Cued visual selection of targets with and without orientation contrast

Hans-Christoph Nothdurft
Visual Perception Laboratory (VPL) Göttingen, Germany

The method of cued visual selection (CVS) was used to measure the identification speed of targets in different surrounds. Local orientation contrast is known to modulate neural responses in the visual system; we should expect targets with orientation contrast being faster identified than targets surrounded by similar lines. This was confirmed in the data of five (of six) observers in the main experiment of the study. Performances were measured with line arrays in which individual lines ("targets") were cued at various delays after stimulus onset. At all tested cue delays from 100 ms to 5 s after stimulus onset, targets in *popout* configurations were faster identified than cued targets in *uniform* configurations. At shorter delays (0 ms and 50 ms), differences were absent or not significant. Targets in border-like configurations, which 37.5% of the orientation contrast of popout targets, were identified at intermediate rates. Surrounds also affected the time course of growing target identification with increasing presentation time (after the cue), basically confirming predictions made from neural population responses at various delays after stimulus onset. The sixth subject behaved differently and produced opposite results. It was conjectured that this might have been due to crowding and an inability to quickly identify targets at the tested eccentricity. To test this hypothesis, this and two other subjects were tested in a modified experiment, in which target eccentricity was reduced and cued targets were presented closer to the fixation spot. This modification was sufficient to generate performance differences similar to those obtained with the other observers in the main experiment. Altogether, the experiments have uncovered perceptual variations in target identification that were predicted from neuronal response differences between uniform and popout targets. CVS has thus been proven a useful method for looking into details of neural processing in the visual system with psychophysical methods. *© Author*

Citation: Nothdurft, H.C. (2017). Cued visual selection of targets with and without orientation contrast. VPL-reports, 7, 1-22, www.vpl-reports.de/7/

INTRODUCTION

Recently published experiments with the cued visual selection (CVS) paradigm (Nothdurft, 2017) suggest that exogenous cuing may provide access to the neural representation of perceptual processes. When line arrays were shown in which one line (then named the “target”) was cued and had to be identified, performance varied strongly with the cue delay after stimulus onset. This indicates a timed access of visual perception to the underlying neural representation at the occurrence of the cue. For example, observers needed much shorter target presentation times to identify the target when the cue was presented shortly (100-300 ms) after stimulus onset (when transient neural responses were presumably strong) than when the cue was presented earlier or later (when neural responses had either not yet started or were diminished again).

To validate the model behind these observations, I searched for a timely process in visual perception that is known to generate responses differences which themselves are not perceived. Based on my own experience with popout from feature contrast (e.g., Nothdurft, 1993, 2015), in particular from orientation contrast (e.g., Nothdurft, 1991, 1992) I thought it might be worthwhile to look into the dynamics of target identification when the target is surrounded by either the same or orthogonal lines. From cell recordings we know that lines surrounded by orthogonal lines evoke stronger responses in many cortical neurons than lines surrounded by similar lines (e.g., Nelson & Frost, 1978; Knierim & Van Essen, 1992; Li & Li, 1994; Sillito, Grieve, Jones, Cudeiro, & Davis, 1995; Kastner, Nothdurft, & Pigarev, 1997; Nothdurft, Gallant, & Van Essen, 1999). The response differences are likely caused by orientation-specific suppression from regions outside the "classical" receptive field (see also Nothdurft, 1997; Li, 2002; Gao, Mahadevan, & Vasconcelos, 2008) and represent one form of contextual modulation that was frequently associated with an increased salience of targets with high feature contrast compared to targets with little or no feature contrast (cf. Nothdurft, 1993, 2015). Contextual suppression is graded and may also enhance the responses at texture borders (Lee, Mumford, Romero, & Lamme, 1998; Lamme, Rodriguez-Rodriguez, & Spekreijse, 1999; Nothdurft, Gallant, & Van Essen, 2000; Rossi, Desimone, Ungerleider, 2001; Marcus & Van Essen, 2002). In orientation it is strongest when a line is entirely surrounded by similarly oriented lines (the *uniform* target configuration) and smallest when all surrounding lines are orthogonal (the *popout* target configuration). The strength of suppression is "in between" when half of the surrounding lines are identical to the target, the others orthogonal, as it would be the case at texture borders between areas of differently orientated lines (the *border* target configuration) (Nothdurft, Gallant, & Van Essen, 2000).

Recordings have also shown that the onset of orientation-specific suppression from the surround is shortly delayed against the onset of activation from the center of a neuron's receptive field (Knierim & Van Essen, 1992; Lamme, 1995; Zipser, Lamme, & Schiller, 1996; Lamme, Rodriguez-Rodriguez, & Spekreijse, 1999; Nothdurft, Gallant, & Van Essen, 1999; Smith, Kelly, & Lee, 2007). The reported latencies of orientation-specific suppression vary between about 15 and 30 ms but it is not clear whether suppression from the surround is systematically delayed through the neural net or is due to a general time-distance relationship which would then delay the suppression from the naturally more distant "surround" (Bair, Cavanaugh, Movshon, 2003; but see Smith, Bair, & Movshon, 2006). In psychophysical studies, I could predict and confirm the delay by flickering targets and surrounds; targets that flickered slightly out-of-phase with their surrounds, were better detected than targets that were flickering in-phase (Nothdurft, 2002b).

Given all these informations, I wondered if it would be possible to reveal the differences in the neural responses to uniform, border, and popout target configurations also in a CVS experiment. The principle idea was to present targets in either one of these conditions and measure if CVS performance would differ.

The population responses in Figure 1 illustrate the effects we might expect. The data are from Nothdurft, Gallant, & Van Essen (1999); recordings had been made in macaque area V1 under anesthesia. The figure shows the mean response of 32 "OC" cells to an optimally oriented line when the surrounding lines (outside the classical receptive field) had either the same or orthogonal orientation (uniform and popout conditions, respectively). The population response to the popout target is notably larger than the response to the same target in the uniform configuration. This difference begins 63 ms after stimulus presentation (Fig. 1), about 15 ms after the onset of the response, and holds on for the whole length of recordings (recorded up to 500 ms). OC cells ("orientation contrast") were identified by their differential response to such stimuli and should thus have produced a particular strong response difference. But recordings have been made under anesthesia and responses differences may actually be smaller than to be expected in an active, attentive visual

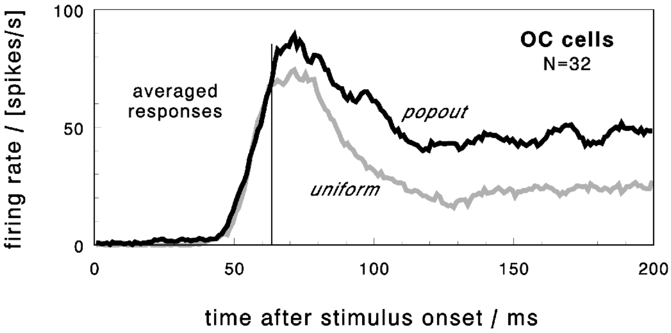


Figure 1. *Population responses of "OC cells" in macaque area V1 to optimally orientated lines in uniform and popout configurations* (modified from Nothdurft, Gallant, and Van Esssen, 1999). Responses were averaged from 32 neurons that had been classified to respond better to an optimally oriented line in the receptive field center when lines in the surround outside the classical receptive field were oriented orthogonal ("popout") than when they were oriented in parallel ("uniform") to the center line.

system. Under anesthesia, OC cells made up about a quarter of recorded neurons (Nothdurft, Gallant, & Van Essen, 1999); in the alert monkey, the proportion was slightly larger (32%; Knierim & Van Essen, 1992).

If the behavioral identification of cued targets would be based on, and temporally related to the neural responses of OC cells in area V1, we should expect that targets in popout configurations are better (and faster) identified than targets in uniform configurations, since the neural signals to both targets differ in strength. The difference should be present all over the response, after a delay of about 63 ms after stimulus onset. Before that delay, targets in uniform and popout configurations did not evoke different responses (and before about 58 ms after stimulus onset, they did not evoke any response at all). Thus, if we assume that the decision process in target identification requires the accumulation of neural signals over some time, we should expect that the identification of a uniform target requires longer target durations for a reliable answer than the identification of a popout target with its, on average, larger response in area V1. And given the transient response characteristics of neurons in Figure 1, the required accumulation time for reliable target identification should be particularly short, when targets are cued soon after stimulus onset so that target analysis will mainly fall in the response peak.

METHODS

Overview

The main goal was to measure target identification from cued visual selection (CVS) when targets were presented in either uniform, border-like, or popout configurations. The difference is illustrated in Figure 2. Test patterns showed regular arrays of randomly tilted oblique lines. One of these lines (the *target*, in one of these three configurations) was later cued. To measure the speed of cued target identification, test patterns remained visible for a limited time after the cue (referred to as *target presentation time,* or *target duration*) and then were masked (Fig. 2b). For each tested cue delay, target presentation time was systematically varied to obtain performance ratings between 50% (chance) and 100% correct. These ratings were later fitted with cumulative Gaussian curves to evaluate the target presentation time at 75% correct responses (Δt_{75}), for every cue delay.

Stimuli

Test patterns were 7 x 7 regular arrays of oblique lines that were individually tilted to the left or right (±45°); the orientations were randomly assigned in every new test pattern. In the course of a trial, one of these lines was cued thus becoming the target that the observer had to identify. Cues occurred at various delays (0-5 s) after test pattern onset. The orientation assignment of individual lines in a pattern was not entirely random but was modified (Fig. 2a) in such a way that the eight lines surrounding the target did either share the target's orientation (the *uniform* condition), were all orthogonal to it (the *popout* condition), or were split to form a texture border with 5 same and 3 orthogonal lines (the *border* condition). Texture borders were randomly oriented in vertical or horizontal orientation with the target on its left or right side and above or below the border. The orientations of all remaining lines in the pattern were chosen randomly. Observers were not informed about the different test conditions, and the differences were not intuitively obvious, since the random orientation assignment with the majority of lines (39 of 48) had generated many almost similar line configurations as those under test. Different test conditions were intermingled and the target location was not known until the cue occurred. Even the author, who is an experienced observer in such tasks and also did know about test conditions and the algorithms of pattern construction, could not reliably predict the target location from the brief inspection of line patterns before the cue. While true popout or uniform target configurations (all 8/8 surrounding lines are parallel or orthogonal to the target) occurred by chance in only 0.4% of the random configurations, nearly popout or uniform configurations (with 7/8 or 6/8 identical lines in the surround) were more frequent (about 3% and 11%, respectively). Such configurations are almost as easily recognized as popout or uniform.

All stimuli (Fig. 2) were white on gray background (no color), except for the green fixation cross and stimulus frame, and were generated with standard DOS VGA technique. They were displayed with 100 Hz repetition rate on a Sony monitor 73 cm (±1.5 cm) in front of the observer. Variations in the viewing distance were due to the dispense of chin rest and bite bar; instead observers sat relaxed with the head conveniently leaned against the wall. Monitor position was fixed. This gave a constant viewing distance for every subject; across subjects, however, the viewing distance varied slightly with the

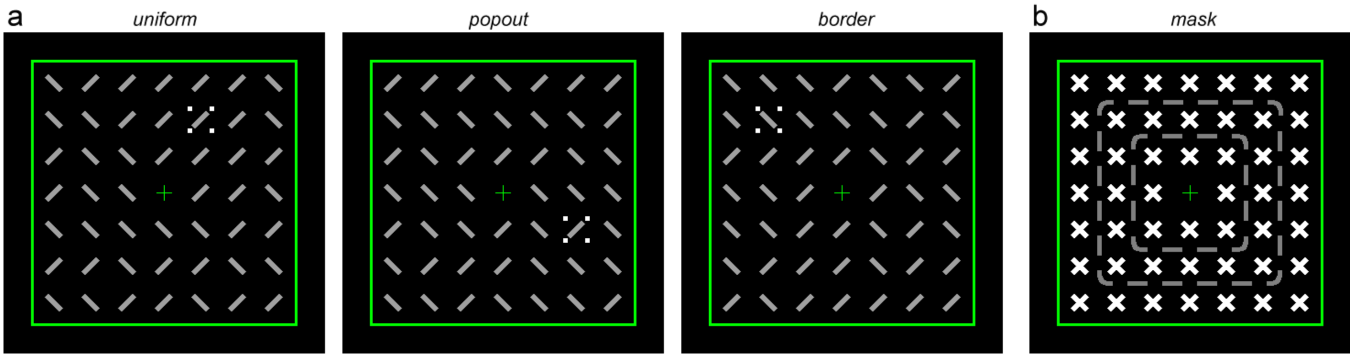


Figure 2. S*timulus examples*. **a**. Typical test patterns with cued targets in uniform, popout, and border-like arrangements; **b**. mask. Subjects had to fixate the central cross. At various delays after stimulus onset, a four-dot cue was shown (20 ms, here superimposed), and subjects were asked to indicate the orientation of the cued target. The presentation time between occurrence of the cue and the pattern replacement by the mask was systematically varied to measure performance variations between target configurations and at different cue delays. Cue (and target) locations were restricted to a rectangular band around the fixation cross (between the dashed lines in *(b))*; in a modified version of the experiment, they were restricted to the inner area next to the fixation point. Dashed lines are only shown for illustration; they were not visible in the experiment. Labels of target configurations in *(a)* refer to the eight lines surrounding the target; they could have same ("uniform") or orthogonal orientation ("popout") to the target, or display an intermediate "border" configuration with five lines parallel and three orthogonal to the target. Borders were generated in four different configurations, horizontal below (as shown) or above, and vertical to the left or right of the target (not shown). All other lines in the pattern were randomly assigned to one of the two oblique orientations. Aim of the experiment was to look for differences in target identification rates between the different target configurations.

different head sizes. All size computations below are based on the mean viewing distance of 73 cm. Line elements in the stimulus patterns were 50' x 12.5' at 18 cd/m² on a background of 9.5 cd/m². Raster width was 1.5 deg; no positional jitter was applied. The line array was presented within a rectangular frame (10.6 deg x 10.6 deg) as illustrated in Figure 2. Masks were made from superimposing the two orthogonal lines that could occur at each raster position and were enhanced in brightness (25 cd/m²). The central line element in the raster was left out and replaced by a central fixation cross (38' x 38'; 37 cd/m²; green). Target selection was provided by “four-dot cues" (cf. Fig. 2a), an arrangement of four small squares, each 11' x 11' and displaced 38' from the target center in the four oblique directions. Such cues have earlier been shown to produce fast and reliable results in target identification without deteriorating target visibility by spatially overlapping cues (Nothdurft, 2016). Cues had a luminance of 58 cd/m² and were always presented for 20 ms and then switched off.

Subjects

Altogether six students (19-26 yrs, 2 females), who were paid for the time they spent in experiment, and the author (male, 68 yrs) performed as subjects in different parts of the study. All except the author were naïve as to the purpose of the experiment, and all had, at least, normal or corrected-to-normal visual acuity.

Procedures

Training, test sequences and blocking. Previous studies (e.g., Nothdurft, 2002a, 2016, 2017) have reported considerable improvements in performance at the beginning of experiments. Therefore, all subjects who had not yet performed in this or a similar task, were given two or three initial training sessions to become familiar with the task and improve their ability to identify targets from short presentations, before the measurements began. The final measurements intended to establish performance rating curves with usually 10-15 different target durations at each of typically 10 cue delays for all three target configurations; that gave a total of 300-450 test conditions each of which had to be repeated 30 to 50 times. It is obvious that this large sample had to be diminished wherever possible and that test presentations had to be blocked. In the first tests of every subject, blocking was made over same target durations across different delays; every condition was then typically tested 5-10 times in a random sequence. These data provided a skeleton of performance variations across cue delays and target durations, that was filled and extensively tested in subsequent runs. The later runs were individually adjusted for each subject to include his or her most interesting target durations (where variations affected rating performance) and avoid too many test conditions far outside that interval. These runs were blocked for the same cue delay, and usually covered 5 repetitions of each test condition (various target durations at this delay). The different target conditions uniform, border, and popout were always intermingled within a run; and it was attempted to include runs with all different cue delays, in random order, in every session. The goal of testing was to obtain reliable performance data over the interesting ranges of target durations at every tested cue delay.

Data analysis. Runs were intermingled in sessions of two hours. Data were collected in up to ten sessions from each subject. When performance ratings were considered to be continuous and reliable, usually with 30-50 repetitions of every test condition, data collection was terminated and curves were fitted with cumulative Gaussian functions to evaluate the target presentation time for 75% correct responses (Δt_{75} values). With all repetitions, each such fit is typically based on 300-600 stimulus presentations at a given cue delay.

Identification of single targets. In addition to targets in uniform, border, and popout configurations at various cue delays, one curve was taken with a single (cued) target with no surround. Since a single line is immediately classified as target, only the 0 ms delay (simultaneous presentation of line and cue) was tested. The test procedure was similar as with targets in line arrays (different presentation times before the target was masked). These tests were blocked in separate runs, which were randomly inserted into the main test sequences. The increase of identification performance with increasing target duration provided an independent measure of target visibility at target onset, without surround.

Responses. Every run started with the presentation of the fixation cross and the stimulus frame, which both remained visible during the whole run. Trials began with the onset of the test pattern upon which, eventually after a delay, the cue (as target marker) was shown for 20 ms. After the target presentation time, the line pattern was masked (Fig. 2b). Observers were supposed to perform the same task throughout all tests, to identify the cued target and to indicate its orientation as being tilted to the left or right by pressing left-hand and right-hand keys, respectively, on a computer board. After key pressing, there was a break of about 1s, before the next trial started. Subjects were not forced to give fast responses but could take all time they wanted before responding.

Fixation. Subjects were asked to fixate a central cross on the screen; eye movements were registered by means of a camera focused upon the observer's eyes. Reliable fixation performance was extensively controlled for during the first sessions of each subject and regularly also in all later sessions. All subjects quickly learned to perform the task under strict fixation, in particular as gaze shifts did not provide any benefits with short target presentation times (cf. Fischer, 1987; Fischer *et al*., 1993).

Modified 2AFC. The task itself was a two-alternative forced choice (2AFC) with two minor modifications: (i) Subjects could reject a single trial if they had noticed that they were not attentive, had been distracted during the trial or had lost fixation, particularly during long cue delays. Rejected trials were put back to the sample of trials to be tested in the current run and were later repeated with a new pattern. Subjects were instructed not to use this key if they had been attentive but simply could not identify the target. In that case they were asked to guess. In fact, all subjects made very little use of this modification; but the possibility to reject a trial instead of being forced to guess if they had been inattentive or distracted, was considered helpful. (ii) Subjects could change a given response immediately after it was entered by pressing a special correction key. This modification was particularly helpful at the beginning of the experiment, when subjects were not yet fully familiarized with the response keys and had erroneously hit the wrong key. Some subjects made occasional use of this function in early sessions but rarely or not at all in later ones.

Target locations. It has been shown that target identification rates depend on target eccentricity (e.g., Nothdurft, 2017). To reduce performance variations within identical conditions (but still keep the uncertainty about cue locations sufficiently high), targets were not distributed all over the pattern but presented at selected locations. This was also necessary to provide a full line pattern context at each tested location. The selected target locations are indicated by dashed lines in Figure 2b; only lines within the marked rectangular ring were selected as targets. Subjects were not informed about this restriction. At each of these locations, targets were surrounded by a complete set of eight neighboring lines.

Additional experiment

The main test as described above was carried out by six observers. The performance of one subject, however, suggested carrying out an additional test series with reduced target eccentricity. In this experiment, cues were only presented at locations in the central dashed area of Figure 2b and thus marked targets in the immediate neighborhood of the fixation cross. Only uniform and popout configurations were tested in this modification, but because of the missing line at the location of the fixation cross, targets at these locations were not *fully* surrounded by parallel or orthogonal lines.

RESULTS

The study intended to search for performance differences in the identification of shortly presented targets with different surrounds. For that, line patterns (Fig. 2) were shown in which the targets were embedded in either similar or orthogonal lines (*uniform* and *popout* configuration, respectively). To evaluate intermediate effects, a third condition (*border* configuration) was included, in which five of the surrounding lines had similar, the other three orthogonal orientations. Possible cue and target locations were restricted (as indicated in Fig. 2b) to reduce the performance variations from different eccentricities (Nothdurft, 2017). Beside the target and immediately surrounding lines, all other lines were individually assigned to randomly left- or rightwards tilted oblique orientations (±45°). As described in the Methods section, this has led to the simultaneous occurrence of several border-like and nearly uniform or popout line configurations in every pattern. Which of the many lines in the test pattern finally served as target was only known when the cue had occurred. In addition, the three target conditions were intermingled so that observers could not predict which line in a test pattern was likely to be selected as target. *Cue delays* measured between pattern onset and presentation of the cue, and *target durations* measured from cue onset until the mask occurred were systematically varied over the tests.

The main test series of the study was performed by six subjects including the author (two female). One subject, however, performed entirely different to the other subjects

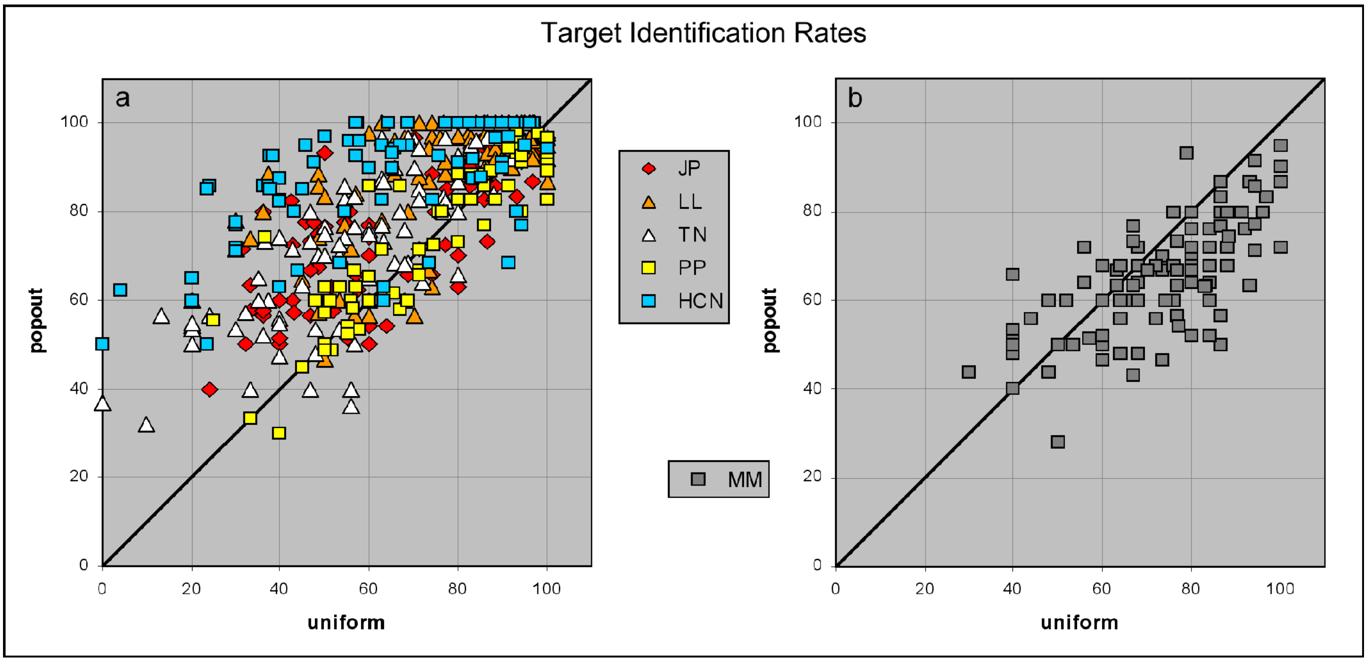


Figure 3. *Scatter diagrams of target identification rates when targets were presented in uniform or popout configurations.* **a., b.** Individual data of all six subjects. Every data point represents a pair of ratings obtained with the same target duration. Ratings were performed by altogether six subjects (different symbols and colors). Most subjects showed a strong, overall preference for targets in popout configurations which could be better identified than targets in uniform configurations. In subject MM, however, the preference was reversed *(b)*, and the subject was therefore excluded from the main analysis (but see Figures 12-16).

and was excluded from analysis of these tests. Possible reasons for her deviations are discussed below. In a second series of the experiment, certain parameters were modified and performance was then tested again with this and two additional subjects one having already participated in the main experiment. Modifications and re-tests will be reported below.

General performance differences between target configurations

Figure 3 shows the target identification ratings of all six subjects for uniform and popout target configurations. The figure includes only data pairs from measurements with the same target presentation time. Additional tests with presentation times at which only one of the two target configurations was tested, are not included. For example, targets in the popout configuration had often to be tested with shorter durations than uniform targets to reach 50% performance. Vice versa, uniform targets often required longer target durations to reach 100% performance, which were not tested with popout targets if these had already been correctly identified at shorter presentation times. Furthermore, all data pairs with 100% performance in both target conditions are omitted in the figure.

The scatter plot in Figure 3a reveals a strong bias; on average, targets in popout configurations were better identified than targets in uniform configurations, at the same presentation time. The mean performance deviation between both configurations over the 467 data points in Figure 3a is 12.70 ± 0.74 % (popout better than uniform) and highly significant (two-sided paired t-test; $p<0.0001$). This bias was found in the rating data of five subjects with individual mean deviations from 5.03 ± 1.31 % (PP) to 24.17 ± 2.12 % (HCN) and was highly significant also on an individual subject's basis ($p<0.0001$; for PP, $p<0.0005$). The scatter plot in Figure 3b, however, shows the notably different performance of subject MM; the mean deviation here is –7.14 ± 1.20 % (uniform better than popout) and is also statistically significant ($p<0.0001$). The reason of this diverging behavior is not yet clear. The systematically better performance with targets in uniform than in popout configurations, however, does not allow us to interpret it as an eventually poor performance of this subject (which might have reduced the bias but should not have turned it into the opposite direction). I will come back to this subject in a later section and will meanwhile restrict all further analysis to the data of the remaining five subjects with consistent response characteristics.

Figure 4 shows the analogue comparison between performance ratings with targets in uniform and targets in border configurations. The general bias of data points to the upper left half of the graph is far less pronounced than in Fig. 3a. The mean performance shift between border and uniform target configurations over the 491 data points in Figure 4 is 3.37 ± 0.54 % (border better than uniform). Although this deviation is highly significant, too (two-sided paired t-test; $p<0.0001$), this was not the case in the individual data of all subjects. Mean deviations between uniform and border target configurations varied from 1.33 ± 1.13 % (JP, not significant) to 7.84 ± 1.37 % (TN, highly significant; $p<0.0001$) and even included a bias in the opposite direction, i.e. an on average better performance with uniform than border-like target configurations (PP; –2.44 ± 1.05 %; $p<0.05$). The data of subject MM (not shown) did not reveal any systematic bias in border-to-uniform comparisons.

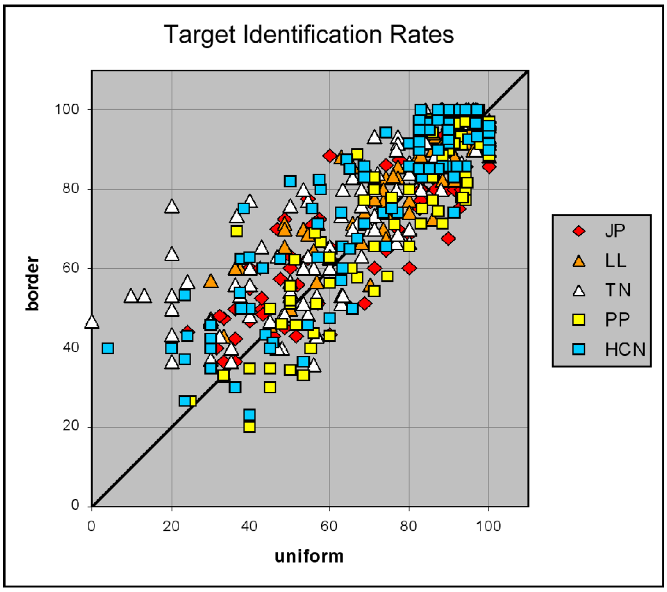


Figure 4. *Scatter diagrams of target identification rates when targets were presented in uniform or border configurations.* Same subjects as in Figure 3a. Each data point represents a pair of ratings obtained with the same target duration. The general performance shift is significant but less pronounced than in Figure 3a.

Delayed contextual modulation

In Figure 1 the population responses to popout and uniform targets begin to differentiate about 63 ms after stimulus onset. To see if a similar time delay was also apparent in the performance bias of Figure 3, analysis was repeated on two sub-samples of data in Figure 3a; sample A in which cue delays plus target durations did not exceed 70 ms after stimulus onset, and sample B covering all other test conditions (cue delay + target duration ≥ 70 ms). Sample A included 22 data points, sample B 445. Scatter analysis revealed that there was indeed no significant performance shift between uniform and popout target configurations in sample A (mean shift 6.30 ± 3.90 %; *p>0.12*), whereas the shift in sample B was slightly increased (mean shift 13.02 ± 0.75 %; *p<0.0001*) compared to the total sample. While the missing significance of the shift in sample A alone must not be overemphasized because of the small number of data points, the simultaneously increased shift in sample B indicates that it were mainly non-biased cases that had been removed from the overall sample in Figure 3a. The transition seems to be continuous, however. When the cutting time between samples A and B is increased beyond 70 ms, the mean shift between target conditions in sample B still increases and then decays; the largest shift (13.13 ± 0.78 %) is obtained with a cutting time of 100 ms. At the same time the difference in sample A continuously grows and begins to become significant.

Altogether, the data indicate that performance differences between popout and uniform target configurations are not immediately present at stimulus onset but occur with a delay (like the differences in the population responses in Fig. 1). However, given the shortness of this delay and the small number of test conditions restricted to it, the evidence from direct performance differences alone is not particularly strong. There will be further evidence from the analyses below.

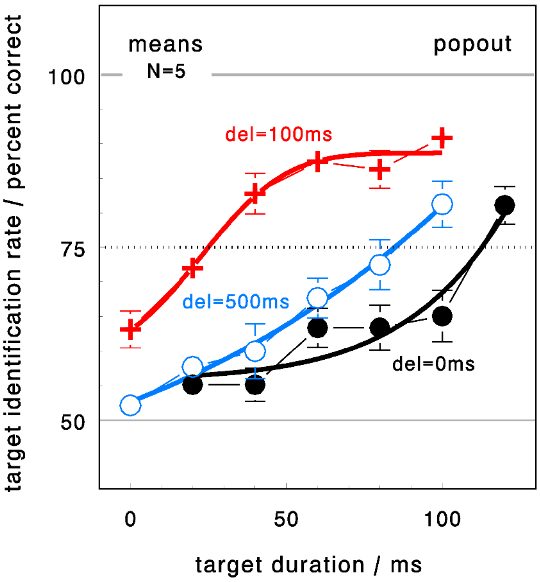


Figure 5. *Target identification with increasing target duration.* Mean data (and s.e.m.) of all subjects (except MM) with popout targets. At cue delay 0 ms, targets had to be presented fairly long before identification rates increased; at cue delay 100 ms performance increased immediately. At cue delay 500 ms, performance increased constantly and moderately. Thick curves are nonlinear fits to the data (see legend of Fig. 6). Horizontal gray and dashed black lines in this and all following figures indicate the valid performance range between chance level (50%, no target identified) and perfect performance (100%, all targets identified), and the performance level in-between (75%, correct target identification in half of the trials).

Cumulated signals

In my earlier CVS study (Nothdurft, 2017) I have shown that target information is generally not immediately available with the onset of the cue. Even when a stimulus pattern is shown several seconds before the cue, the target must still remain visible for a short "presentation time" after the cue if observers want to make a reliable evaluation of its orientation (cf. the demo on *www.vpl-goettingen.de/cvs/*). One plausible interpretation was that neural signals must be accumulated over time to reach a certain level that would qualify the visual system to make that decision. Since neural responses vary over time and since the needed accumulation time varied accordingly with the cue delay after stimulus onset, it seemed plausible to relate cumulative performance to the strength of underlying neural signals (for details, see Nothdurft, 2017).

From the population responses of OC cells in Figure 1 we can formulate several immediate expectations from such a model. (i) Targets cued right at the beginning of the response train (i.e., simultaneously with stimulus onset) should require longer presentations until the relevant neuronal signals begin to grow and a decision on target orientation can be made. (ii) On the contrary, for targets that are cued during the peak of the neuronal response (e.g., at a delay of 100 ms after stimulus onset) the relevant information should be quickly accumulated, as the neuronal signals encoding the target's orientation are particularly strong. (iii) Targets that are cued later (e.g., 200ms or 500ms after stimulus onset) should reveal smaller but continuous increments in identification performance, since neurons might still be active but fire less strongly than at the initial peak response. This is indeed what was found. Figure 5 shows the mean performance of all five subjects with popout targets at selected cue delays, that differ in the predicted way. Although the curves look somewhat incomplete (which is due to the fact that not all five subjects had been tested with exactly the same presentation times outside the plotted range), the curves reveal the characteristic courses predicted from the neural population response in Figure 1.

On the same assumption (stronger neuronal responses lead to faster target identification) we can predict that targets in popout configurations should need shorter accumulation time for reliable responses than targets in uniform configurations. Or, in an equivalent formulation, at a given presentation time popout targets should be better identified than uniform targets, which was already shown in Figure 3a.

Performance variations with different delays and target configurations

The different increases of identification rates with increasing target duration are illustrated in Figure 6. Performance variations across different cue delays and target surrounds are exemplarily shown in the data of three subjects (left-hand, middle, and right-hand graphs, respectively). For each subject, two response curves at different delays are plotted, both for targets in uniform (Fig. 6a) and targets in popout configurations (Fig. 6b). All curves rise with increasing target durations, i.e., subjects could identify targets better the longer these remained visible after application of the cue; but slopes differ considerably. Each subject could identify targets at short cue delays (100 ms or 150 ms) much better, i.e., from shorter presentations, than targets at long cue delays (2000 ms or 5000 ms). This is remarkable, since the full stimulus presentation time at long delays was about 20 times as long (before the cue) as that at short delays; nevertheless, targets were more quickly identified after short than long delays. It indicates that the long stimulus presentation before the cue did, in general, not improve (fasten) the target identification process. The other important observation in Figure 6 are the differences between uniform (Fig. 6a) and popout targets (Fig. 6b). Curves plot performance variations at the same cue delays (same colors); the curves obtained with popout targets are always shifted to shorter target durations than the corresponding curves obtained with uniform targets, from the same subject. The differences are particularly obvious when target durations at 75% correct responses are compared (indicated by vertical lines). These "Δt_{75}" values strongly differ between short and long cue delays and also between uniform and popout target configurations.

Despite the, in principle, similar variations across subjects, there are notable differences between the exact presentation times needed for target identification (note the different time scales in Fig. 6). Subject HCN, for example, required generally much shorter target durations than subject JP, for same performance. Similar variations across subjects were reported in my earlier study (Nothdurft, 2017). Please also note the particularly high performance of subject HCN with popout targets at cue delay 150 ms. Even with a target duration of 0 ms, his identification rate was far above chance, so that "negative" target durations had to be included to measure rating variations down to chance (50%). Apparently, the subject could, at this delay, even identify targets that were already masked when selected by the cue.[1]

The examples in Figure 6 are selected to illustrate the differences between uniform and popout targets and the variations seen with short and long cue delays, that were consistently found across subjects. But it is also important to underline that the differences between target configurations were not seen at all delays. From Figure 1 we expect that differences between target configurations should be absent at short cue delays. This was indeed the case.

---

[1] A possible explanation of this "weird" phenomenon was already given in Nothdurft (2017). If neural responses are particularly strong and effective, observers with high sensitivity (and the need for only short accumulation times) may still receive target signals when the mask is already shown. As seen in Figure 1, the cumulative response at 150 ms would grow much faster than the cumulative response at 0 ms. The mask, however, is nothing else than a disturbing stimulus (here, the presentation of an additional, orthogonal line) that should make further target analysis impossible. Given the different strength of responses at 150 ms (target) and 0 ms (orthogonal line in the mask), signals representing the target orientation should still dominate at zero and even some "negative" target durations.

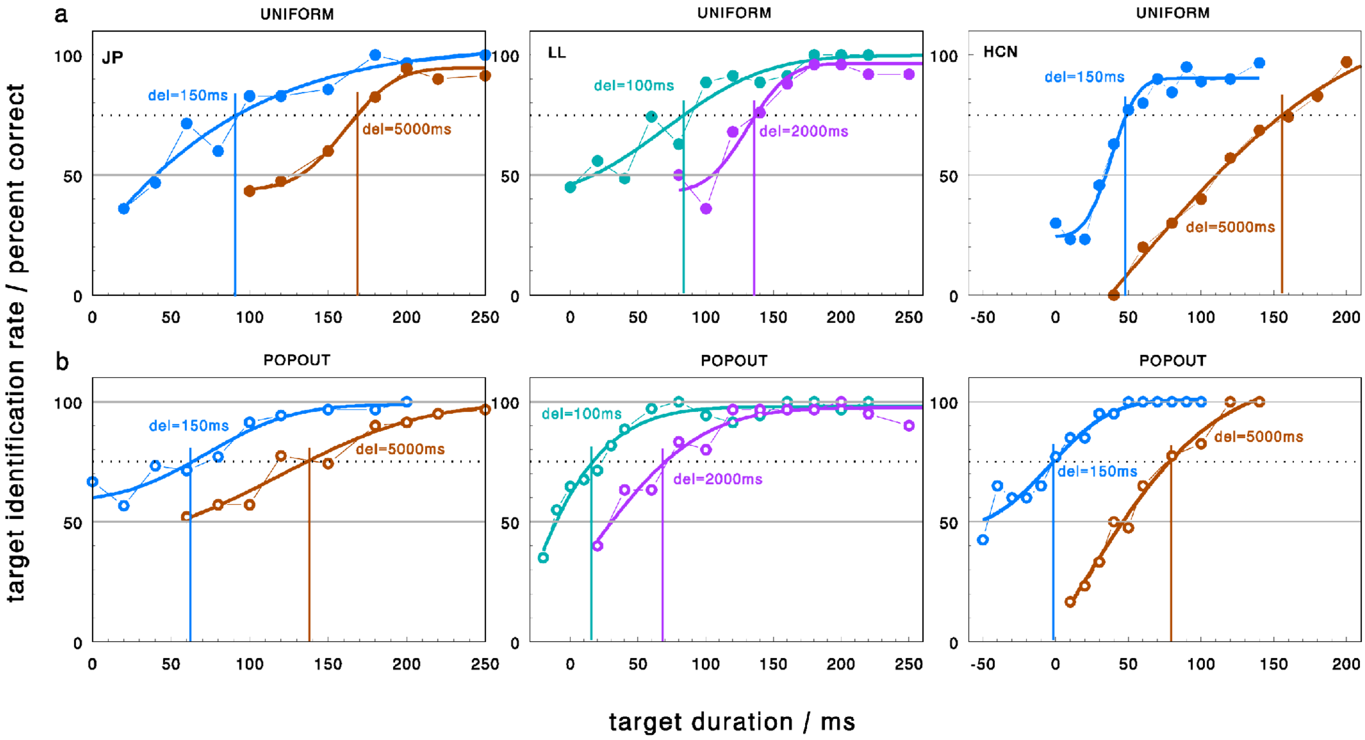


Figure 6. *Examples of cumulative performance curves with increasing target presentation time.* Individual data from three subjects (columns) tested with targets in **a.** uniform and **b.** popout configurations. Curves show the increasing identification rates with increasing target durations in different conditions. Each graph presents the data from two cue delays ("del"), one short (blue or cyan), one long (brown or violet); target identification rates generally increased faster at the short delays. Targets in uniform *(a)* and popout configurations *(b)* were tested at the same cue delays (same colors). Target identification rates generally increased faster with popout than with uniform targets. For later analysis, each curve was non-linearly fitted with a Gaussian cumulative function, from which the target durations at 75% correct were taken (Δt_{75} values), here indicated by vertical lines in the according colors. Please notice the time scale differences between subjects and the measured ratings at negative target durations with subject HCN (see text).

Figure 7 shows again *mean ratings* of all five subjects, now with all three target configurations at various cue delays. To increase the number of data points that cover measurements from *all* subjects for this comparison, a few missing data from two subjects were extra- and interpolated from measurements nearby. Subject HCN had generally been tested with shorter target durations than the other subjects; his missing ratings at long target durations (≥120 ms) were set to 100% if he had already reached 100% performance with shorter target durations at this delay. Subject JP was tested with target durations 120 ms, 150 ms, and 180 ms, whereas three other subjects (not HCN) had been tested with target durations 120 ms, 140 ms, 160 ms, and 180 ms, in this range. To obtain means from all five subjects from durations >120 ms, virtual identification rates of subject JP at 140 ms and 160 ms were obtained from interpolation; this only applied to cue delay 2000 ms (right-hand graph in Fig. 7). All other target durations plotted in Figure 7 were obtained from real measurements.

Consistent with the expectations from the population responses (cf. Fig. 1), all target configurations were identified with rather similar ratings at cue delay 0 ms (Fig. 7, left-hand graph). With increasing delays, the speed of target identification grows faster for popout (red curves) than for uniform (black) or border target configurations (green). This difference holds up to the longest delays tested. For targets in border configurations, rating differences to uniform targets were not consistently found and are best seen in the mean data at delays 500 ms and 2000 ms (right-hand graphs in Fig. 7). Statistical analysis (paired two-sided t-tests of all paired data points at a given delay, including data outside the plots) revealed significant *($p<0.05$)* to highly significant *($p<0.005$)* differences between uniform and popout target configurations, except for delays 0 ms *($p>0.3$)* and 1000 ms *($p>0.08$)*. The differences between uniform and border target configurations were generally not significant *($p>0.07$)*. This pattern slightly changes when more data points are included in the comparison, e.g., by including measurements taken from, at least, three (of the five) subjects. The differences between uniform and popout target configurations then became significant *($p<0.05$)* or highly significant *($p<0.005$)* for *all* delays except synchrony (delay 0 ms), and differences between uniform and border target configurations were significant ($p<0.05$ or $p<0.01$) for, at least, some delays. In the mean data, thus, the delay of contextual modulation is confirmed.

Predictions from simulations

To estimate the expected strength of rating variations with different target configurations, I have performed a simulation based on the population responses in Figure 1. I was particularly interested in the differences between uniform and popout targets (for which responses had been measured) and in the dynamics of target identification at various moments after stimulus onset. This was simulated by integrating (accumulating) the responses in Figure 1, beginning at various delays from stimulus onset. To compare the results, an arbitrary (and constant) threshold was set and the accumulation time (target duration) from cue onset until this threshold is reached was calculated (Fig. 8a). As expected (and experimentally confirmed by the data in Figure 5), the accumulation time to reach the threshold diminishes towards the response peak and increases again when cues are further delayed. The new aspect here are the differences between uniform and popout targets. For short cue delays, both responses reach the threshold at similar durations. From a certain delay on, however, popout targets need shorter presentation times than uniform targets to reach the threshold. The cue delay at which these differences occur depends on the threshold and, hence, on the required accumulated time. If the integration of neuronal signals has to reach into time periods where responses begin to differentiate (> 63 ms in Fig. 1), the required integration intervals for popout and uniform targets will differ. In the simulation shown in Figure 8, the threshold was set low enough so that differences between uniform and popout targets did not yet become pronounced at cue delays shorter than 50-60 ms after stimulus onset. The differences in accumulation time between targets then begin to establish around a delay of 60 ms, grow up to a delay of 100 ms, and from there on remained about constant until the end of the simulation (cue delay 300ms; later response data are not shown in Fig. 1). Note that Figure 8a does not directly represent

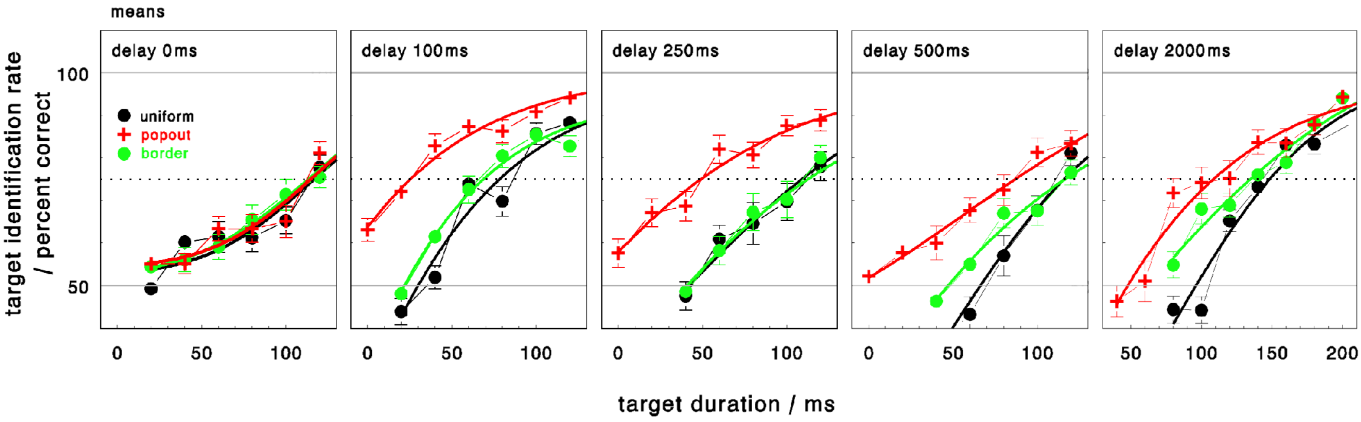


Figure 7. *Target identification rates with increasing target duration for different target configurations.* Mean data (and s.e.m.) of all five subjects (except MM). For targets cued at pattern onset (delay 0 ms), ratings with different target configurations are almost identical (left-most graph). From delay 100 ms on, popout targets were better seen than uniform or border targets. Differences between uniform and border targets were generally less pronounced and are here only seen at delays 500 ms and 2000 ms. For target durations ≥120 ms, the data of two subjects had to be extra- or interpolated to obtain mean values from all five subjects (see text). Fits and performance level markers as in Figures 5 and 6.

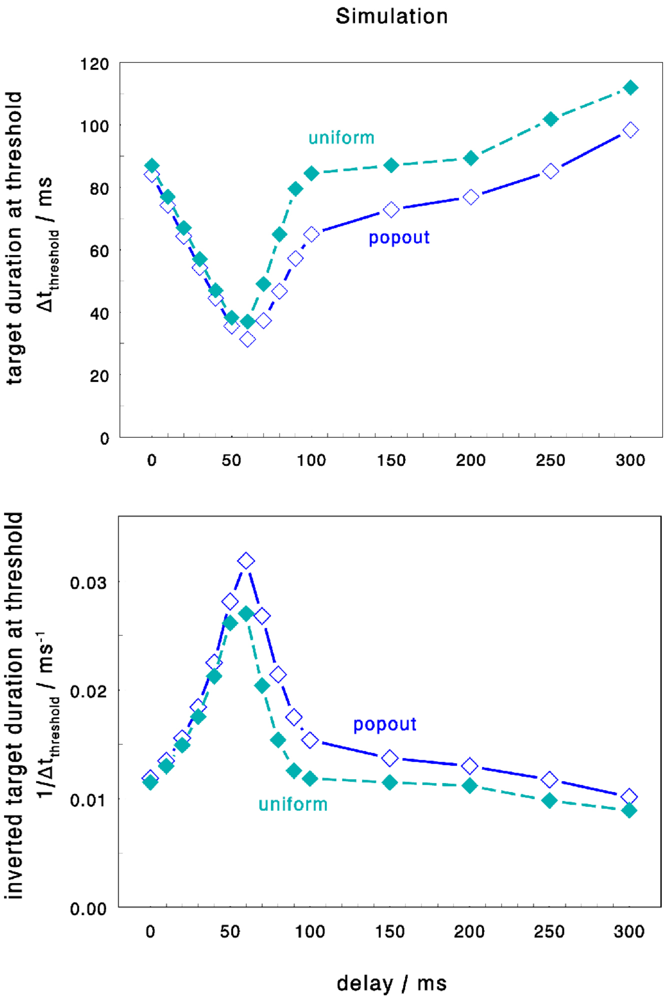


Figure 8. *Simulation.* The population response in Figure 1 (originally measured for 500 ms) was used to generate cumulative performance curves at different delays. An arbitrary threshold (resembling constant performance rates) was set, and the required cumulation time ("target duration") to reach this threshold was calculated. **a.** Target durations to reach the threshold vary in a characteristic manner with the increasing delay when accumulation starts. Durations are particularly short when response accumulation starts in the response peak (delay 60 ms; cf. Fig. 1) and increase at shorter or longer cue delays. **b.** Inverse target durations at threshold coarsely resemble the peak characteristics of the original response (for details, see text).

response differences between uniform and popout targets but shows differences in the required accumulation time until targets in either configuration would have generated a similarly strong signal to make a decision about their orientation.

Since, in this simple model, the integration time to reach a constant signal is inversely related to the strength of the neuronal response at this moment, *t*,

response (t) • integration-time (t) = threshold,

⇒ response (t) = threshold / integration-time (t)

(cf. Nothdurft, 2017), we can compute the inverse value 1/integration-time (with the threshold set to 1) and obtain the presumably underlying response (Fig. 8b). In the experimental data, the result is uncertain with respect to the unknown signal threshold and necessarily "smeared out" the longer the integration-time has to be to reach this threshold.

Measured target durations at 75% correct identifications

To obtain analogue curves from the experimental data, I fitted the measured performance data of each subject at every tested cue delay with Gaussian cumulative functions and computed the presentation time at which subjects could identify half of the targets. The according performance level is 75% correct (between 50% for chance and 100% for perfect performance), and the associated presentation time was therefore labeled Δt_{75}.

Δt_{75} values are already indicated in Figure 6 (vertical lines); they reduce each rating curve to a single value that can then easily be compared between target conditions. Δt_{75} values directly reveal the performance differences at short and long cue delays (curves in different colors in Fig. 6) and between targets in uniform and popout configurations (Figs. 6a and b). The computation of Δt_{75} values made it occasionally necessary to include "negative" target durations when subjects had produced particularly high identification rates with short durations at certain cue delays (cf. subject HCN in Fig. 6b).

Δt_{75} variations with different target configurations

In the last part of analysis, I shall look at systematic performance differences over all tested delays. As mentioned above (and also visible in Figure 6), Δt_{75} values differed between subjects; to compare performance variations across subjects it is therefore necessary to look into details of the individual ratings. Figure 9a shows the measured Δt_{75} values of all five subjects; the first 500 ms are repeated in better resolution in Figure 9b. Inverted (1/

Δt_{75}) curves are shown in Figure 10; please notice the different scales. Despite considerable variations across subjects, certain characteristics are shared by all of them. Let us first look at the variations immediately after stimulus onset (Figs. 9b and 10b). The deepest inflections of Δt_{75} data in Figure 9 and strongest peaks in the $1/\Delta t_{75}$ curves in Figure 10 were obtained with popout targets; inflections are deeper (peaks in Fig. 10 larger) than those with uniform or border target configurations (except for subject PP).

In most subjects (but not LL and HCN), the deepest inflections reach the Δt_{75} values obtained with a single target cued at stimulus onset (dashed horizontal lines). This indicates that popout targets were identified about as fast as single targets. Only for subjects LL and HCN, the popout configuration further improved (and accelerated) the target identification.

With targets in border configurations, performance differences were less consistent. Three subjects (LL, TN, HCN) produced shorter Δt_{75} values for border than for uniform targets with at least some cue delays; in the remaining two subjects, these differences were small and variable (JP) or strong but reversed (PP). Performance ratings of subject PP were exceptional, as his Δt_{75} data do not reveal any systematic differences between popout and uniform targets; both target configurations were better seen than targets in border configurations.

Similar performance variations between target configurations were seen at longer cue delays (Figs. 9a and 10a). While four subjects produced a consistent difference between uniform and popout targets, with popout targets being faster identified

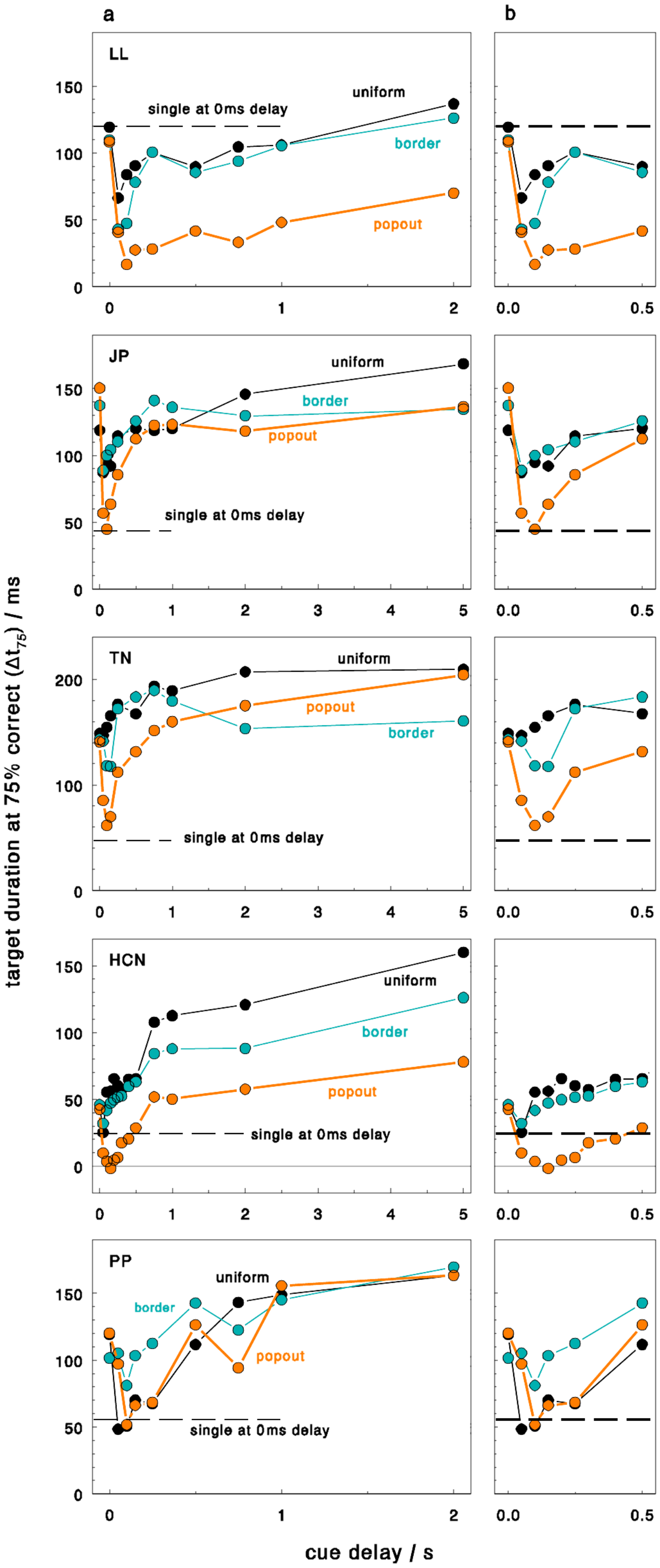
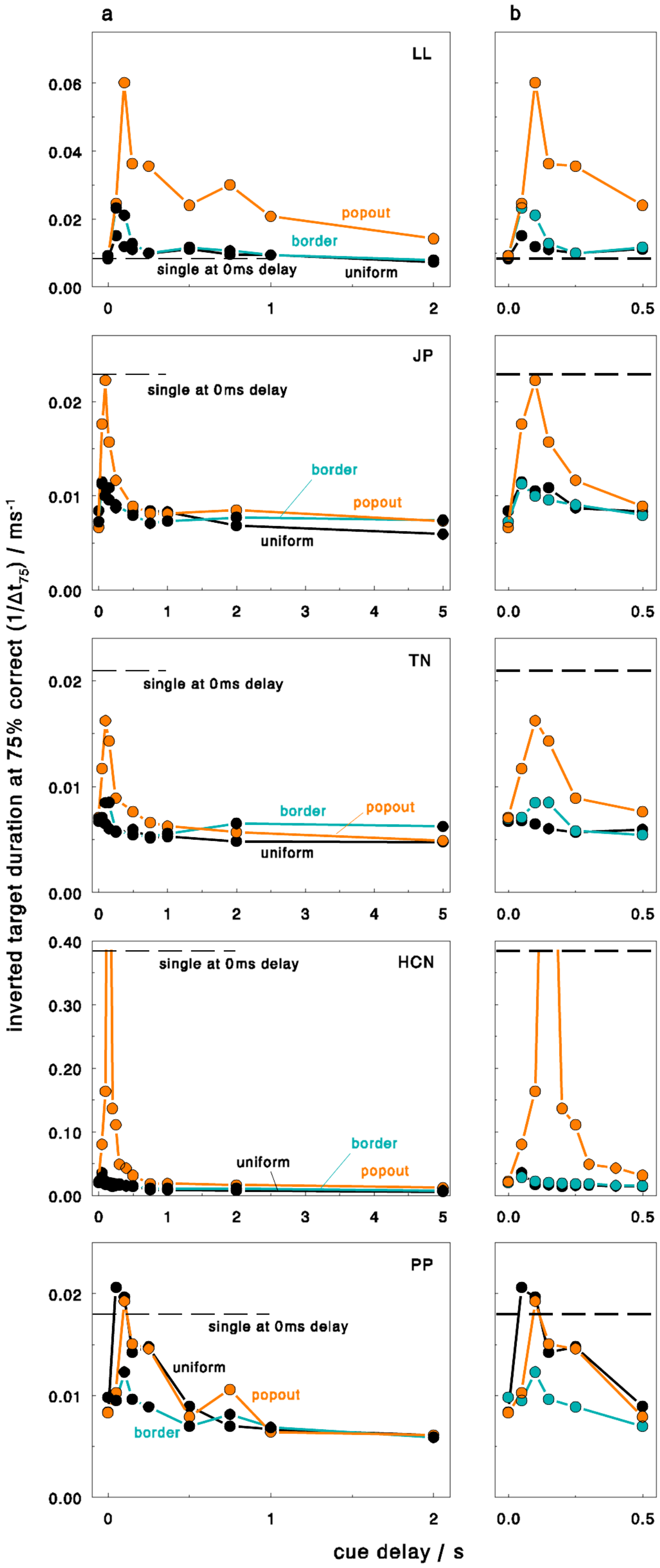


Figure 9. *Δt_{75} variations of all subjects (except MM) for uniform, border, and popout target configurations;* **a.** full range of tested cue delays, **b.** short cue delays enlarged. Δt_{75} values were obtained from Gaussian cumulative fits to individual data sets at every tested delay and for each target configuration (cf. Fig. 6). They are plotted in separate curves for popout, uniform and border targets. Δt_{75} values obtained with a single cued line are shown as black horizontal dashed lines. Curves show similar and diverging characteristics across subjects (for details, see text). Most subjects (except PP) required shorter presentation times (Δt_{75} values) to identify targets in popout than targets in uniform or border configurations. Please note the different scales.



than uniform targets (smaller Δt_{75} values), the ranking of uniform and border configurations was not consistent. Only subject HCN (who has carried out the most runs and the largest number of test repetitions of all subjects in the study) showed the expected ranking between uniform, border, and popout target configurations at all cue delays (except 0 ms and 50 ms).

The individual variations are smoothed in the means (Fig. 11). The differences between uniform and popout configurations are consistent and strong at all cue delays beyond 0 ms; Δt_{75} and $1/\Delta t_{75}$ values of the "in between" border configurations (less suppression from the surround than in the popout configuration) are "in between", too, and particularly at short cue delays not notably different from the uniform condition.

This general pattern is confirmed in the *statistical data analysis*. While in the means, the differences between Δt_{75} values of uniform and popout configurations are highly significant ($p<0.0005$; paired two-sided t-test) and even improve when the two shortest cue delays (0 ms and 50 ms) are excluded, the differences between uniform and border configurations are not significant ($p>0.17$). In the individual data, however, the pattern looks different. Uniform vs. popout differences were significant for only three subjects ($p<0.00001$, for LL and HCN, without the data from 0 ms and 50 ms delays; $p<0.005$, for subject TN) and just significant for subject JP ($p<0.05$). Differences between uniform and border configurations were only weakly significant ($p<0.05$) for subjects LL, TN, and HCN, and not significant for subject JP ($p>0.6$). Data from subject PP showed no significant differences at all between the uniform and other target configurations.

Altogether, the data confirmed the expectations based on neural response differences in area V1. Cued targets in popout configurations were better seen and faster identified than cued targets in uniform configurations. These differences were established in several analyses and were highly significant for four

Figure 10. *Inverted ($1/\Delta t_{75}$) computations of the data in Figure 9;* **a.** full range, **b.** short range enlarged. Note the different scales. The $1/\Delta t_{75}$ transform of the negative value from subject HCN (at delay 150 ms) was set to 1 (see text).

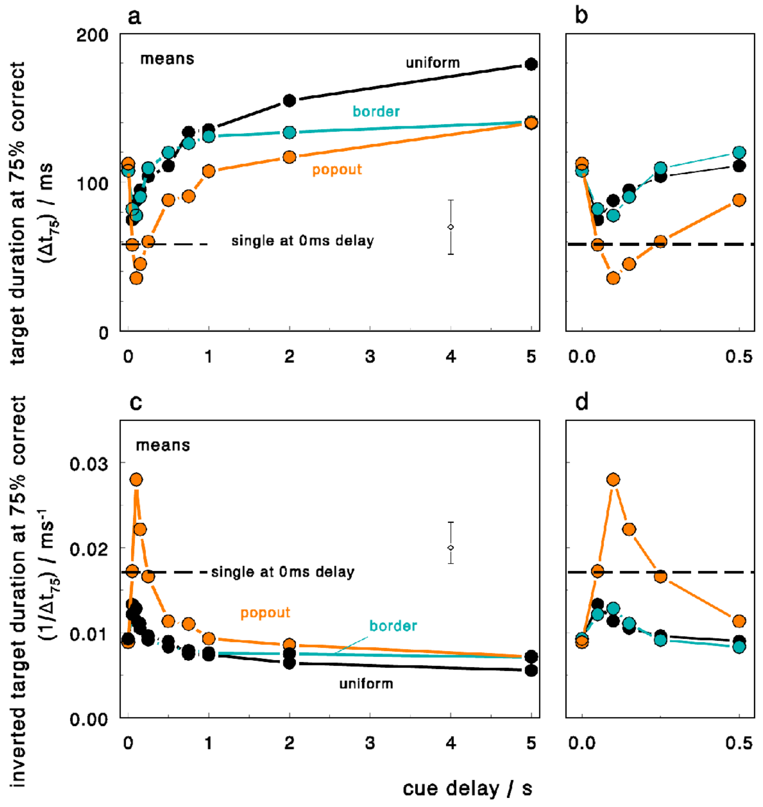


Figure 11. *Means of the data in Figure 9.* **a., b.** Means of individual target durations at 75% correct (Δt_{75} values) plotted over the full range of tested delays *(a)* and over a short range with enlarged resolution *(b)*. For readability, the s.e.m. of individual data points is averaged (not plotted in *(b)*). **c., d.** Inverted $1/\Delta t_{75}$ transforms of the mean data in *(a)*, plotted over the full delay range *(c)* or a short range enlarged *(d)*. For readability, the s.e.m. of individual data points is averaged; inverted transforms of the means+sem and means-sem in *(a)* generate an asymmetric s.e.m. in *(c)*. Presentation as in Figures 9 and 10.

observers. The differences between uniform and border targets, however, were less pronounced, did not consistently occur, and could be verified in only three subjects.

Failures

Two of the originally six subjects (MM and PP) did not produce the expected response variations. In one (PP), the Δt_{75} differences between uniform and popout conditions were small and inconsistent, although the subject had, on average, identified targets in popout conditions better than targets in uniform conditions, as the scatter analysis had shown (Fig. 3a). The other subject (MM) had produced responses properties opposite to those of all other subjects and revealed better target identification with targets in uniform than in popout configurations (cf. Fig. 3b). Since she was a highly motivated subject in the experiment, the reason for this deviation was likely not simple disinterest or inattention, which should anyway have led to many incorrect responses but not to a systematic deviation in the opposite direction. Since she had confirmed that she could see and locate the cues, one (hypothetical) explanation might be that her performance was affected by spatial interference between neighboring lines ("crowding"). If she had difficulties to quickly identify targets that were accompanied by *different* lines nearby, she might have often been wrong in popout or border conditions, but should have better identified targets in uniform configurations in which all neighboring lines share the target's orientation. This should have let to performance ratings in an opposite ranking to what was expected, which was indeed the case (Fig. 12; see also Fig. 3b).

Modification of experiment: reduced target eccentricity

To overcome possible strong crowding effects, the experiment was modified and possible target locations were restricted to fixation-near positions (within the inner dashed frame in Fig. 2b). This restriction had the

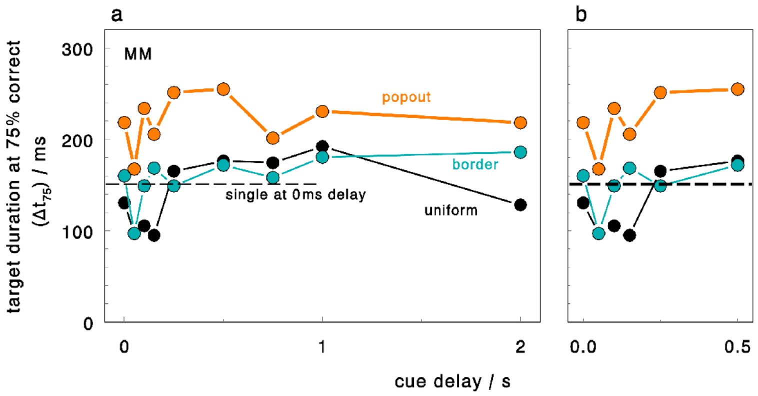


Figure 12. *Performance of subject MM.* Presentation as in Figure 9. As already seen in Figure 3b, the subject systematically identified uniform targets better than popout targets, contrary to the performance of most other subjects, and thus required longer target presentations to identify popout than uniform targets. This should be expected if the subject could not identify targets that were too closely surrounded by different lines (crowding). In a modified version of the experiment, cues and targets were therefore presented closer to the fixation point so that crowding effects were reduced.

disadvantage that the target was not entirely surrounded by lines, since the area of the fixation cross now became part of the surround. For that reason, testing was restricted to (incomplete) uniform and popout conditions but did not include border configurations. The range of tested cue delays was slightly reduced; all other test conditions were identical to those in the main experiment.

The modified experiment was carried out by altogether three subjects, subject MM (who had produced diverging scatter data in Fig. 3), subject PP (who did not show consistent performance variations in the Δt_{75} test of the main experiment), and a new subject, OC (22, male).

Figure 13 shows the scatter data of different target conditions for each subject; as before, only data are included for which the uniform and popout target configurations were tested with identical durations. As is obvious from the comparison of Figures 3b and 13a, performance of subject MM had notably changed in the new experiment. While she had better identified uniform than popout targets in the main experiment, she now revealed the expected bias for popout targets when targets were presented closer to the fixation cross and target eccentricity was reduced. The mean deviation changed from –7.14% ± 1.2% (preference for uniform targets) in Figure 3b to 3.28% ± 1.5% (preference for popout targets) in Figure 13a. Scatter data of the other two subjects are shown in Figure 13b and c. Mean deviations are 3.48 ± 1.25 % (PP) and 2.66% ± 1.11% (OC), resulting in a total mean deviation of 3.11% ± 0.76% for all three subjects. Although the biases look less spectacular than those in Figure 3a, all deviations are significant (paired two-sided t-tests; $p<0.05$, for subjects MM and OC;

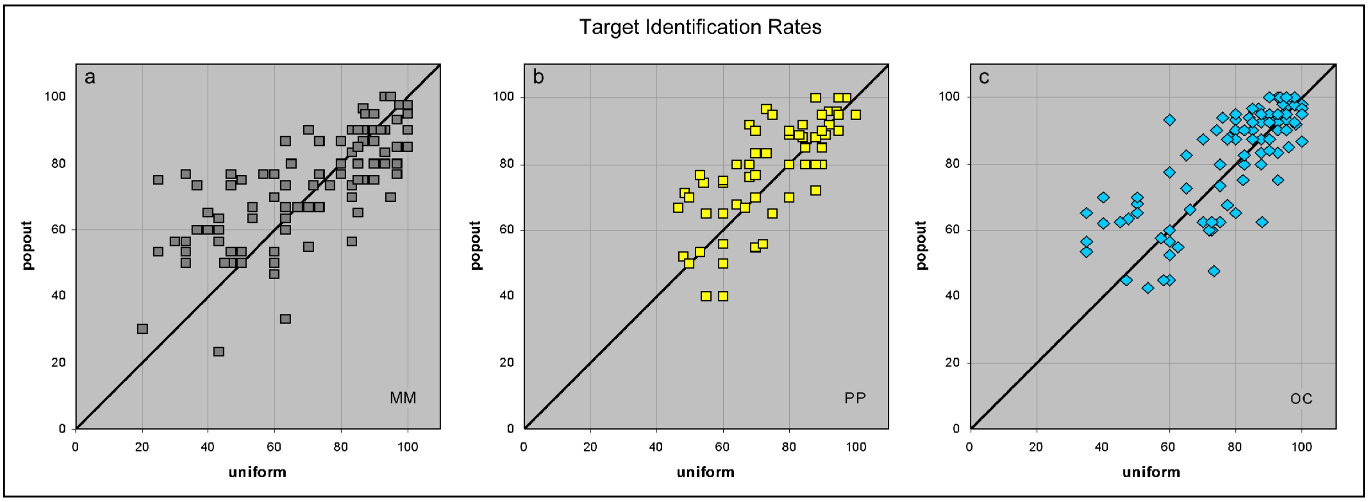


Figure 13. *Scatter diagrams of subjects tested in the modified experiment.* **a.** subject MM; **b.** subject PP, **c.** subject OC. The preference for popout over uniform targets was less strong than in Figure 3a but was seen with all subjects, now also with subject MM.

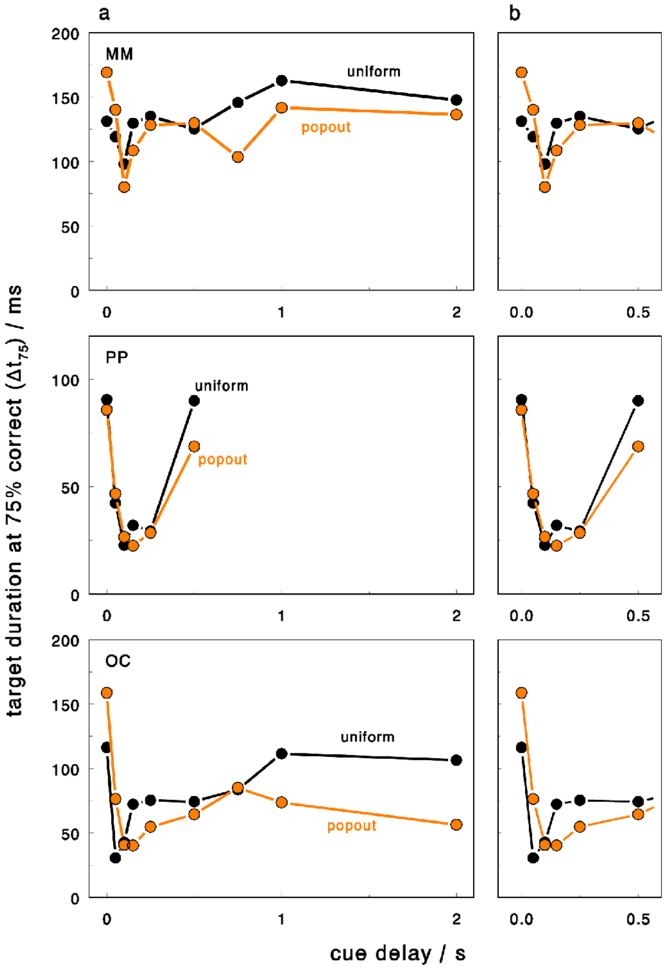


Figure 14. *Δt_{75} variations in the modified experiment;* **a.** full presentation of tested cue delays, **b.** data from short cue delays enlarged. Presentation is similar to that in Figure 9. Single lines and targets in border configurations were not tested. Curves show consistent rankings at certain delays (when tested).

$p<0.01$, for subject PP; and *$p<0.0001$*, for the total sample, N=274).

Figures 14 and 15 show the Δt_{75} and $1/\Delta t_{75}$ values of all subjects tested in the modified experiment. While the curves from uniform and popout target conditions are generally less distinct than in the best examples of Figure 9, certain differences are consistently found in all subjects. For example, at delays larger than 50-100 ms (3rd to 4th data points in Figs. 14b and 15b), popout targets generally required shorter presentation times than uniform targets, for 75% correct identification. Only a few data points from individual subjects did not follow this ranking (e.g., 500 ms, subject MM; 750 ms, subject OC). Furthermore, the ranking seems to be consistently reversed at short delays (<100 ms).

This general pattern is also revealed in the mean data of all three subjects (Fig. 16). The curves reflect the predictions made in Figure 8. The differences between Δt_{75} values for uniform and popout targets are highly significant in the means (paired, two-sided t-test, *$p<0.001$;* cf. Fig.16). In individual subjects (Fig. 14), however, only the performance differences of subjects MM and OC reached weak statistical reliability (*$p<0.05$*) when analysis was restricted to certain cue delays (≥100 ms).

Altogether, the modified experiment showed that even subjects who failed to produce the expected variations in the main experiment did reveal such variations in a simpler task with reduced target eccentricity.

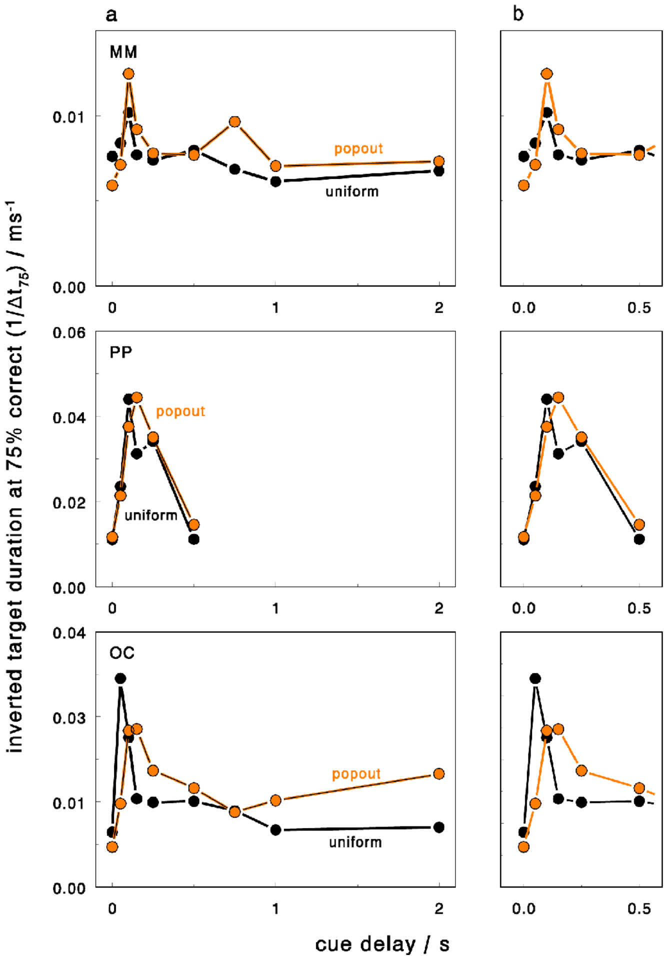


Figure 15. *Inverted ($1/\Delta t_{75}$) computations of the data in Figure 14.* Similar presentation as in Figure 10 (but without border cases).

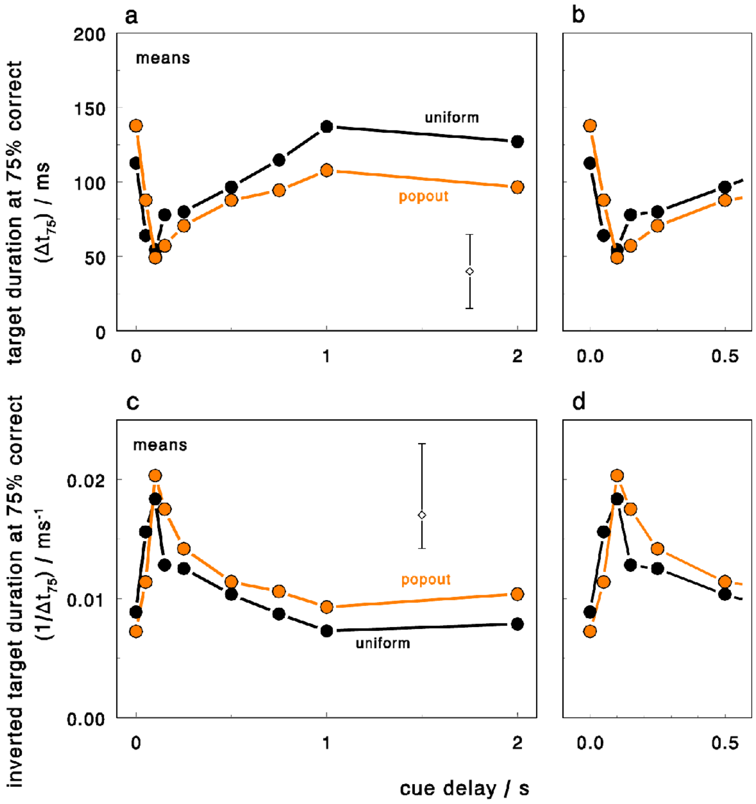


Figure 16. *Mean data from the modified experiment;* **a., b.** means of the Δt75 values in Figure 14, plotted over the full range of tested delays *(a)* and over a short range with enlarged resolution *(b)*. **c., d.** Inverted 1/Δt75 transforms of the mean data in *(a)*, plotted over the full delay range *(c)* or a short range enlarged *(d)*. Presentation as in Figure 11; error bars show the mean s.e.m. (not plotted in *b*). For cue delays at and above 100 ms, curves reveal systematic preferences (shorter Δt75 values) for popout than for uniform target configurations. Note that means from cue delays ≥750 ms are from two subjects only.

DISCUSSION

The study has shown that different target configurations are indeed distinguished in visual analysis even when observers are not aware of these differences. Targets in popout configuration, displaying local orientation contrast to their neighbors, were faster identified than targets in uniform configurations where neighboring lines share the target orientation and notably suppress the neuronal responses. Targets in border configurations represent an intermediate step between these two extremes, and also showed intermediate performance in the tests. Interestingly, the observed performance variations did in many details reflect the predictions made on population responses in area V1. The differences needed time to develop but then could be seen with similar strength over cue delays of up to 5 seconds. While identification rates with uniform and popout targets differed strongly in all seven tested subjects (although with some of them only at fovea-near locations), the differences between uniform targets and targets in border configurations were generally smaller and became statistically significant in only three of six tested subjects. Orientation contrast in border conditions is reduced compared to the orientation contrast in popout conditions (3 of 8 neighboring lines versus 8 of 8 neighboring lines), but still has been reported to evoke notable response differences in neurons in areas V1 and V2 (Nothdurft, Gallant, & Van Essen, 2000; Rossi, Desimone, Ungerleider, 2001; Marcus & Van Essen, 2002). In fact, given the vivid responses of cortical neurons to texture borders (Lamme, 1995; Zipser, Lamme, & Schiller, 1996; Lee, Mumford, Romero, & Lamme, 1998; Lamme, Rodriguez-Rodriguez, & Spekreijse, 1999) I was a little surprised that border configurations had not generated stronger and more consistent differences to uniform target configurations in the present experiments. But one must keep in mind, that the tested border-like target configurations in the present study did not generate true perceptual impressions of global borders in the test patterns (cf. Fig.2). The labels *uniform*, *popout*, and *border* did only refer to the local neighborhood of lines

ISSN:2364-3641

immediately around the cued target, which itself was embedded in a random line pattern outside that configuration. It would be interesting to see if uniform vs. border-like target configurations might lead to stronger differences in target identification when borders are continued over a wider distance than tested here.

Altogether, the results confirm the expectations put into the method of cued visual selection (Nothdurft, 2017) as well as into the predictions made in the Introduction. CVS provided a useful way to look at neural activity and to document differences that are perhaps not recognized perceptually. Beyond of that, the experiments showed that there are psychophysically measurable differences between uniform and popout target configurations, which, apart from associated saliency effects, had so far only been documented in neurophysiological recordings. The speed of target identification in these configurations was directly explained from response differences in neural populations between these patterns.

The locus of extracted information

While the measured performance variations seem to closely reflect certain neural response variations in area V1, we cannot be sure from which processing level the cued information is really taken and accumulated for the behavioral decision. The fact that performance improvements are delayed at cue delay 0 ms but quickly rise at cue delay 100 ms (Fig. 5) suggests the existence of a response latency and a subsequent transient peak in the relevant neural signal, as it was observed in neurons of area V1 (Fig. 1). However, the neural encoding of cues is also delayed and a too simple model of the relative timing of cues and V1 responses might not be adequate to explain the dynamics of perceived target identification. Furthermore, from the V1 responses in Figure 1 (obtained from anesthetized animals) we should expect best performances around 60-110 ms after stimulus onset, followed by a fast decay up to 120 ms, whereas the Δt_{75} data revealed strongest inflections around 100 ms to sometimes 300 ms. Of course, the measurements are made in different species and neural signals must first be accumulated to reach 75% performance, which may partially compensate for this difference. But why should CVS performance be directly related to neural activity in V1? Context-modulated responses have been seen in several visual areas, e.g. in V4 (Burrows & Moore, 2009), and many other visual areas are directly or indirectly driven by signals from area V1, Given the short delays in neural processing cascades, we should expect similar, or only little postponed performance data if target identification were based on the neural responses of other visual areas. Latency differences between visual areas are small and generally too small to be resolved with the present CVS data.

It is however interesting to see that the performance of best observers fits closely to the time limits evaluated for orientation discrimination in area V1 (Berens *et al.*, 2012). Based on the population code obtained from recordings in behaving monkeys, the authors found that two orientations could already be discriminated after 30-80 ms, using the spike counts in time windows of only 30 ms. They also noticed that the readout of orientation is most accurate during the transient phase of the neural response. This is exactly what was predicted, and found in the present study. Similarly short accumulation times were, however, only found in particularly sensitive and highly trained subjects (HCN; see also RUB in Nothdurft, 2017). Most observes, even after some training, needed longer presentation times to produce reliable decisions.

Salience variations?

A target with orientation contrast (as in the popout configuration) is salient (cf. Nothdurft, 1991, 2006; Wolfe, 1998) and attracts attention (Joseph & Optican, 1996; Nothdurft, 1999, 2002a; Turatto & Galfano, 2000; Zenon, Ben Hamed, Duhamel, & Olivier, 2008; see also Treue, 2003) which may then fasten target analysis and identification (Nothdurft, 2002a, 2006; Töllner, Zehetleitner, Gramann, & Müller, 2011; Sundberg, Mitchell, Gawne, & Reynolds, 2012). This generates an interpretation problem with the present findings. One could argue that the faster identification of popout targets might have been due to their increased salience, which might then have faster attracted attention than other targets, in particular targets in uniform configurations. In other words, the observed differences in identification speed might reflect differences in target *salience* but not necessarily differences in the neural encoding. It is not easy to reject that argument, since salience itself has convincingly been related to the enhanced activity in the primary visual cortex (Li, 2002; Zhang, Zhaoping, Zhou, & Fang, 2012). Many labs (including mine) have shown that salience is generated from several visual properties, including orientation contrast and additional cues, and both may guide attention to the target (Nothdurft, 2002a; Turatto *et al.*, 2000; Zenon, Ben Hamed, Duhamel, &

Olivier, 2008). Furthermore, different salience effects may add (Nothdurft, 2000; Zehetleitner, Krummenacher, & Müller, 2009) so that the popout targets in the present study, being salient from both orientation contrast and the cue, should have been more salient than the uniform targets being salient only from the cue. But given the strength of differences between these targets in the identification rates of some observers in the present study, the popout-defined salience should have been huge, if performance differences were only due to differences in target salience. Cue-defined salience was identical for both target configurations. In every test pattern, however, there were several popout-like items (cf., e.g., Fig. 2a) which should then all have attracted attention and thus should have rendered target selection rather difficult (e.g., Huang & Pashler, 2005; Koch, Müller, & Zehetleitner, 2013; Liesefeld, Liesefeld, Töllner, & Müller, 2017). This was obviously not the case. Furthermore, salience from orientation contrast was constantly present in a test pattern and was not modulated in time, whereas the measured identification rates were strongly modulated, in synchrony to the assumed variations in neural responses. Last not least, subjects were not aware and did generally not recognize the difference between cued targets in uniform or popout configurations, opposed to what a very strong salience difference would predict.

Wang and colleagues (Wang, Chen, Yan, Zhaoping, & Li, 2015) have explicitly studied cue-induced response variations in V1 neurons of the monkey and have found a transient increase of neural activity after the onset of the response. Different to the present study, cues were presented *before* the target. For two reasons, however, their findings cannot account for the performance variations seen in the present work. First, Wang and colleagues report that their cue-induced responses quickly adapt; they disappeared when the monkey became familiar with that cue, and could only be evoked again with novel cues. In the present experiments, however, neither cues nor targets changed and observers were soon quite familiar with both. Second, and more important, the response modulations reported by Wang *et al.* were synchronized to the occurrence of the cue, whereas the Δt_{75} modulations in the present study occurred in synchrony with the onset of the test pattern. It is important to recall this difference. It underlines that CVS is not primarily measuring cuing effects in vision but is using the cues as a kind of perceptual window to study neuronal processes evoked by the stimulus.

Failures

I have no clear evidence yet why some subjects performed worse than others. While training obviously improved target identification rates for short target durations, I have not yet identified which part in the experimental procedure was mainly affected by exercise. All subjects reported right from the beginning of experiments that they were able to detect and localize the cues, but nevertheless failed to identify the cuedtargets when presentation time was too short. During the first sessions, this ability noticeably improved so that thereafter reliable measurements could be made. But the training process was not yet finished at that time, improvements seem to rise asymptotically, and small changes could still be seen after several sessions. However, since all test conditions were intermingled in the experiment, these minor improvements should have applied to all conditions in a similar way.

In the main experiment, one subject (MM) had produced performance ratings that surprisingly differed in the opposite direction to that of all other subjects. But when target eccentricity was reduced and cued targets were presented closer to the fixation point, performance improved and response differences switched into the expected direction. While cue locations in the main experiment were restricted to 3.0 - 4.2 deg in the periphery, this eccentricity in the used raster width of 1.5 deg should have been sufficient to generate spatial interference (crowding) from neighboring lines when targets had to be identified (Strasburger, Rentschler, & Jüttner, 2011). I have made a similar observation in my earlier CVS study, where I also had to restrict the range of possible target locations for one subject to obtain reliable performance (Nothdurft, 2017). While visual acuity of all subjects had been measured to be normal, these two subjects might have been particularly sensitive to crowding which was reduced by choosing targets located closer to the fixation point. Thus it seems that CVS may suffer from a difficulty of certain observers, and it should be helpful to find that out early in experiment.

In general, however, the CVS experiments reported here were not particularly difficult to perform but required considerable patience (also from the experimenter) and a constant level of concentration on the observers' side. It is important (and partly also the responsibility of the experimenter) to hold observers at this level throughout experiments. The present study included a rather large number of test conditions (that might be smaller in other studies), which all had to be intermixed and frequently repeated to obtain reliable and consistent results. That required many experimental sessions in which subjects might have been particularly alert, or sometimes perhaps a little drowsy from previous work. I assume that some variations in the reported results might reflect the possibility that the level of attentive concentration of some subjects could perhaps not be held constant over all sessions.

CONCLUSIONS

The method of cued visual selection (CVS) of targets in line arrays was used to look into visual processes in the brain. The identification of targets presented in different configurations has revealed behaviorally measured performance differences that were correlated to variations in neural activity. CVS has thus been proven to be a useful tool to link behavioral and perceptual observations with neural processes that have been identified in single cell recordings. Although behavioral testing can be rather extensive and time-consuming, as was the case in the present study, the results could not easily be obtained in another way and thus seemed worth that elaboration.

REFERENCES

Burrows, B.E., & Moore, T. (2009). Influence and limitations of popout in the selection of salient visual stimuli by area V4 neurons. *Journal of Neuroscience, 29 (48),* 15169-15177. doi: 10.1523/JNEUROSCI.3710-09.2009

Berens, P., Ecker, A.S., Cotton, R.J., Ma, W.J., Bethge, M., & Tolias, A.S. (2012). A fast and simple population code for orientation in primate V1. *Journal of Neuroscience, 32 (31),* 10618-10626. doi: 10.1523/JNEUROSCI.1335-12.2012

Bair, W., Cavanaugh, J.R., & Movshon, J.A. (2003). Time course and time-distance relationships for surround suppression in macaque V1 neurons. *Journal of Neuroscience, 23,* 7690-7701.

Fischer, B. (1987). The preparation of visually guided saccades. *Reviews of Physiology Biochemistry and Pharmacology, 106,* 1-35.

Fischer, B., Weber, H., Biscaldi, M., Aiple, F., Otto, P., & Stuhr, V. (1993). Separate populations of visually guided saccades in humans - reactions-time and amplitudes. *Experimental Brain Research, 92,* 528-541.

Gao, D., Mahadevan, V., & Vasconcelos, N. (2008). On the plausibility of the discriminant center-surround hypothesis for visual saliency. *Journal of Vision, 8(7)*:13, 1-18, http://journalofvision.org/8/7/13/, doi:10.1167/8.7.13.

Huang, L., & Pashler, H. (2005). Quantifying object salience by equating distractor effects. *Vision Research, 45 (14),* 1909-1920.

Joseph, J.S., & Optican, L.M. (1996). Involuntary attentional shifts due to orientation differences. *Perception and Psychophysics, 58,* 651-665.

Kastner, S., Nothdurft, H.C., & Pigarev, I.N. (1997). Neuronal correlates of pop-out in cat striate cortex. *Vision Research , 37,* 371-376.

Knierim, J.J. & Van Essen, D.C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology, 67,* 961-980.

Koch, A.I., Müller, H.J., & Zehetleitner, M. (2013). Distractors less salient than targets capture attention rather than producing non-spatial filtering costs. *Acta Psychologica, 144,* 61-72. doi: 10.1016/j.actpsy.2013.04.023

Lamme, V.A.F. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience, 15,* 1605-1615.

Lamme, V.A.F., Rodriguez-Rodriguez,V., & Spekreijse, H. (1999). Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of macaque monkey. *Cerebral Cortex, 9,* 406-413.

Lee, T.S., Mumford, D., Romero, R. & Lamme, V.A.F. (1998). The role of the primary visual cortex in higher level vision. *Vision Research, 38,* 2429-2454.

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences, 6,* 9-16.

Li, C.Y., & Li, W. (1994). Extensive integration field beyond the classical receptive field of cat's striate cortical neurons – classification and tuning properties. *Vision Research, 34,* 2337-2355.

Liesefeld, H.R., Liesefeld, A.M., Töllner, T., & Müller, H.J. (2017). Attentional capture in visual search: Capture and post-capture dynamics revealed by EEG. *NeuroImage, 156,* 166-173. doi: 10.1016/j.neuroimage.2017.05.016

Marcus, D.S., & Van Essen, D.C. (2002). Scene segmentation and attention in primate cortical areas V1 and V2. *Journal of Neurophysiology, 88,* 2648-2658. doi: 10.1152/jn.00916.2001.

Nelson, J.I., & Frost, B.J. (1978). Orientation-selective inhibition from beyond the classic visual receptive field. *Brain Research, 139,* 359-365.

Nothdurft, H.C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research, 31,* 1073-1078.

Nothdurft, H.C. (1992). Feature analysis and the role of similarity in pre-attentive vision. *Perception & Psychophysics, 52,* 355-375.

Nothdurft, H.C. (1993). The role of features in preattentive vision: comparison of orientation, motion, and color cues. *Vision Research, 33,* 1937-1958.

Nothdurft, H.C. (1997). Different approaches to the coding of visual segmentation. In: Jenkins, M., & Harris, L. (Eds.) *Computational and biological mechanisms of visual coding* (pp. 20–43). New York: Cambridge University Press.

Nothdurft, H.C. (1999). Focal attention in visual search. Vision *Research, 39,* 2305-2310.

Nothdurft, H.C. (2000). Salience from feature contrast: Additivity across dimensions. *Vision Research, 40,* 1183-1201.

Nothdurft, H.C. (2002a). Attention shifts to salient targets. *Vision Research, 42,* 1287-1306.

Nothdurft, H.C. (2002b). Latency effects in orientation popout. *Vision Research, 42,* 2259-2277.

Nothdurft, H.C. (2006). Salience and target selection in visual search. *Visual Cognition, 14(4-8),* 514-542. doi: 10.1080/13506280500194162

Nothdurft, H.C. (2015). The role of features in salience and grouping: luminance and disparity. *VPL-reports, 3,* 1-32, www.vpl-reports.de/3/

Nothdurft, H.C. (2016a). Spatial cue effects in Visual Selection. *VPL-reports*, *4,* 1-21, www.vpl-reports.de/4/

Nothdurft, H.C. (2017). Cued visual selection - a tool to study the dynamics of neural processes in perception? *VPL-reports, 6,* 1-24, www.vpl-reports.de/6/

Nothdurft, H.C., Gallant, J.L., & Van Essen, D.C. (1999). Response modulation by texture surround in primate area V1: Correlates of "popout" under anesthesia. *Visual Neuroscience, 16,* 15-34.

Nothdurft, H.C., Gallant, J.L., & Van Essen, D.C. (2000). Response profiles to texture border patterns in area V1. *Visual Neuroscience, 17,* 421-436.

Rossi, A.F., Desimone, R., & Ungerleider, L.G. (2001). Contextual modulation in primary visual cortex of macaques. *Journal of Neuroscience, 21 (5),* 1698-1709.

Sillito, A.M., Grieve, K.L., Jones, H.E., Cudeiro, J. & Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature, 378,* 492-496.

Smith, M.A., Bair, W., & Movshon, J.A. (2006). Dynamics of suppression in macaque primary visual cortex. *Journal of Neuroscience, 26 (18),* 4826-4834. doi: 10.1523/JNEUROSCI.5542-06.2006

Smith, M.A., Kelly, R.C., & Lee, T.S. (2007). Dynamics of response to perceptual pop-out stimuli in macaque V1. *Journal of Neurophysiology, 98,* 3436-3449. doi: 10.1152/jn.00441.2007

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision, 11(5):13,* 1–82, http://www.journalofvision.org/content/11/5/13, doi:10.1167/11.5.13.

Sundberg, K.A., Mitchell, J.F., Gawne, T.J., & Reynolds, J.H. (2012). Attention influences single unit and local field potential response latencies in visual cortical area V4. *Journal of Neuroscience, 32 (45),* 16040-16050. doi: 10.1523/JNEUROSCI.0489-12.2012

Töllner, T., Zehetleitner, M., Gramann, K., & Müller, H.J. (2011). Stimulus saliency modulates pre-attentive processing speed in human visual cortex. *PLoS ONE 6(1),* 1-8: e16276. doi: 10.1371/journal.pone.0016276

Treue, S. (2003). Visual attention: the where, what, how and why of saliency. *Current Opinion in Neurobiology, 13,* 428-432.

Turatto, M., & Galfano, G. (2000). Color, form and luminance capture attention in visual search. *Vision Research , 40,* 1639-1643.

Turatto, M., Benso, F., Facoetti, A., Galfano, G., Mascetti, G., & Umiltá, C. (2000). Automatic and voluntary focusing of attention. *Perception & Psychophysics, 62 (5),* 935-952.

Wang, F., Chen, M., Yan, Y., Zhaoping, L., & Li, W. (2015). Modulation of neuronal responses by exogenous attention in macaque primary visual cortex. *Journal of Neuroscience, 35 (39),* 13419-13429. doi: 10.1523/JNEUROSCI.0527-15.2015

Wolfe, J.M. (1998). Visual search. In: Pashler, H. (Ed.) *Attention* (pp. 13-73). Hove, UK: Psychology Press.

Zehetleitner, M., Krummenacher, J., & Müller, H.J. (2009). The detection of feature singletons defined in two dimensions is based on salience summation, rather than on serial exhaustive or interactive race architectures. *Attention, Perception, & Psychophysics, 71 (8),* 1739-1759. doi: 10.3758/APP.71.8.1739

Zenon, A., Ben Hamed, S., Duhamel, J.R., Olivier, E. (2008) Spatial and temporal dynamics of attentional guidance during inefficient visual search. *PLoS ONE 3(5):* e2219. doi: 10.1371/journal.pone.0002219

Zhang, X., Zhaoping, L., Zhou, T., & Fang, F. (2012). Neural activities in V1 create a bottom-up saliency map. *Neuron, 73,* 183-192. doi: 10.1016/j.neuron.2011.10.035

Zipser, K., Lamme, V.A.F. & Schiller, P.H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience, 16,* 7376-7389.