to the second, both the spatial phase and SF of distracters varied between displays. To further ensure that the threshold was based upon the comparison of display 1 and display 2, rather than the comparison of display 2 with an implicit standard, extreme final SF values were removed by adjusting the SF increment value εf by the following rule:

$$\text{If } (\varepsilon_1 f + \delta f/2 > \max \delta f) \text{ or } (\varepsilon_1 f + \delta f/2 < \min \delta f) \text{ then } \varepsilon_1 f := -\varepsilon_1 f$$

3. Results.

3.1. Experiment 1: Set size effects for spatial frequency (SF) discrimination in "all change" displays

The purpose of experiment 1 was to measure capacity limits in two-frame SF discrimination of a cued target in multi-element Gabor displays. Set size effects were measured for cues placed before the first display, within the ISI or after the second display.

3.1.1. Experiment 1: Methods

The set size for experiment 1 was varied between 1 and 8 by varying the number of 0.25 deg Gabor elements (actual set size). The target on which the judgement was to be based was selected in each case by means of a 200 ms visual cue. The cue came before the first display, during the ISI, either at 300 ms or 1700 ms, or after the second display.

3.1.2. Experiment 1: Results and discussion

For a target cued before the first display (figure 2a,b), the curve was flat (Palmer, et al. 1993; Wright, et al., 2000). Thus pre-cueing allows selection of a single target with no interference from distracters. For cues placed within the ISI or after the second frame, thresholds rose with set size. When plotted on log-log co-ordinates, the data for set sizes 2-8 could be fitted reasonably well by a straight line. This allowed a comparison of set size effects and thresholds for two further cueing conditions: post-cueing after the second frame, and cueing in the ISI before the second (Figure 2a,b). Thresholds for the post-cued target are higher at all set sizes than those for the ISI-cued target. However, slopes are similar for targets cued before or after the second display.

Figure 2 a,b Near Here.

Figure 3 Near Here.

Data from three more observers was obtained for mid-ISI cues at 300msec and 1700msec after the first frame. The results show there is no absolute capacity limit in terms of a fixed number of items that can be processed, but rather there is a graded set-size effect. The absolute thresholds differed markedly for the three observers suggesting real individual differences in the memory task, but the set-size slopes were similar both for different observers and different cue timings. A two-way ANOVA on the individual thresholds gave only a main effect of set size ($F(3,33)=5.53$, $p<0.005$) and there was no significant interaction between timing and set size slope.

Set size slopes are steeper than predictions from signal detection theory (SDT) for identification of an odd target in an array (Palmer, Ames and Lindsey, 1993) both for cueing in the ISI and cueing after the second display. SDT is essentially an unlimited-capacity theory, and the unlimited-capacity predicted slope is around 0.3, but the observed log-log slopes (0.59 – 0.86) are more consistent with a well-known limited capacity search model (Palmer, 1994; Verghese and Nakayama, 1994) where a value of 0.75 is expected. Our results suggest strong limitations in the capacity to selectively compare even small numbers of stimuli across a time interval. Conversely, it was confirmed that cueing before the first display gives essentially flat set-size slopes (Fig 2a,b) so that the limitation is not perceptual (Palmer, et al. 1993). Taken together, this suggests that loss of information in CD occurs both (a) with the initial encoding of multiple stimuli and (b) with processing of the second display. In experiment 2, we attempt to locate more precisely in time this loss of information.

3.2. Experiment 2: effects of cue timing on the discrimination of spatial frequency (SF) change in "all change" displays

The purpose of Experiment 2 was to analyse the effect of cueing at varying times through the two-interval exposure. Set size was kept constant, unlike experiment 1, but a more detailed analysis of cue timing was carried out. In this way it was intended to trace the increase in threshold at varying stages of the display sequence.

3.2.1. Experiment 2: Methods

Data are combined for three experimental sessions. In Conditon 1), a four-Gabor display was used, with a single cue dot specifying the target element. The stimulus sequence consisted of two 150 ms displays separated by a 2000 ms ISI. Four different cue-times were used: (i) immediately before the first stimulus display (ii) immediately after the first stimulus display (iii) immediately before the second stimulus display (iv) immediately after the second stimulus display. Only one stimulus location was cued on any trial. Different cue delays were measured in different counterbalanced blocks of 54 trials. In condition 2) a four Gabor display was used, and the single cue was presented at different delays within the ISI. Condition 3) was the same as condition 2 but with eight Gabor patches.

3.2.2. Experiment 2: Results and discussion

Figure 4 shows the results for each of three observers. The low threshold for cueing before the first display (-350ms) is consistent with other known effects of pre-cueing on attention (Nakayama and Mackeben, 1989). Cueing immediately after the first stimulus (0ms) gives higher thresholds, but is nevertheless relatively efficient, and this suggests that a high-capacity iconic representation is available. After the 2000ms ISI and before the second display (2000ms), any iconic representation is likely to have faded, and the reduced effectiveness of cueing in condition 1 suggests a limited capacity VSTM. A consistent decline in performance in all subjects for cueing after the second display (2150ms) suggests interference or competition from the second display. Conditon 2) and 3) tested specifically whether information is lost over the duration of ISI, as would be expected from a fading iconic memory. The results showed a very shallow gradient of threshold change within the ISI. ANOVA failed to show a significant effect of cue timing. There is similar performance for all ISI values, suggesting that the information is already in VSTM rather than in a decaying iconic representation, and that the capacity of VSTM extends to multiple items. The gradient across ISI was however flatter in condition 2) than condition 1), and this may be due to learning effects (condition 2 was measured after condition 1 for all participants). Observers MW and ST showed improvements relative to condition 1) for late cueing in ISI only, whereas LA showed overall improvements in thresholds but most marked for late in the ISI.

Figure 4 a,b,c Near Here.

Our finding of cueing benefits late into the ISI refutes the idea of CB as due to limited visual selection alone. The effectiveness of a cue during the ISI relative to a cue following the second display confirms that a relatively complete representation of the initial stimuli is present during the ISI. We find that there is little decay of SF in VSTM over 2000 ms ISI. Also, experiment 1 showed that cueing before the first display gives essentially flat set-size slopes. Extrapolating from these experiments to CB phenomena generally, “overwriting” effects in CB (Becker, et al., 2000), might be more significant quantitatively than the initial selection of a sparse representation of the visual input (O’Regan, 1992) or the decay of a representation within VSTM (Landman, Spekreijse & Lamme, 2003).

3.3. Experiment 3: Effects of masking before or after cue

Masks placed in the stimulus sequence were found to worsen SF discrimination. Cueing in the ISI could reduce the masking of the target representation by selectively strengthening the representation of the target patch. Hence a mask placed before the cue should have a more significant masking effect than a mask placed after the cue. By keeping cue position constant and placing the mask before or after the cue we can determine whether the representation of display 1 differs before and after the cue. Two types of mask were used: a noise mask and a Gabor mask, that is a mask consisting of Gabor patches similar in form to the display items. Both masks would be expected to affect an iconic representation. A Gabor mask may also disrupt the representation of SF in VSTM if it involves higher level processing consistent with the target stimulus. In two-interval SF discrimination experiments, Magnussen, Greenlee, Asplund and Dyrnes (1991) showed that a grating presented midway during a 10s ISI raises the SF discrimination threshold, independently of orientation but dependent on the relative SF of test and mask stimuli. Similar masking effects were found by Lalonde and Chaudhuri (2002) for single grating stimuli varying in orientation or SF. They placed masks before the start of 2AFC trials and found interference particularly when the mask was relevant to the task and then only on the same attribute: there was no masking effect when the mask was involved in an orientation judgement and the 2AFC task was for SF. The SF masking effect was eliminated when the target SF and mask SF matched, and masking increased with SF difference, unlike perceptual masking. The masking range in both studies was consistent with the bandwidth of SF channels. However, since masking occurs independent of other attributes (such as orientation) whereas psychophysical SF channels or V1 cells are tuned to multiple attributes, this suggests that interference occurs at a higher level of processing than these early SF channels (Magnussen, 2000). Magnussen & Greenlee (1997) found that thresholds were raised several fold when making simultaneous discriminations on the same attribute relative to simultaneous discriminations on separate attributes. However, performance did not depend on the relative spatial frequencies of the targets as with memory masking. Based on these findings, Magnussen and colleagues propose that VSTM may be operating via a parallel set of memory analysers tuned to single stimulus attributes. Retention of more than one stimulus leads to interference as resources are limited within each attribute store (Magnussen, 2000). We suggest that change detection is generally affected by resource limitations in short-term memory and that focused attention on a target is necessary to both reduce noise from other detectors and effectively allocate resources.

From this model of VSTM, we predict a susceptibility to masking during ISI particularly during distributed attention as opposed to focused attention on a target. This attentional modulation in Experiment 3 is produced by varying the relative timing of cues and masks.

3.5.1. Experiment 3: Methods

The displays all contained 4 Gabor patches. A target cue (200 ms) was always presented at 1000 ms after the first stimulus display – the mid-point of ISI. In addition to this, on two-thirds of trials, a mask was presented at 500 ms or 1500 ms after the end of the first display. The different mask conditions were tested in separate blocks. In condition (i) a noise mask was used, consisting of 100% contrast random noise with 50% black and white pixels each subtending approximately 1 arc min. It covered the whole screen area (8 deg x 5 deg) and had a duration of 100 ms. The mask in condition (ii) was similar in duration but consisted of Gabor patches identical in size and location to the stimuli themselves (but with orientations and spatial frequencies randomised). Conditions (i) and (ii) were tested separately with the noise mask condition completed before the Gabor mask.

3.5.2. Experiment 3: Results and discussion

The presentation of a noise mask produced little difference in performance between the three conditions (figure 5). Some effect of the noise mask was expected at 50 ms, but ANOVA showed no overall effect of noise mask timing on thresholds. The Gabor mask however was effective at 50 ms and continued to produce a significant masking effect at 500ms in all three observers - too long a delay to attribute to iconic masking (figure 5). This was confirmed by ANOVA on the individual thresholds for the conditions: no mask, late mask and early mask. The results indicated a significant effect of mask timing ($F(2,35)=9.37$, $p<0.01$). Although (according to observers' reports) it was easy to ignore the middle Gabor (mask) display and attend to the first and last Gabor displays only, the Gabor mask nevertheless had a strong influence on thresholds.

This "memory masking" effect may reflect interference between memory analysers coding different spatial frequencies at target locations. The effectiveness of the Gabor mask was slightly greater when presented before rather than after the cue. Cueing the target should allow the target signal to be stronger due to less distracter competition / noise and thereby reducing susceptibility to masking.

In other words, if the target signal is strong then the effects of interference due to "overwriting" will be minimal since there will be high residual target signal.

The lack of masking effects for similar-SF stimuli in the previous studies may have reflected the observer's approach to the task. In Lalonde and Chaudhuri's (2002) design, the target and mask had matching SF on a third of trials. Since the mask preceded the first target, the mask may have acted as a cue for the target SF range with varying cue validity, producing no interference when target and mask SF were similar. For studies using masking during ISI, the mask may enhance detector activity causing enhancement when target and mask have similar SF. Therefore the lack of SF-dependent interference for dual discriminations need not indicate differing levels of representation.

Figure 5 a,b,c Near Here.

To summarise, a noise mask should affect iconic representations only, whereas a Gabor mask is expected to show SF interference or "memory masking" effects (Magnussen, 2000). The implication of our results is that by 500 ms, the stimulus had already been coded in VSTM, which would not be disturbed by a noise mask. Overall, VSTM appears to be resistant to the effects of a noise mask, but less resistant to the effects of a Gabor mask, consistent with a "memory masking" effect.

3.6. Experiment 4: Effect of mask timing relative to the stimulus sequence

If a cued target is protected against masking, then the effect of a Gabor mask should be small. Provided there is adequate time between the cue and the mask, the representation of the target Gabor should be selectively strengthened in VSTM. In Experiment 4(i), the target was cued immediately following the first display. A Gabor mask was then presented at different times following the cue to determine the minimum time required to transfer information into a durable form of VSTM not susceptible to masking. A similar mask timing design was used in Experiment 4(ii) except the cue was presented immediately preceding the second display, at the end of the ISI, as opposed to at the beginning of the ISI in Experiment 4(i). The combination of a late cue with an earlier Gabor mask enabled the testing of the hypothesis that the representation is vulnerable when the target is not pre-cued and resources are stretched. Again, the Gabor mask was presented at varying time intervals during the ISI, allowing an estimate of the minimum time required to transfer information into a durable form of VSTM, not susceptible to masking. In Experiment 4(ii), all the stimuli needed to be transferred into VSTM before the mask, rather than the cued stimulus alone as was the case in 4(i). Performance is compared with a noise mask condition (Experiment 4(iii)). If a noise mask interferes with iconic memory alone, then unlike the conditions with the Gabor mask, there should be effects of the mask only when placed very early in ISI despite the cue appearing only at the end of the ISI.

3.6.1. Experiment 4: Methods

The displays contained 4 Gabor patches, as described for Experiment 3. A mask was presented at different possible timings during ISI: 0, 100, 500, 1000 and 1500 ms after first display.

This was compared with performance on a no mask condition. There were 3 separate cueing and mask conditions:

i) Gabor mask with an early cue. The cue was presented immediately following the first display. The Gabor mask was presented at delays of between 0-1500 ms relative to the end of the 200ms cue.

ii) Gabor mask with a late cue – The mask was presented at delays of between 0ms and 1500 ms after the first display. It was followed by a cue at 1700ms, which lasted for 200 ms, and then after a 100ms gap the second stimulus display appeared.

iii) Noise mask with a late cue - the display sequence was identical to (ii) with a mask between 0 and 1500 ms and a cue at 1700 ms.

3.6.2. Experiment 4: Results and discussion

The results for all three mask and cue conditions are given in Figure 6. Thresholds differed between the Gabor mask conditions with early and late cues (F(1,50)=25.6, p<0.001). These differences are suggestive of the Gabor mask interfering with the retention of multiple stimulus information in VSTM. When cueing was early, the Gabor mask had little effect unless it appeared in the first 100ms (Observers MW, LA). Cueing allows efficient and precise transfer of items into VSTM. In contrast, the thresholds obtained with an early Gabor mask and a late cue, were comparable to the thresholds obtained in experiment 1 and 2 for post-cued targets. In this respect, the effect of the Gabor mask is similar to the effect of the second display. These trends were evident in all three subjects' data, though observer ST seemed to be able to transfer information into a durable form more completely than MW or LA. Thresholds were higher in the late cue condition than for the early cue condition, because information about all four Gabors in the first display (or a sample of this information) would have to be transferred into VSTM. Early masks may disrupt even the initial selection of items into VSTM, particularly when the process is slowed down by larger set sizes. This transfer of information into VSTM may take several hundred ms (Observers MW and LA, Experiment 4(i)). The extent to which VSTM is disrupted by a mask decreases with mask delay, showing a transition from a transient to a more durable representation. This is either a transfer from iconic storage to VSTM or else it reveals a consolidation effect within VSTM.

Figure 6 Near Here.

The noise mask had a detrimental effect on performance for observers MW and LA, when it was presented very early (<100ms). The effect of the noise mask with a late cue was less than that of the Gabor mask with a late cue. This was evident from ANOVA showing differential effects of mask time on the two mask conditions. A two-way ANOVA on noise mask versus Gabor mask and mask timing gave a significant effect of mask time (F(4,40)=7.50, p<0.001), an almost significant effect of mask type (F(1,40)=3.93, p<0.075) and a significant interaction between mask time and mask type (F(4,40)=4.27, p<0.01). The noise mask interfered with only the iconic representation.

The results confirm that cueing a target allows selective transfer of information into durable VSTM (Gegenfurtner and Sperling, 1993). Non-selective transfer (before cueing) in trials with a Gabor mask is asymptotic over a relatively long time course (approximately 500ms) which is a longer interval than generally ascribed to iconic memory and suggests the strengthening of a representation in VSTM. Furthermore, cueing may allow not only the efficient transfer of items between iconic and short-term memory but also an improvement of the target representation in VSTM by reducing distracter noise / increasing target signal (Lu & Dosher, 1998). This would account for the improvement between 0 and 500ms for the post-cued Gabor condition despite the delay being too long for iconic masking. We suggest that VSTM is subject to an initial consolidation process that improves target signal over a time-span of several hundred milliseconds. We also propose that VSTM can hold multiple items but is subject to an overall resource limit leading to weaker signal with set size. After cueing, only one Gabor patch needs to be represented in VSTM whereas before cueing, information about four Gabor patches is needed. The processing resources required to represent four stimuli are distributed producing weaker stimulus representations and allowing greater interference by a SF mask. It is noteworthy that the effect of the Gabor mask was less disruptive to performance than was the second display in the post-cueing condition. The mere presence of a second stimulus does not erase the representation of the first stimulus, and the overwriting by the second display in the post-cue condition appears more complete. This reflects active processing of the display in order to identify the SF of the cued target, whereas the Gabor mask is seen but not actively processed. The limiting factor in comparing information between displays is the difficulty in retaining an initial representation; once new stimuli are presented, overwriting of earlier VSTM representations occurs.

4. General Discussion

Even with eight Gabor patches presented simultaneously, consistent thresholds were achieved by experienced observers for SF discrimination of a target cued after a 2000ms ISI and before the second display. The capacity of VSTM is often described as about four items (Luck and Vogel, 1997, Cowan, 2000, Lee and Chun, 2001, Landman et al. 2003). If this were strictly the case, we would not expect to see a gradual change in thresholds for display sizes between 2 and 8. Rather than a limit based on a fixed number of items, the accuracy with which the SF of each stimulus can be represented in VSTM decreases with set size.

For single stimuli, properties such as orientation (Magnussen, Landro and Johnsen, 1985), spatial offset (Fahle & Harris, 1992) and contrast (Lee and Harris, 1996) show gradual decay in VSTM. Properties such as SF (Regan, 1985, Magnussen et al.. 2000) or visual motion (Wright and Gurney, 1995, 1999) are retained (for single stimuli) with even longer persistence and with a precision equal to that of perceptual judgements. Thus it is unlikely that capacity limitation in VSTM consists simply of a temporal decay of neural activity.

Limited capacity effects in VSTM emerge strongly with multiple stimuli both for orientation (Landman, et al. 2003) and for SF (present results). Set size slopes in the storage phase of VSTM (0.35-0.7) were substantial. It may be that when representing multiple items, stimulus features need to be linked to their respective items – the ‘binding problem’. Treisman (1996) suggests that such linking of features to items uses a spatial map. It is only by allocating attention to specific positions in the spatial map that features can be combined into objects. This is supported by the greater difficulty of SF discrimination in “all change” Gabor displays, relative to displays in which there is a net change in SF across multiple stimuli (Wright, et al., 2001), in which a global SF discrimination is possible. Conversely, in change detection and CB experiments, a spatial representation must normally be maintained in order to determine which items are to be compared with which across time. This matching of features to locations is computationally demanding, producing limits on the number of objects that can be represented without error.

Magnussen (2000) proposed that VSTM consists of a set of memory analysers closely associated with the analysers used for perception. Whereas perceptual analysers are tuned to multiple stimulus dimensions, including both SF and orientation, it is argued that memory analysers are tuned to single stimulus dimensions, with little interference between those dimensions. For example, SF memory masking is not orientation dependent (Magnussen, et al. 1991). Memory analysers also show certain limited capacity effects, including a limit to the ability to carry out multiple discriminations. In simultaneous discrimination tasks with two gratings, Magnussen and Greenlee (1997) showed that thresholds were higher when the same dimension (SF or contrast) was discriminated in both stimuli, relative to conditions where SF discrimination was required for one grating and contrast discrimination for the other grating. In our experiments, although we measured discrimination on a single dimension (SF) we found a limitation in processing multiple stimuli.

We propose that the SF of stimuli in the first display (and other stimulus properties) is stored in VSTM in a more or less coarse coding at a certain level of activation of memory analysers. A Gabor mask, or a second stimulus display introduces activation in an overlapping population of analysers, reducing the discriminability of signals in VSTM arising from the first activation, and this is the principal limitation on performance in change discrimination. Interference between a current and a remembered stimulus may be thus be a consequence of the sharing of neural mechanisms (Jha, 2002). A spatial cue prior to the Gabor mask or second display prevents the SF analysers in VSTM for the cued location from becoming too low in activation to be discriminated from their neighbours. Thus the local SF information is preserved. Overwriting by a second display may be stronger than overwriting by a Gabor mask because attention is directed at the second display, and withdrawn from the Gabor mask, so the activations produced by all stimulus patches in the second display are greater. For complete dominance of the second stimulus to occur, the second activation must greatly exceed the first, and attention is a factor in modulating activation. In our experiments, we know that VSTM was well controlled and that cueing was the only method for locating the target. CB experiments generally require the changing stimulus to be identified or located, so the limited capacity of VSTM is involved. We propose that distributed attention (Shaw, 1980) and overwriting both contribute to CB. The strength with which each initial item is represented in VSTM depends on the number of items, and is reduced further by the processing of the second display. Stimuli are not replaced in VSTM; they just become much less discriminable. This is analogous to the pooling effect that occurs for “crowded” visual stimuli (Solomon and Morgan, 2001). A cue can insulate against interference from a mask as it allows the selective activation of the target's analyser to be increased within VSTM. The suggestion that there is wider pooling of activation where attention is distributed across several items, fits with the finding that there is no sudden breakdown of performance for large set sizes. Instead, thresholds gradually increase with set size. Thus although some information is lost on encoding into VSTM, the greatest information loss results from interference. These results are consistent with a graded interference between perceptual input and items in VSTM.

Acknowledgements

This work was made possible by a project grant from the Wellcome Trust.

References

Alston, L. & Wright, M.J. (2002). Overwriting of visual short-term memory (VSTM) in change blindness. *Perception* 31: 81-81 Suppl. S 2002

Becker, M. W., Pashler, H., & Anstis, S. M. (2000). The role of iconic memory in change detection tasks. *Perception*, 29, 273-286.

Cowan, N. (2000) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioural and Brain Sciences*, 24, 87-185.

Fahle, M. & Harris, J.P. (1992). Visual memory for vernier offsets, Vision Research, 32, 1033-1042

Gegenfurtner, K. R. & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology, Human Perception and Performance*, 19, 845-866.

Greenlee, M.W. & Magnussen, S.L. (1998). Limited capacity mechanisms of visual discrimination. Vision Research, 38, 375-385.

Hole, G. J. (1996). Decay and interference effects in visuospatial short-term memory. *Perception*, 25(1), 53-64.

Jha, A.P. (2002). Tracking the time-course of attentional involvement in spatial working memory: an event-related potiential investigation. *Cognitive Brain Research*, 15, 61-69.

Lages, M. & Treisman, M. (1998). Spatial frequency discrimination: visual long term memory or criterion setting? *Vision Research*, 38, 557-572.

Lalonde, J. & Caudhuri, A. (2002). Task-independent transfer of perceptual to memory representations during delayed spatial frequency discrimination. Vision Research, 42, 1759-1769.

Landman, R., Spekreijse, H. & Lamme, V.A.F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research,* 43, 149-164.

Lee, D. & Chun, M.M. (2001) What are the limits of visual short-term memory: objects or spatial locations? *Perception and Psychophysics*, 63, 253-257.

Lee, B. & Harris, J.P. (1996). Contrast Transfer Characteristics of Visual Short-term Memory, Vision Research, 36, 2159-2166.

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

Luck, J. L., & Vogel, E. K. (1997) The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279-281.

Magnussen, S. (2000). Low-level memory processes in vision. *Trends in Neurosciences*, 23, 247-251.

Magnussen, S., & Greenlee, M. W. (1997) Competition and sharing of processing resources in visual discrimination. *Journal of Experimental Psychology: Human Perception and Performance,* 23, 1603-1616.

Magnussen, S., Greenlee, M. W. Asplund, R., & Dyrnes, S. (1990) Perfect visual short-term memory for periodic patterns. *European Journal of Cognitive Psychology,* 2, 345-362.

Magnussen, S., Greenlee, M. W. Asplund, R., & Dyrnes, S. (1991) Stimulus-specific mechanisms of visual short-term memory. *Vision Research,* 31, 1213-1219.

Magnussen, S., Landro, N. I. & Johnsen, T. (1985). Visual half-field symmetry in orientation perception. Perception, 14, 265-274.

Nakayama, K. & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, 29, 1631-1647.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology,* 46, 461-488.

Palmer, J., Ames, C. T. & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance,* 19, 108-130.

Palmer, J. (1994). Set-size effects in visual search: the effect of attention is independent of the stimulus for simple tasks. *Vision Research,* 34, 13, 1703-1721.

Palmer, J. Verghese, P. & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, 40, 1227-1268.

Scott-Brown, K. C. Baker, M. R. & Orbach, H. S. (2000). Comparison blindness. *Visual Cognition,* 7, 253-267.

Scott-Brown, K. C. & Orbach, H. S. (1998). Contrast discrimination, non-uniform patterns and change blindness. *Proceedings of the Royal Society of London, B,* 265, 2159-2166.

Shaw, M. L. (1980) Identifying attentional and decision-making components in information processing. In R. S. Nickerson (Ed.), *Attention and Performance VIII, Vol 8, 277-295*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Solomon, J. A. & Morgan, M.J. (2001) Odd-men-out are poorly localised in brief exposures. *Journal of Vision*, 1, 9-17.

Tatler, B.W. (2001). Characterising the visual buffer: real-world evidence for overwriting early in each fixation. *Perception*, 30, 993-1006.

Treisman, A. (1996) The binding problem. *Current Opinion in Neurobiology*, 6(2), 171-178.

Verghese, P. & Nakayama, K. (1994). Stimulus discriminability in visual search. Vision Research, 34, 2453-2467.

Wright, M.J. & Gurney, K.N. (1995) The discrimination of dynamic orientation changes in gratings. *Perception*, 24, 665-679.

Wright, M.J. & Gurney, K.N. (1999) Visual discrimination of direction changes based on two types of angular motion. *Vision Research,* 39, 1927-1941

Wright, M. J., Green, A. and Baker, S. J. (2000). Limitations for change detection in multiple Gabor targets. *Visual Cognition* 7, 237-253.

Wright, M.J., Alston, L. & Popple, A.V. (2001) Set size effects for spatial frequency change and discrimination in multiple targets. *Spatial Vision*, 15, 157-170

Legends

Fig 1. Typical sequence of stimulus frames. The task is to determine the direction of spatial frequency change (increase or decrease) of the cued Gabor target. In this example, the set size is 8 and the cue is placed in the ISI 300 ms after the end of the first display.

Figure 2. Set size effects for SF discrimination of a cued Gabor target for different temporal locations of the cue. Ordinate: SF discrimination threshold expressed as log Weber fraction. Abscissa, log no. of Gabor patches. Straight lines are least squares linear fit to log data. (a) subject MW. (b) subject SM. Estimated set size slopes: MW precue 0.0, ISI cue 0.68, postcue 0.8; SM precue 0.0, ISI cue 0.54, postcue 0.58.

Figure 3. SF discrimination thresholds for Gabor arrays (ordinate) as a function of set size (abscissa). Comparison of the effects of cueing at 300 ms in the ISI with cueing at 1700 ms in the ISI. Data for three observers are shown. Straight lines are least squares linear fit to data plotted on double logarithmic axes. Estimated set size slopes: AM 300 ms: 0.70; 1700 ms, 0.45. ST 300 ms, 0.44, 1700 ms, 0.69. LA 300 ms, 0.65; 1700 ms, 0.52.

Figure 4. Spatial frequency discrimination thresholds (linear ordinate, 75% threshold in c/deg) as a function of cue timing (abscissa, onset of cue in ms relative to onset of first display). Condition 1 (black diamonds) was carried out first, using four 0.45 deg Gabor patches. Condition 2 (white squares) utilised four 0.45 deg Gabor patches and measured thresholds for different timings of the cue within the ISI and was carried out second. Condition 3 (white triangles) utilised eight 0.25 deg Gabor patches and was conducted last. (a) MW, (b) LA, (c) ST.

Figure 5. Thresholds for discrimination of a SF change in the cued target of a 4 Gabor array. The 200 ms cue occurred at the mid point of a 2000ms ISI and could be either preceded or followed by a mask. Two types of mask were used: a full-screen noise mask at 100% contrast (black diamonds), or a Gabor mask (grey squares). Observers: MW, LA, ST.

Figure 6. SF change discrimination varying mask timing during ISI. Black diamonds – Gabor mask with late cue (bold lines); Grey squares, Gabor mask with early cue (dashed lines); Open triangles – noise mask with late cue (dotted lines). Observers LA, MW and ST.

.

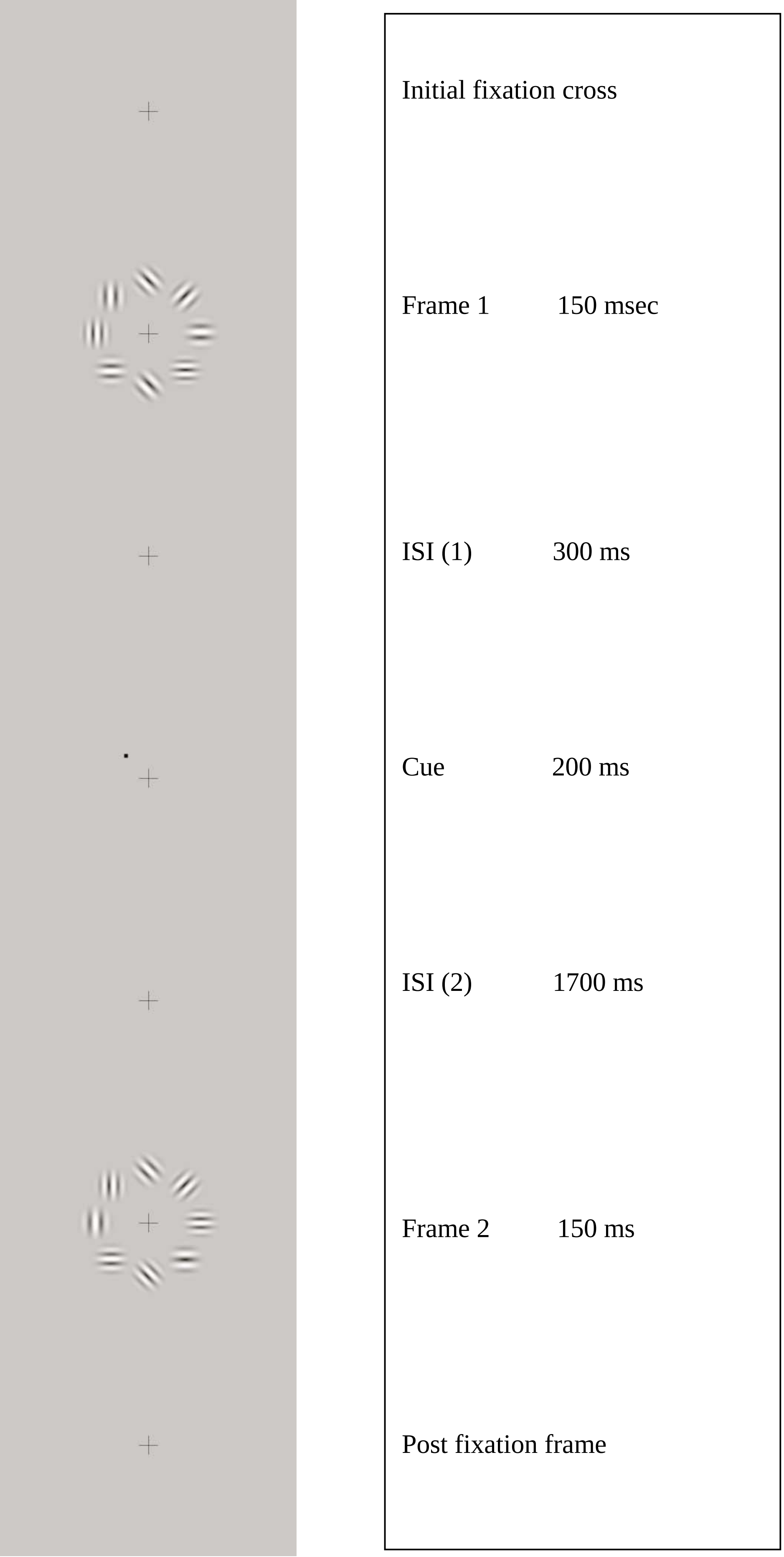
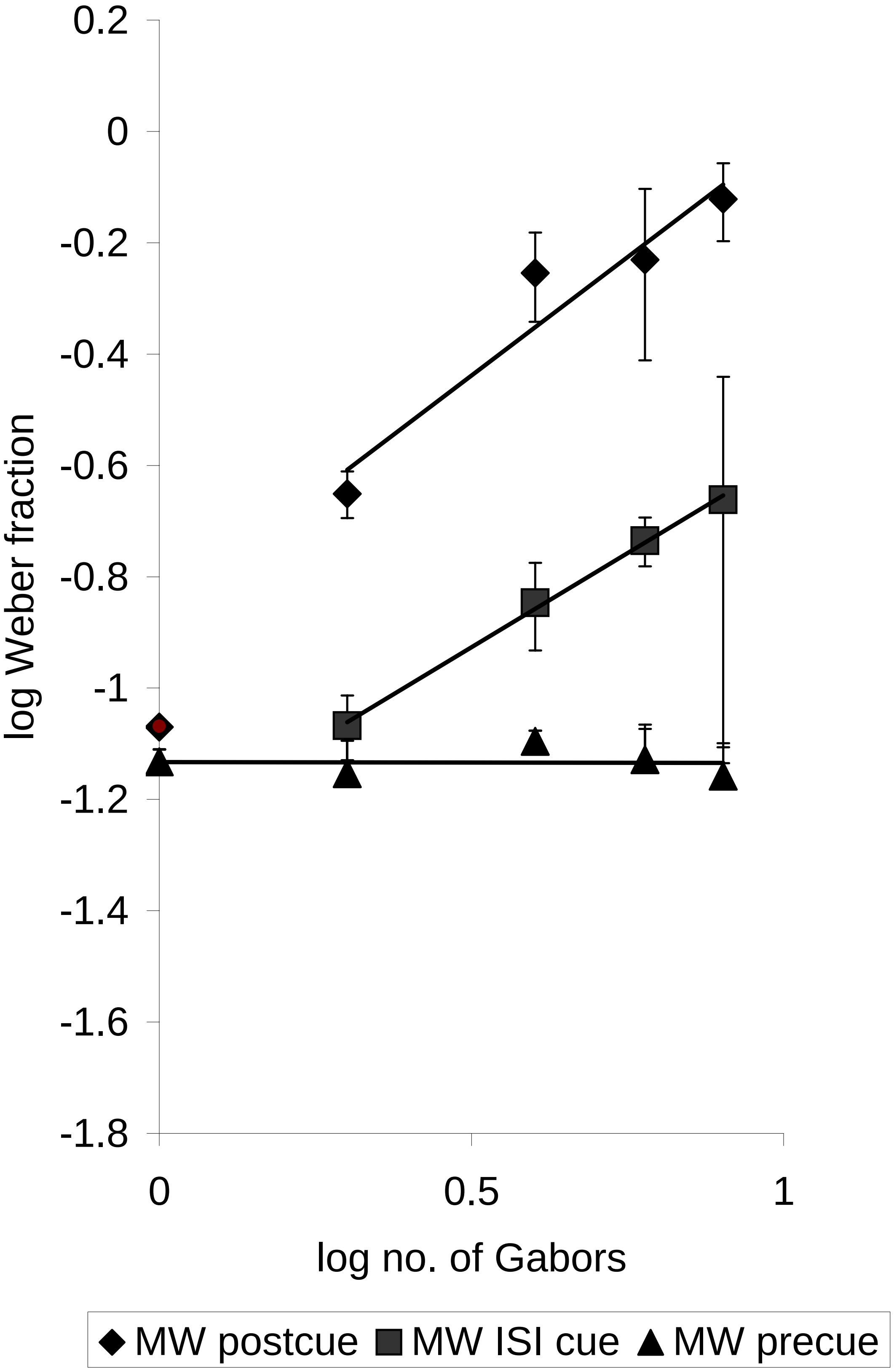
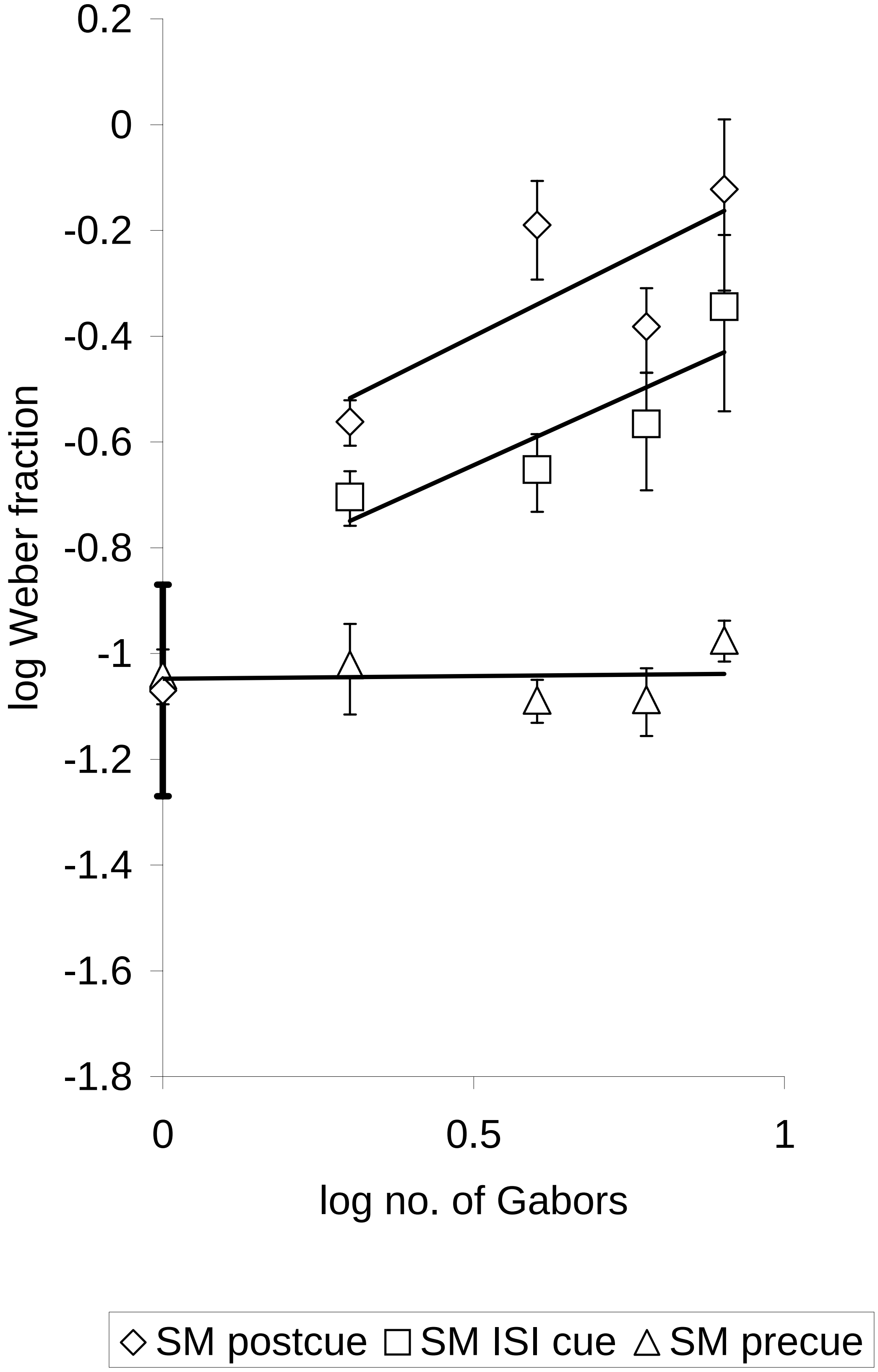
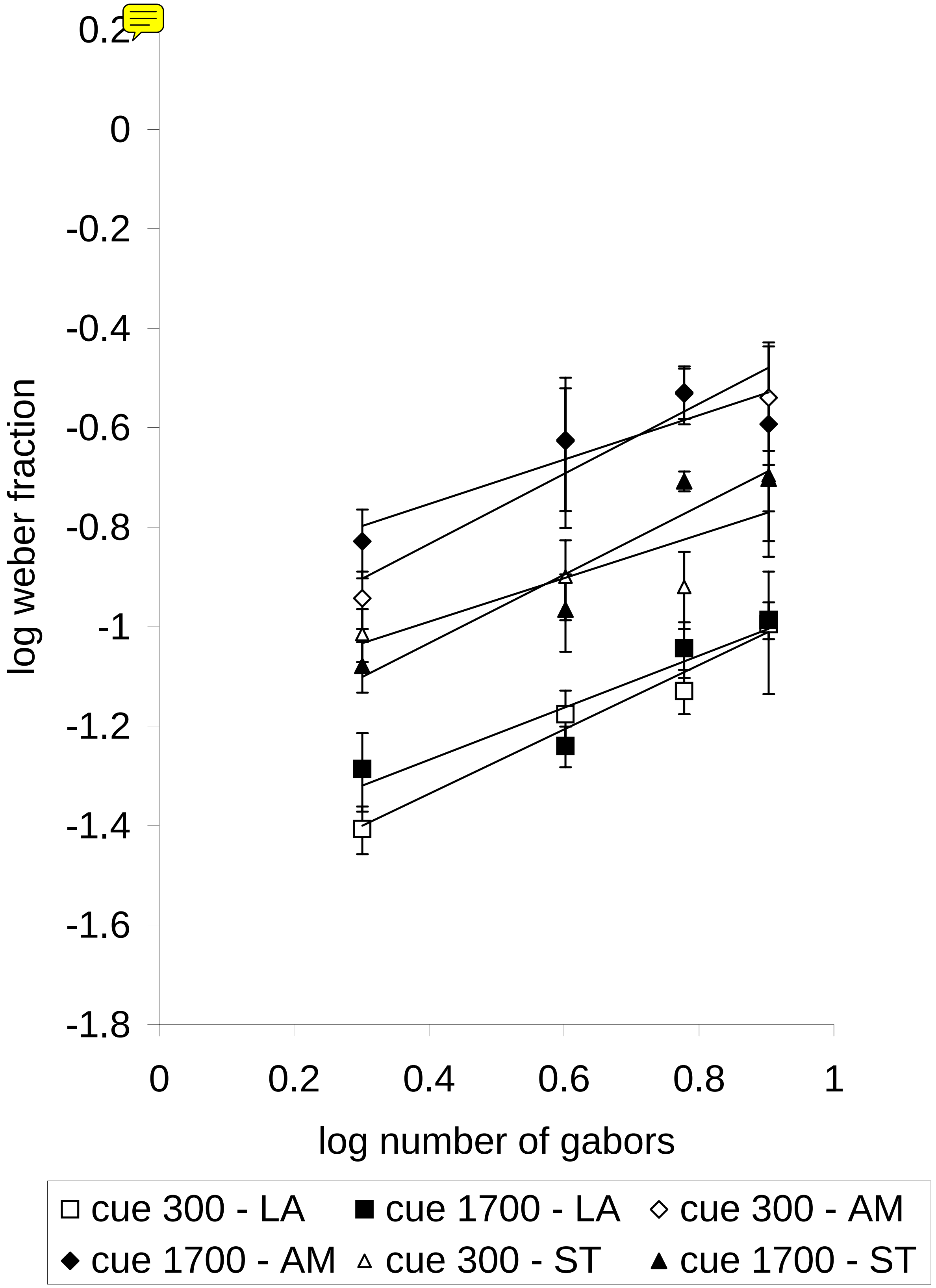


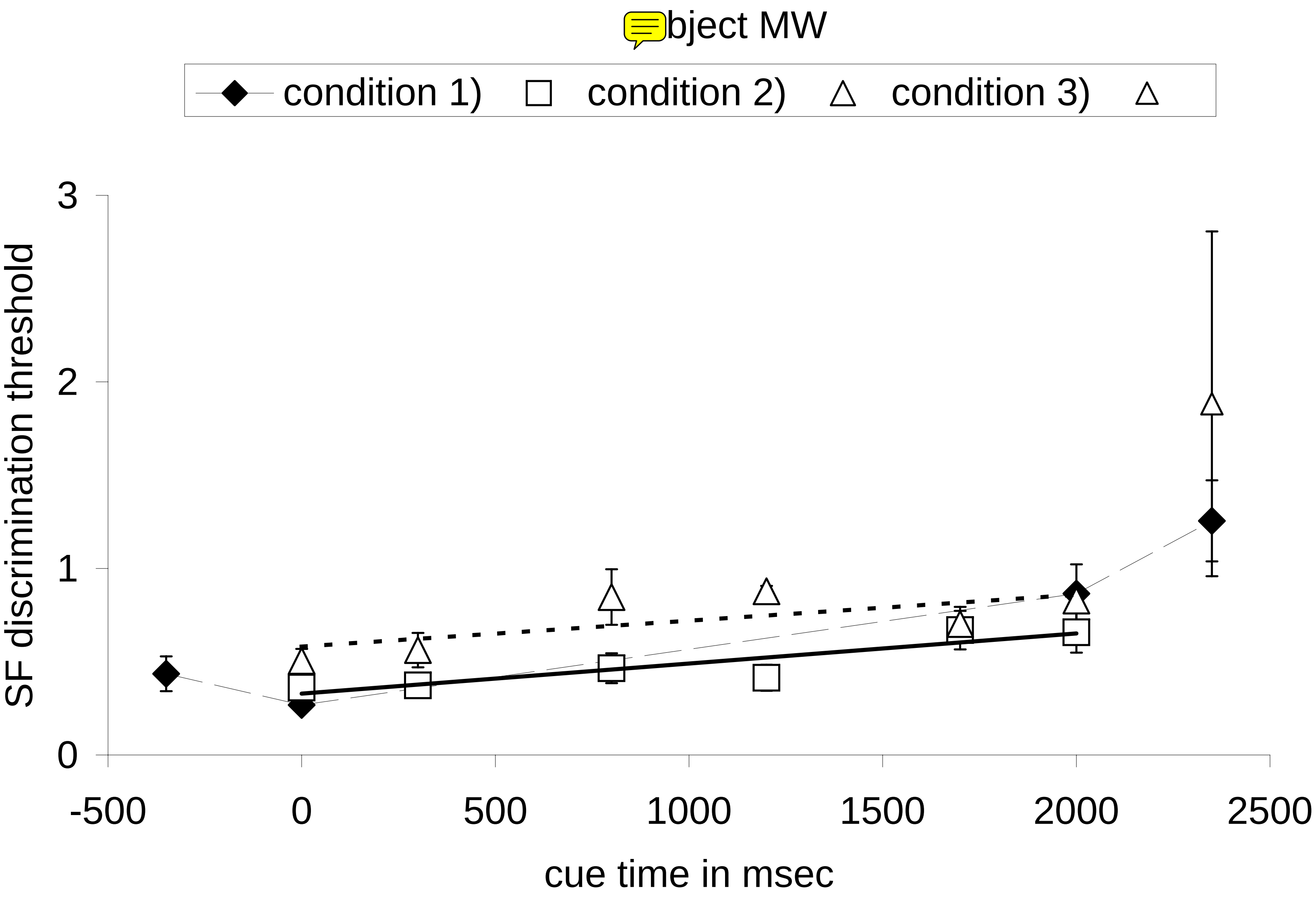
Fig 1. Typical sequence of stimulus frames. The task is to determine the direction of spatial frequency change (increase or decrease) of the cued Gabor target. In this example, the set size is 8 and the cue is placed in the ISI 300 ms after the end of the first display.

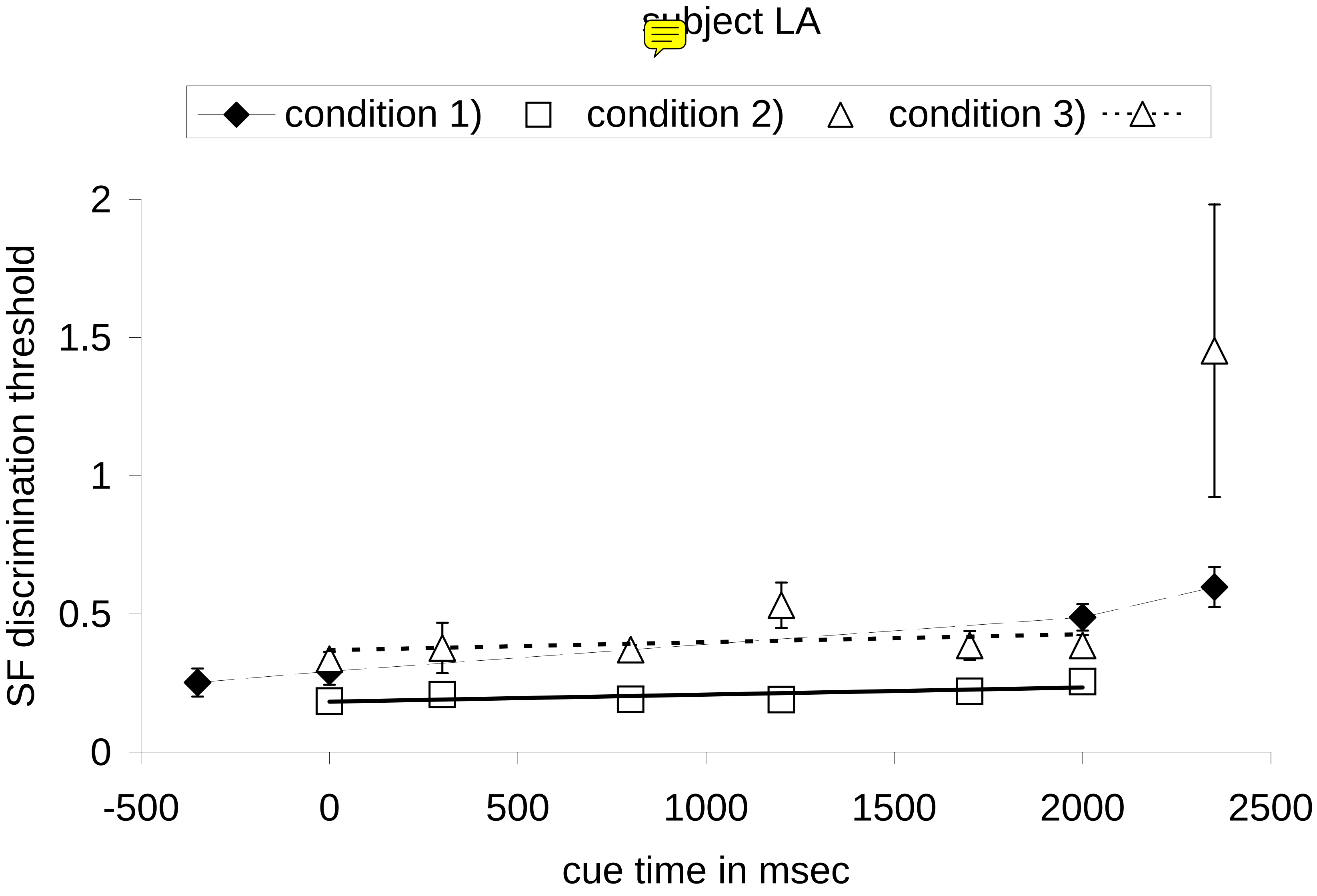

0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
-1.6
-1.8
log Weber fraction
0
0.5
1
log no. of Gabors
◆ MW postcue
■ MW ISI cue
▲ MW precue

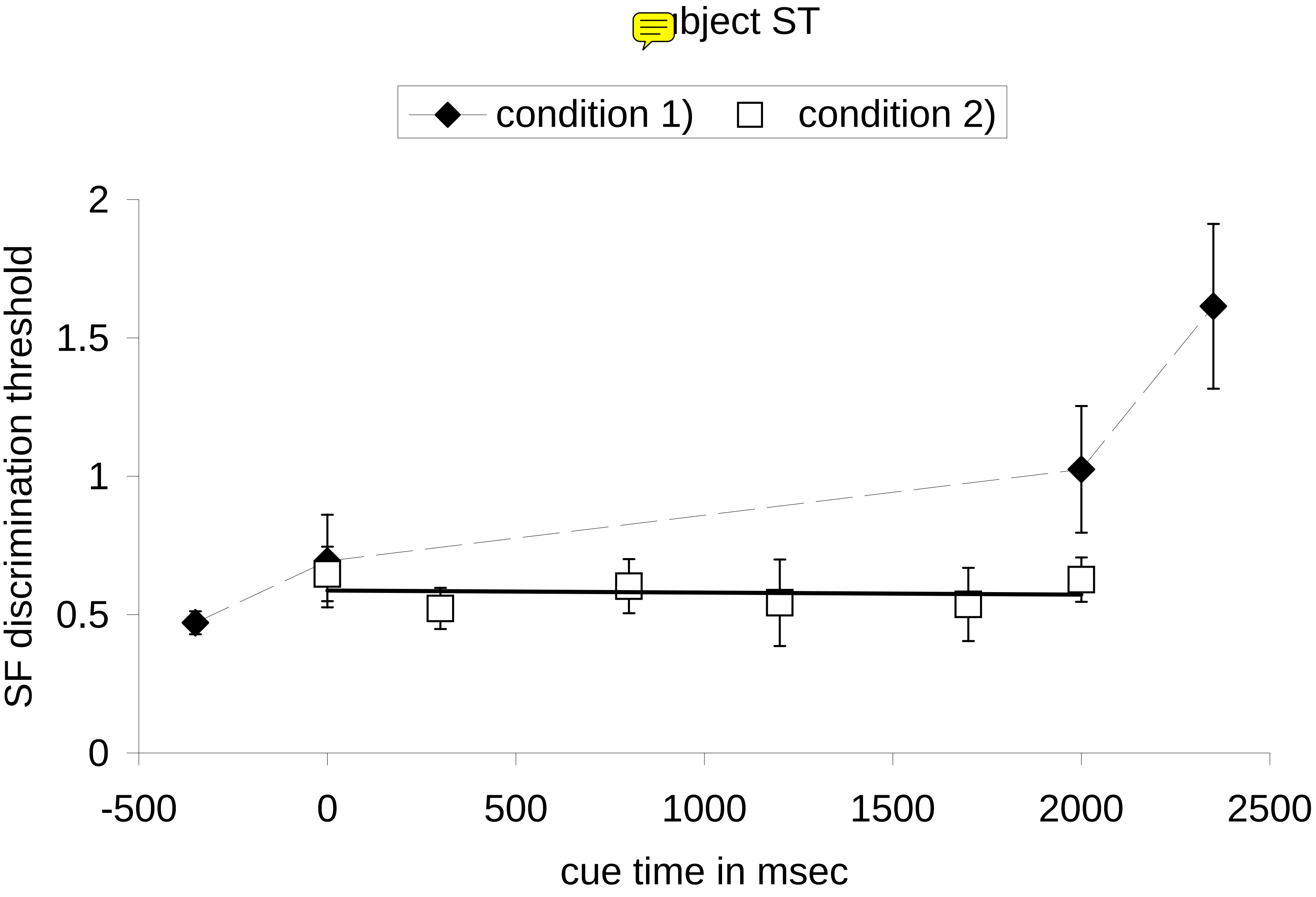
I

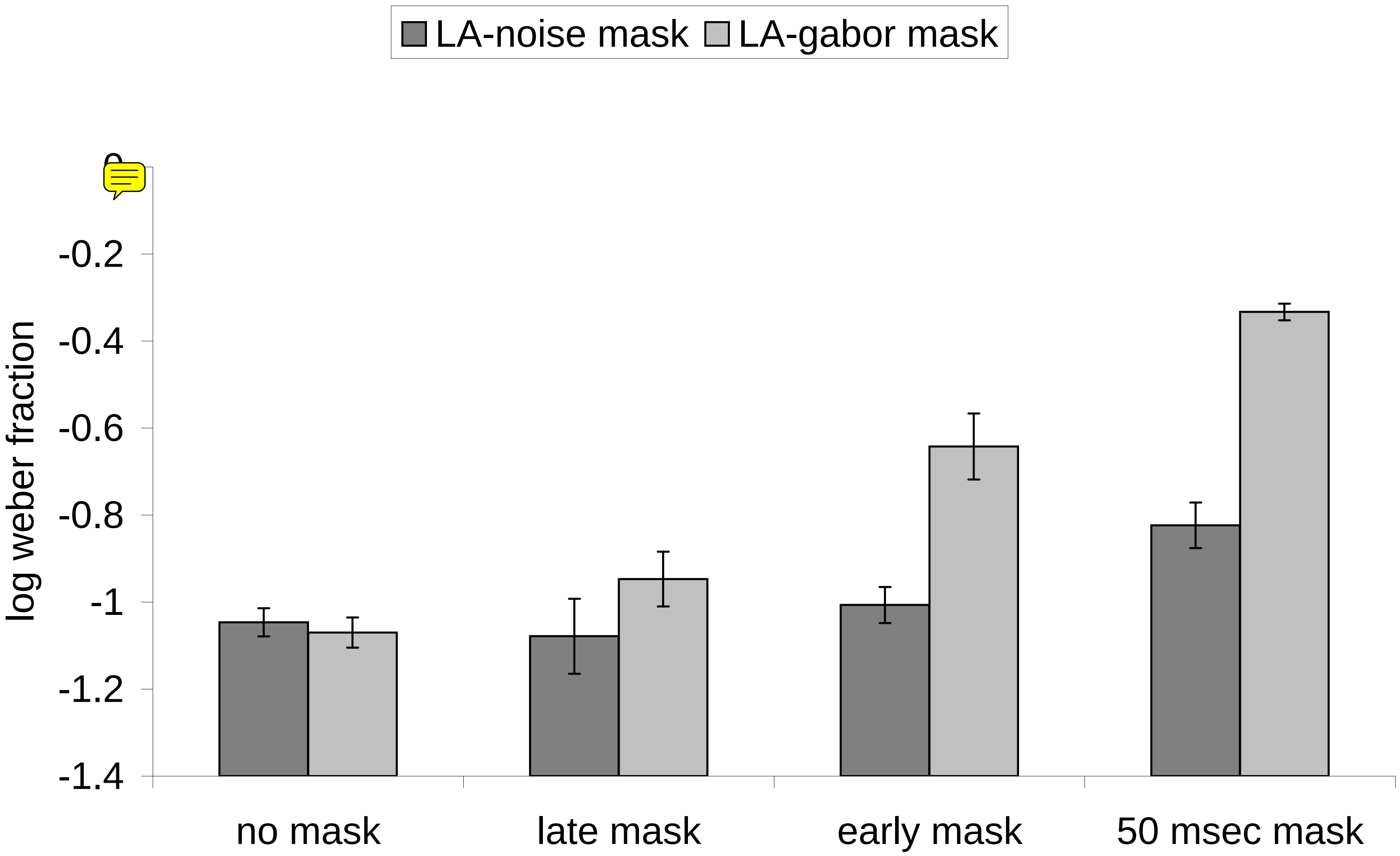

0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
-1.6
-1.8
log Weber fraction
0
0.5
1
log no. of Gabors
◇ SM postcue
□ SM ISI cue
△ SM precue

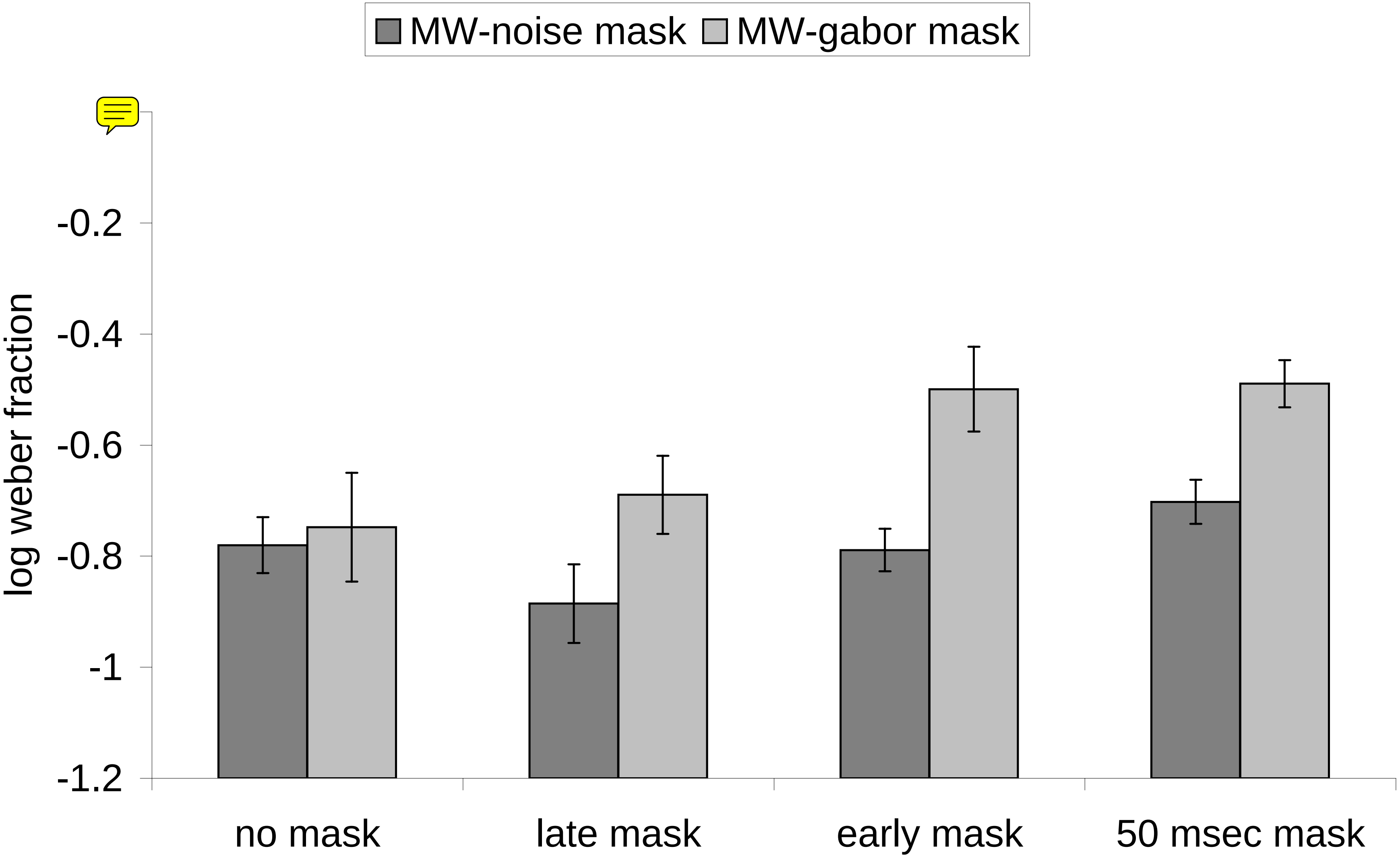

log weber fraction
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
-1.6
-1.8
0
0.2
0.4
0.6
0.8
1
log number of gabors
□ cue 300 - LA
■ cue 1700 - LA
◇ cue 300 - AM
◆ cue 1700 - AM
△ cue 300 - ST
▲ cue 1700 - ST

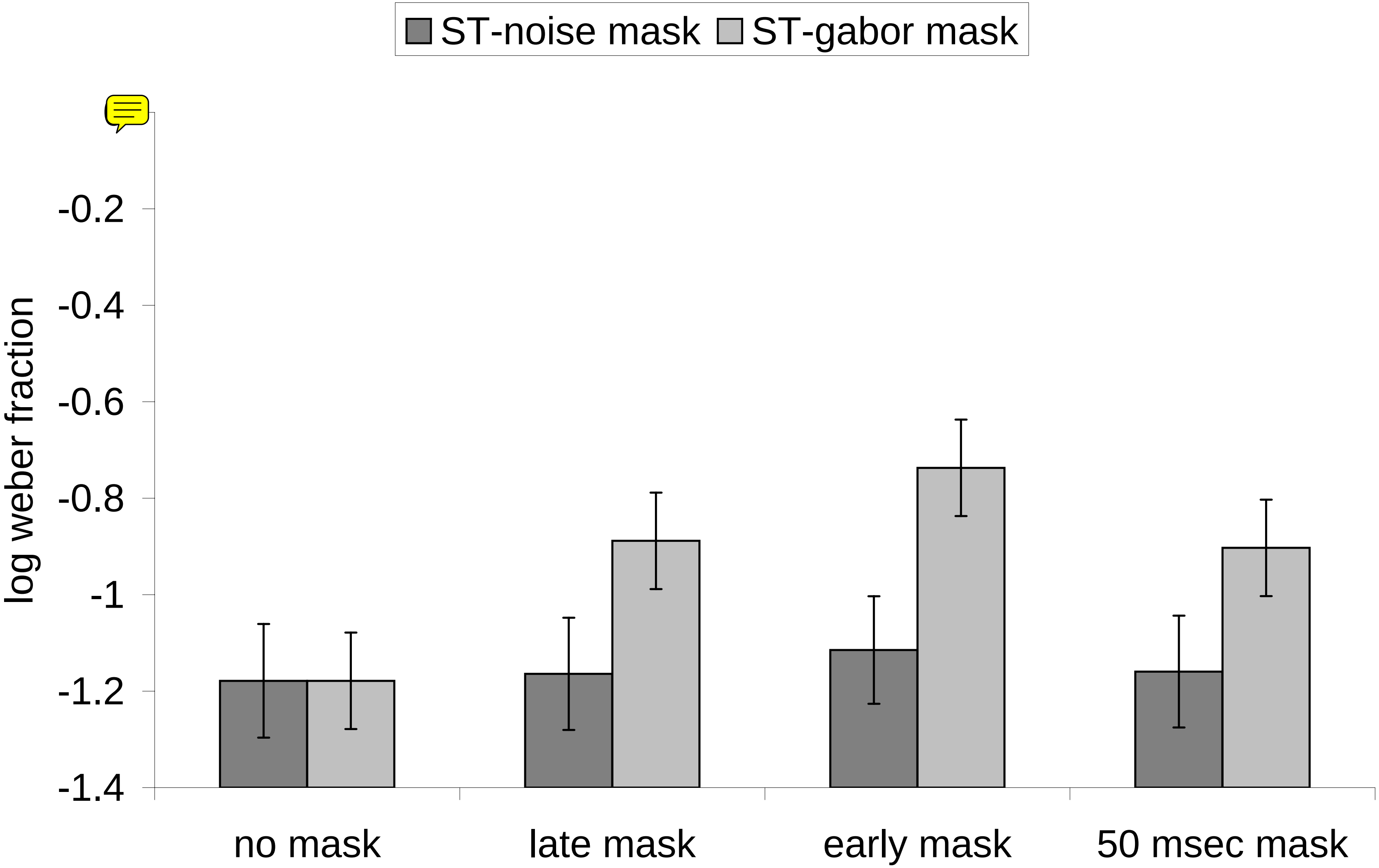

bject MW
condition 1)
condition 2)
condition 3)
3
2
1
0
SF discrimination threshold
-500
0
500
1000
1500
2000
2500
cue time in msec

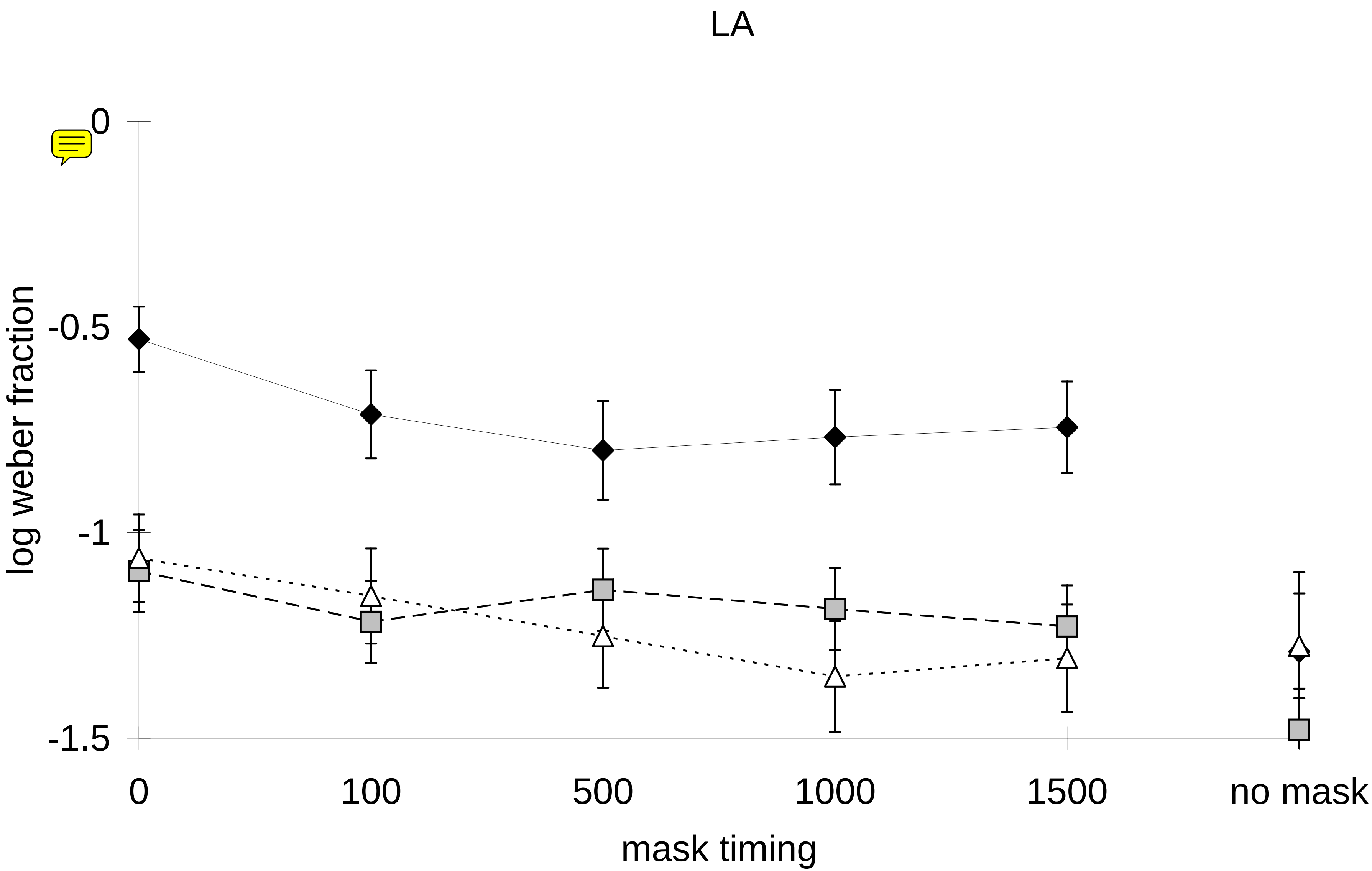

subject LA
condition 1)
condition 2)
condition 3)
SF discrimination threshold
2
1.5
1
0.5
0
-500
0
500
1000
1500
2000
2500
cue time in msec

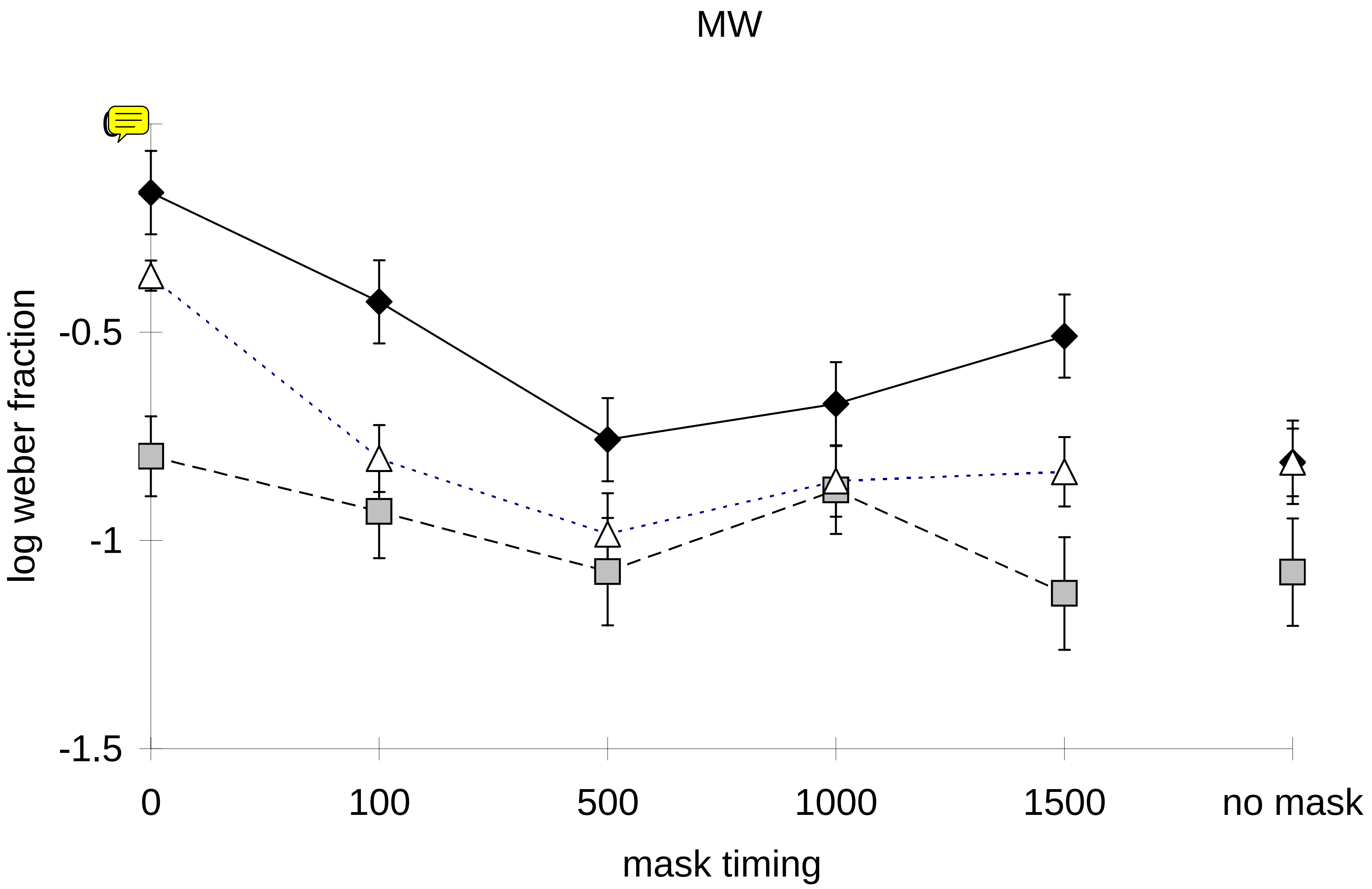

bject ST
condition 1)
☐ condition 2)
SF discrimination threshold
2
1.5
1
0.5
0
-500
0
500
1000
1500
2000
2500
cue time in msec

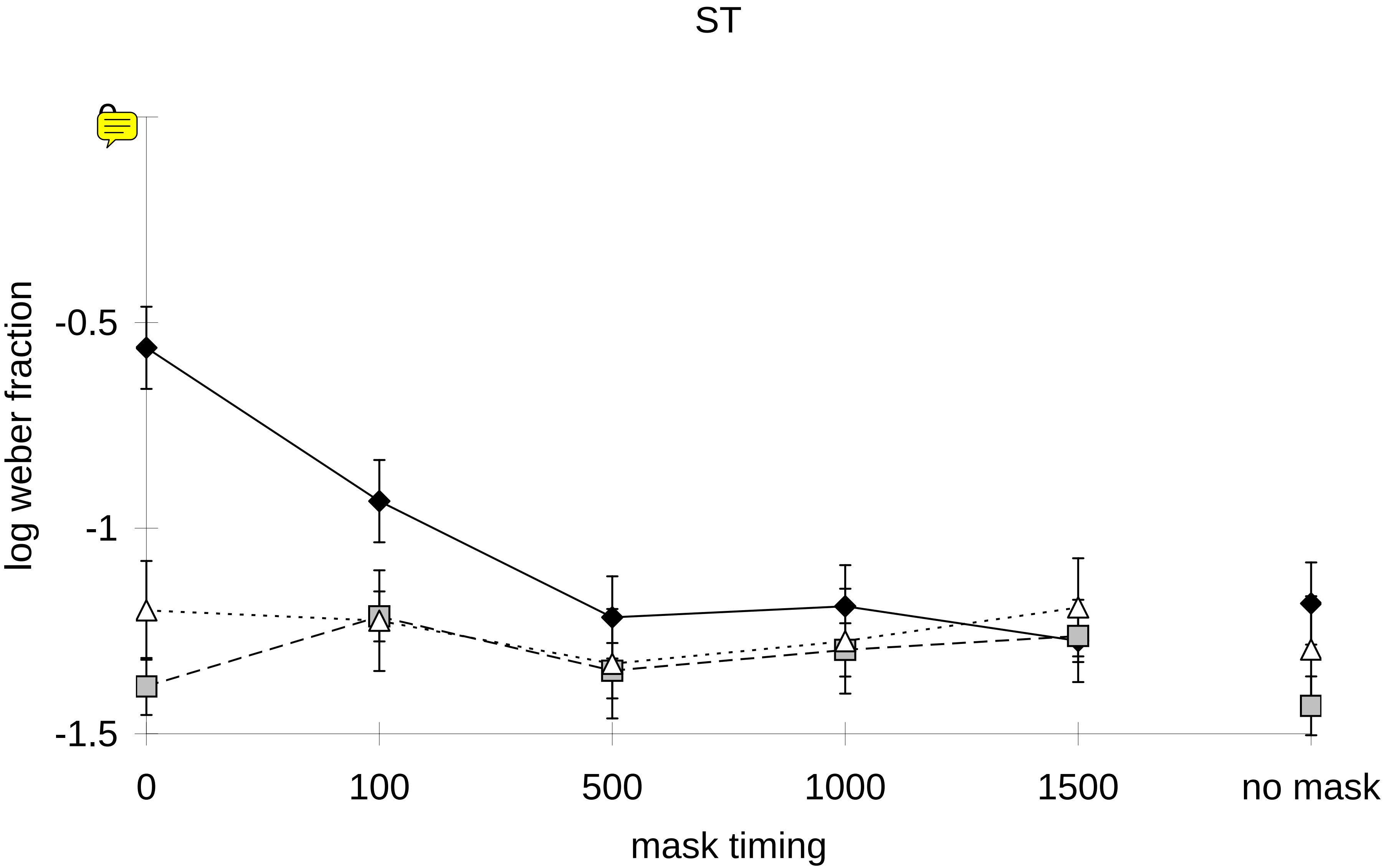

■ LA-noise mask
□ LA-gabor mask
log weber fraction
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
no mask
late mask
early mask
50 msec mask


MW-noise mask
MW-gabor mask
log weber fraction
-0.2
-0.4
-0.6
-0.8
-1
-1.2
no mask
late mask
early mask
50 msec mask


ST-noise mask
ST-gabor mask
log weber fraction
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
no mask
late mask
early mask
50 msec mask


LA
0
-0.5
-1
-1.5
log weber fraction
0
100
500
1000
1500
no mask
mask timing


MW
log weber fraction
-0.5
-1
-1.5
0
100
500
1000
1500
no mask
mask timing


ST
log weber fraction
0
-0.5
-1
-1.5
0
100
500
1000
1500
no mask
mask timing