Feature contrast in salience and grouping: luminance and disparity

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

The paper expands an earlier study [Nothdurft, H.C. (1993b). The role of features in preattentive vision: comparison of orientation, motion, and color cues. *Vision Research 33(14)*, 1937-1958] into the visual dimensions luminance and depth. Two perceptual phenomena were tested; (1) the ability of observers to detect a single salient target that differs from neighboring items (measurements on salience), and (2) the ability of observers to group salient items to larger figures (measurements on grouping). In each experiment, feature variations were restricted to the studied dimension, here luminance or disparity. Grouping processes were also studied with a monocular depth cue, depth from shadow. All experiments support the earlier notion that salience occurs from feature differences, not features themselves. Even for grouping, feature differences (which make a target salient) are more important than feature identities, at least in short stimulus presentations. *Target salience:* Like with orientation, motion and color, small differences in luminance or disparity are sufficient to make a target stand out from an array of identical items. But when background items themselves vary, target differences must be increased to be detected. With large background variations, even a strong feature contrast may not be sufficient to make the target stand out. *Grouping:* In brief inspections, observers group targets for salience, not for similarity. In the present experiments, however, certain modifications were necessary to make categorically different targets (bright – dark; near – far) equal salient. Apparent depth-from-shadow improved the grouping over that of rotated items with no such depth impression, but a similar difference was found when luminance gradients were replaced by luminance steps that do not resemble shadow or depth.

Published online: 15-Dec-2015

Citation: Nothdurft, H.C. (2015). Feature contrast in salience and grouping: luminance and disparity. VPL-reports, 3, 1-32, www.vpl-reports.de/3/

INTRODUCTION

The important role of salience in the detection of targets (visual search) and borders (texture segmentation) has been frequently demonstrated over the last 50 years (e.g., Beck, 1966, 1982; Julesz, 1962, 1975, 1981; Nothdurft, 2006; Olson & Attneave, 1970; Treisman & Gelade, 1980; Wolfe, Cave, & Franzel, 1989; see also Wolfe, 1998), although the term "salience" was not used in the early studies and it had not always been clear what exactly would make a target or border stand out perceptually. While early theories (Beck, 1982; Julesz, 1975, 1981; Treisman, 1985) had proposed the representation and immediate detection of certain key features, later studies underlined the role of local differences, i.e. *feature contrast* (e.g., Julesz, 1986; Landy & Bergen, 1991; Nothdurft, 1985; Sagi & Julesz, 1987). This was extensively studied and illustrated with oriented lines, where the same feature (e.g., a vertical line) could be made salient or non-salient, and hence shown to pop out or not, depending on how strongly it differed from other items nearby (Nothdurft, 1991, 1992). Also for segmentation, the orientation contrast across a border turned out to be far more important than any similarity of items within regions. In fact, non-segmenting regions with statistically identical distributions of oriented lines could be made to segregate when items were simply re-arranged so that the orientation differences were locally enlarged

(Landy & Bergen, 1991; Nothdurft, 1985; see also Fig. 2.3 in Nothdurft, 1997).

However, the preference for feature differences rather than features themselves (but see, for example, Foster & Westland, 1995) is in conflict with a major proposal of Gestalt psychology according to which one important principle of perceptual organization is the detection and grouping of *similar* objects (e.g., Wertheimer, 1923). That similarity sometimes cannot explain the perceived segmentation or grouping of visual objects has, in fact, been pointed out very early (Beck, 1966). To study the role of orientation vs. orientation contrast in grouping, I explicitly searched for performance differences when subjects had to detect global figures of similar or dissimilar items (Nothdurft, 1992, 1993b). When patterns were briefly presented, subjects could not distinguish between these cases and grouped items just for their salience, irrespective of the features they displayed. Short presentations were essential in these tests, as observers can easily tell apart similar and dissimilar items under long enough inspection time. But the experiment showed that grouping does not require target similarity and that target similarity does not help to group different targets in fast perceptual organization.

The experiments had revealed another interesting property. When the background of items in a pattern is uniform (all non-target lines are parallel), a relatively small orientation difference would be sufficient to make a target or border stand out. But when the non-target lines themselves vary in orientation, e.g., when there is a continuous orientation gradient all over the pattern, the threshold orientation contrast to make a target or border salient is increased. With increasing background variation the threshold orientation contrast increases continuously– until finally, at a background variation of about 30° between neighboring lines, even a maximal orientation step of 90° is not sufficient to let the target perceptually stand out (Nothdurft 1992, 1993c). There are different hypothetical explanations of this phenomenon. One is that only differences of up to 30° between neighboring lines may generate the percept of continuous orientation flow in the background, on which disruptions could then be detected (cf. Field, Hayes, & Hess, 1993; see also Ben-Shahar, 2006). Another hypothesis would be that an orientation contrast of 30º is about as salient as the maximal orientation contrast of 90º, so that on such a background even targets with maximal orientation contrast cannot be more salient than the background items themselves, and hence do not stand out. The latter explanation was, in fact, experimentally confirmed (Motoyoshi & Nishida, 2001; Nothdurft, 1993c) and receives further support from the contextual modulation of neural responses to oriented lines in the primary cortex (Kastner, Nothdurft, & Pigarev, 1997; Knierim & Van Essen, 1992; Li & Li, 1994; Nelson & Frost, 1978; Nothdurft, Gallant, & Van Essen, 1999; Sillito et al., 1995; Zipser, Lamme, & Schiller, 1996). When these neural properties are implemented in models, several observations of popout and texture segmentation can be replicated (Gao, Mahedevan, & Vasconcelos, 2008; Li, 1999, 2002; Nothdurft, 1997).

While the first experiments on feature contrast (beyond luminance contrast) were made on orientation, which is particularly easy to implement, the question about the role of feature contrast in perception was soon extended to other visual features. In several follow-up studies I have transferred the orientation experiments to other feature domains. Four perceptual phenomena were tested if possible; (i) the effect of background variations on the apparent salience of a single target or (ii) a segmenting border, (iii) the role of target contrast on the search characteristics for a given target, and (iv) the ability of subjects to group items by feature identity. Some of these experiments, on color and the direction of motion, have already been published (Nothdurft, 1993b).They showed partly similar and partly different properties to those observed with orientation. In both dimensions, local differences are sufficient to make targets and borders salient, and like in orientation, these differences must be increased when the background is varied, too. In *search*, however, color contrast though helpful for localizing the target seemed to be a less exclusive criterion for fast target detection. Subjects detected a red blob similarly fast when embedded in a green (large hue contrast to neighbors) or yellow-red gradient field (smaller hue contrast to neighbors). This was different to orientation, where subjects needed quite a while to find a vertical target that was not very distinct from its neighbors. However, search performance in color strongly depends on the axis of color differences (D'Zmura, 1991; Nagy & Sanchez, 1990) and tests might not have been optimal to produce a similar effect in color in my experiments. In motion, the search paradigm was not tested, since small movements in a large array of dots created a strong percept of relative movement, so that it was not possible to ask subjects to search for a line moving exactly in a certain direction. In the *grouping* experiments, finally, motion differences generated data very similar to those found for orientation (subjects failed to group targets for similarity), but color targets were not exclusively grouped by contrast but partly also by identity. This was not seen with orientation or motion. Altogether the findings suggested that feature contrast is important for salience in all these domains but that the identification of target properties in the color domain is perhaps not entirely masked by color contrast.

Similar experiments had also been performed on luminance and disparity but so far only been published as a conference abstract (Nothdurft, 1995). These data are now presented here. Experiments addressed only two phenomena of the above list; (i) the effect of background variations on the threshold contrast of a salient target; and (iv) the ability of subjects to group items by similar identity. The (iii) role of target features in search was not studied, for the same reason why it had not been studied with motion direction in the previous study. Both luminance and disparity lack an absolute internal reference. While we can easily search for a vertical line or a red blob, we cannot search for a blob with a certain luminance or occurring at a certain disparity unless a reference is given for comparison. However, a comparison target would confuse the search time characteristics. Note that the situation is different to that for motion. There, we do have an internal reference and should be able to search for, e.g., a target moving exactly upwards or downwards or horizontally to the right or left. But the overall movement of items in the pattern rendered such an estimate extremely difficult, if movements had small amplitudes and dislocation cues could not be used (Nothdurft, 1993b).

Instead, grouping experiments in disparity were anecdotally extended to a monocular depth cue based on the direction of sunlight and the resulting differences in the shadows of concave and convex objects (Ramachandran, 1988; Todd & Mingolla, 1983).

GENERAL METHODS

Overview

The experiments reported here were part of a larger study, some experiments of which have already been published (e.g., Nothdurft, 1992, 1993a-c). Stimuli displayed regular or nearly regular arrays of items that varied in certain properties. The present experiments were based on two major tasks (cf. Figs. 1 and 8). In the *target detection task* (also referred to as *"popout",* like in the earlier study) subjects had to detect a salient item and indicate its location (left or right) relative to the fixation point in the middle of the screen. Aim of these experiments was to estimate the threshold feature contrast under which a target on a given background could reliably be detected. In the *grouping task* subjects were asked to

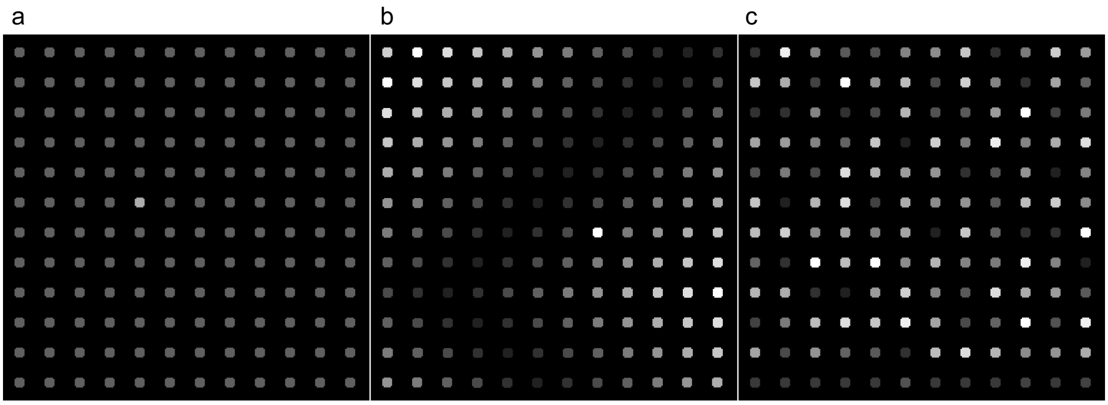

Figure 1. *Examples of stimulus patterns in Experiment 1 (LUM popout).* **a.-c.** Stimuli represented a regular raster of bright blobs on dark background, among which a salient blob at higher luminance contrast was to be detected. Background blobs could be homogeneous *(a)* or display continuous luminance variation between blobs *(b)*. After presentation the stimulus pattern was masked *(c)*. Luminance settings are modified to illustrate target and background variations.

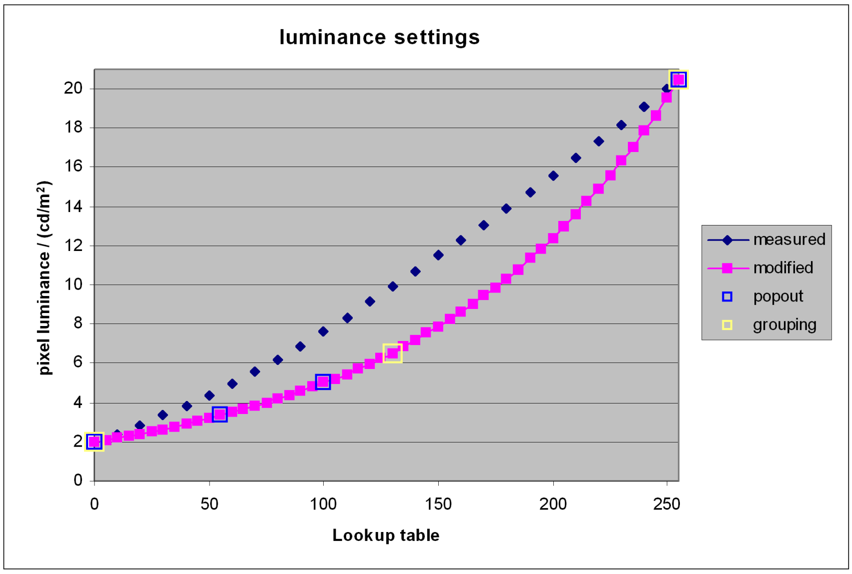

Figure 2. *Luminance settings as used in the LUM experiments.* By re-definition of the lookup table, the measured luminance variations on the screen ("measured") were transformed into a Weber-constant (ΔI/I=constant) luminance increase all over the available luminance range ("modified"). Luminance settings of background and reference targets in Experiment 1 ("popout") and Experiments 3-5 ("grouping") are indicated; see Methods for details.

group several salient items and indicate their global configuration. Aim of these experiments was to study the role of item similarity and feature identities in grouping. As in the previous study, all stimulus items in a particular experiment varied exclusively in the feature domain to be studied, here luminance or disparity; in all other dimensions the stimuli were uniform. That is, the items in the present study were all static (no motion), white or dark (no hue differences) circular blobs (no orientation) and all either appeared at the same depth and varied in brightness (luminance tasks, LUM) or had identical luminance and occurred at different disparities (depth tasks, DIS). Disparity variations were generated by the alternating presentation of different patterns to the observer's eyes. There was another set of stimuli with monocular depth cues (depth from shadow, RAMA) that was tested in an anecdotal addition of the grouping task.

Stimuli

All stimuli were computer generated (LSI-11), using 50 Hz non-interlaced video technique, and displayed on a monitor screen (CONRAC 7211) 2 meters in front of the observer. Patterns were made of blobs (12.7' diameter) that were arranged in a regular or nearly regular fashion across the monitor (7.7 deg x 7.7 deg). Three different raster configurations were used in the experiments, 12 x 12 (luminance popout), 8 x 8 (disparity popout), and 11 x 11 (all grouping experiments) While an exactly regular raster was used for studying feature variations in the LUM domain (raster widths 38' and 41.5', respectively), the blob positions were slightly varied (up to +/- 10.8') in the DIS patterns (raster widths 41.5' and 57.8', respectively) to avoid monocular cues associated with disparity shifts in the two patterns. Experiments on monocular depth cues (RAMA grouping) were performed on a similar jitter (raster width 41.5'). All items and backgrounds in the study were achromatic.

Luminance variations (LUM)

Luminance settings from different computer values were carefully measured and transformed into settings with constant Δluminance/luminance increments (Fig. 2). Internal luminance values were automatically transformed into pixel luminance by means of lookup-tables. Two different luminance settings were used. In the *target detection task* (LUM popout), screen background was 2 cd/m^2 and item luminance varied between 3.3 and 20.4 cd/m^2; that is, all items were brighter than background. In the *grouping task* (LUM grouping), patterns were made to display bright and dark targets, which was achieved by increasing screen background luminance to 6.5 cd/m^2; item luminance varied between 2.0 and 20.4 cd/m^2. Examples of stimuli are shown in Figures 1 and 8.

Disparity variations (DIS)

The smallest disparity that could be produced in the setup was 1.8 minutes of arc (1.8' = 1 pixel shift between the two patterns), a value well above the disparity thresholds of all observers in these experiments (see below). With a pupil distance of 75 mm between the two eyes and the given distance from the monitor (2 m) that would correspond to a depth shift of 2.85 cm (near) and 2.93 cm (far). Disparity variations in the patterns were multiples of this step. All items had a luminance of 20.4 cd/m^2 on 2 cd/m^2 background, but flickering glasses reduced these values to about 10%. Only bright blobs were used. Disparity variations in the DIS patterns were, in principle, comparable to the luminance variations in LUM patterns (cf. Fig. 1 and 8).

Apparent depth cues (RAMA)

Items were circular blobs (27.1' diameter) with luminance gradients that could be rotated (cf. Fig. 17). At certain orientations such elements give the vivid impression of bumps and holes (Kleffner & Ramachandran, 1992; Ramachandran, 1988; Todd & Mingolla, 1983). Blobs with the lighter halves above appear to stand out from the plane (convex); blobs with the lighter halves down are perceived as holes (concave).

Target arrangements

In the *target detection tasks*, the threshold feature contrast of a particularly salient target was measured as a function of the overall feature variation between neighboring items in the background. For that, all blobs were constructed to display continuous feature gradients across the pattern. To avoid too simple figures, gradients were sometimes reversed to change a continuous increment into a decrement of the same size, and *vice versa* (cf. Fig. 1). Gradients reversed automatically at maximum or minimum feature settings (not for disparity variations). Systematic background variations of this sort were generated in luminance, for the LUM experiments, and in disparity, for the DIS experiments. In the RAMA experiments, only selected items, no feature gradients were used in the backgrounds. On top of the feature backgrounds, single targets with an increased feature contrast were generated by adding this higher contrast to the value of a virtual background item at this location. To ensure however, that targets were not added to local peaks of background variation, feature gradients were not reversed at, or in the immediate neighborhood of target locations. For each background gradient, target feature contrast was systematically varied to find the 75% detection rate. All targets were presented on the same feature level of items in the background and hence under similar conditions (5.0 cd/m^2 in LUM popout; zero disparity in DIS popout). This was achieved by taking into account the selected target position, the selected gradient reversal points, and the given background variation, when constructing the patterns. Targets occurred randomly in the left or right half of the pattern at an eccentricity of 1-2 deg, and subjects had to indicate on which side from the fixation point they had detected a salient target.

In the *grouping task*, three or four salient elements were shown each 2.1 deg away from the fixation point (Fig. 8). For DIS grouping, slightly smaller target configurations were used (1.4 deg away from the fixation point). If there were three salient elements, these formed a global triangle composed of identical or not identical elements *(same* or *similar* vs. *different* or *dissimilar).* If there were four salient elements, three of them were identical and formed a triangle; the fourth item was added for confusion. With all patterns, observers were asked to indicate the direction in which the global triangle was pointing. Configurations of similar targets were made from identical bright targets (LUM grouping) or targets occurring at the same convergent disparity (DIS grouping). Configurations of dissimilar targets were made from bright and dark targets (LUM grouping) or targets at convergent and divergent disparities (DIS grouping). For similar salience, these targets were all presented on the same item background level (6.5 cd/m^2 and zero disparity, respectively). This was achieved by using oblique and partly reversed feature gradients in the background, as shown in Figure 8.

While measurements in the target detection task were straightforward, measurements in grouping required some adjustments. On the one side, it had not been elucidative to use only targets that are easily detected and identified. In that case, target similarities could be recognized and reported even if they were not essential for the perceived grouping. On the other hand, it had also not been informative to use patterns in which the salient targets are generally not seen; any differences between similar and dissimilar items might then be invisible. To document performance variations between these extremes, test series were designed in which targets varied from just being detectable to being reliably seen. This was accomplished by the usage of non-uniform backgrounds (cf. Fig. 8a-c) and variable target contrast. Under these conditions, target salience increases gradually with target feature contrast, as measured in the target detection task. The test ranges were optimized in a number of initial tests. Good settings were obtained with moderate feature variations in the background (step 20 for luminance variations; step 1 for disparity variations) and continuous variations of the target feature contrast. A problem, however, were possible salience differences between "similar" and "dissimilar" items which were measured in additional experiments (see below). To compensate for such differences, the grouping task was repeated with stimulus patterns in which the feature contrast of the less salient target was enhanced. All these additional experiments will be explained in the text.

This general procedure was modified for the grouping of depth-from-shadow stimuli (RAMA grouping). Only the four-target conditions (three similar, one different) were tested. Targets were not varied in salience but constantly presented at maximal orientation contrast (cf. Fig. 17a, b). To vary visibility, presentation time was varied between trials. This had produced similar performance variations as did the variation of target contrast in the other grouping tasks. The possible advantage of apparent depth in grouping was tested by comparing the grouping of targets with strong depth cues with the grouping of other targets without such cues.

Subjects

The data presented in this study are from a pool of 13 students in the age of 21 to 33 years (8 female, 5 male), who were paid for the time they have spent in experiment. Most students participated in several but not all experiments. Some of them had also served as subjects in other experiments of the main project (e.g., Nothdurft, 1993a-c). All subjects had normal or corrected-to-normal visual acuity on both eyes and good stereo vision (Titmus and Randot Stereotests). The majority of them could detect and distinguish disparities of up to 20'' (seconds of arc). Only three subjects failed at this level but could reliably detect slightly larger disparities (30'', DN; 40'', GK, KR), which are still far below the smallest disparities tested in the experiments (nearly 2', see above). All subjects were given a thorough introduction and several runs for practicing before the measurements began. After that, subjects were familiar with the basically simple tasks, although not familiar with the aim of the experiments.

Procedures

Subjects were conveniently seated on a chair with a head holder that provided a constant distance from the monitor. All tests were performed binocularly. For disparity experiments (DIS), subjects wore glasses the transmission of which was blocked in alternation, in synchronization with the frame rate of the monitor. With slightly displaced stimuli in the two frames this produced a vivid impression of depth that all subjects could easily see. A few initial trials in every run allowed subjects to adjust to the disparity variations tested afterwards. In all experiments, subjects were asked to fixate a small mark on the screen which remained visible throughout the experiment; this was helpful for performing the tasks under the brief stimulus presentations used. Fixation performance was regularly controlled in initial runs, by watching the eyes of the observers, but only occasionally once subjects had adopted to perform the tasks under fixation.

Stimulus presentation was synchronized to the vertical blank of the video signal. Every trial began with a blank screen; after 1s the stimulus pattern was shown which was typically masked after 100-200ms, or after variable time intervals (RAMA grouping). Variations in the presentation time are discussed below. Masks contained random variations of item settings in the feature under study, that is, blobs of different luminance in the LUM experiments (e.g., Figs. 1c and 8d), blobs at different disparities in the DIS experiments, or items with random luminance gradients at orientations not tested in the current grouping task, in the depth-from-shadow experiments (RAMA grouping; e.g. Fig.17c). After each trial, subjects entered their response into a computer keyboard (keys "1" or "0" on the main part of the keyboard, for salient targets seen on the left or right side of the screen, or keys "4", "8", "6", or "2" on the numeric keyboard, for triangles pointing towards the left, top, right, or bottom). In rare occasions, subjects could reject a trial back to the pool of trials still to be presented, if they had been inattentive during the trial. After the response, data were stored and a new trial started.

Experiments were done in blocks of 100-300 stimulus presentations (for different runs) which usually included all various test conditions on one given background variation; different tasks (target detection or grouping) were performed in separate runs. Thus, e.g., for a given

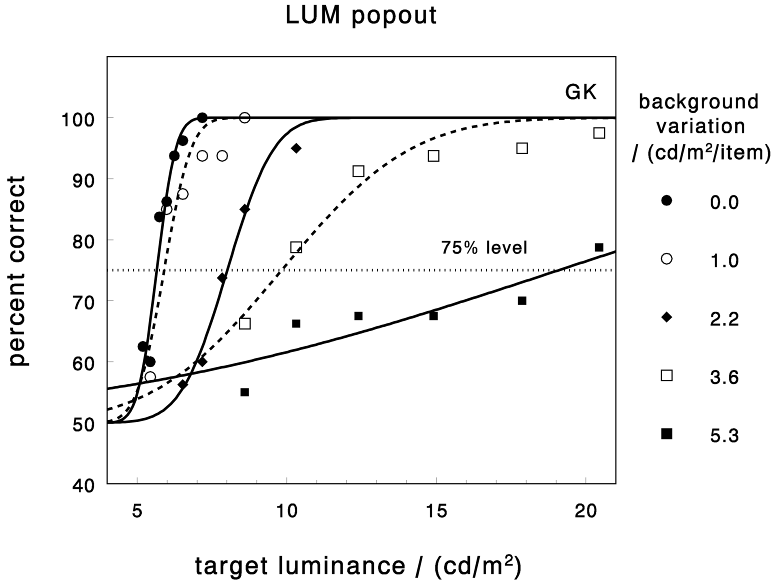

Figure 3. *Performance of subject GK in Experiment 1.* For different luminance gradients in the patterned background (different symbols and curves), detection rates of a target with locally increased luminance contrast are plotted and fitted by cumulative distribution functions. The 75% settings of each curve are taken as the threshold luminance under which a target on the according background variation is reliably detected. Values of background variation refer to increments at the target background level (5.0 cd/m^2).

target detection run, targets were displayed on the right or left side of the fixation point, location and target contrast varied from trial to trial, and so did the exact feature settings of background items, but the mean variation between items in the background had the same magnitude in all stimuli of the run. The run was repeated in other sessions. Background variations of different strength were tested in separate runs. For RAMA grouping, a single run covered all test patterns; presentation time was varied between runs. Several runs were carried out in one experimental session, but subjects could pause whenever they wanted before starting a new run. Experimental sessions were limited to two hours each, maximally one per day. Depending on the task and the subject's performance, up to seven sessions had to be made to finish an entire experiment (including run repetitions). All runs of an experiment were carried out within a few weeks.

Analysis

Performance was accumulated from all repetitions. In data from the *target detection tasks*, the target detection rate was calculated as a function of background variation and target feature contrast; it usually increased from about 50% (chance level) when target contrast was small relatively to background variation, to 100% (perfect detection rate) when target contrast was sufficiently large so that the target was always seen. With large background variations, however, the 100% performance level was often not reached. The data for a various background variations were fitted by nonlinear regression with a cumulative Gauss distribution function, $y=50+25*(1+erf((x-a0)/(a1\sqrt{2})))$, from which the 75% detection rates (a0) were taken, which represent the center of the distribution. These values vary with background variation, as shown below. For data from the *grouping task*, performance with increasing target contrast (LUM or DIS) or increasing presentation time (RAMA) was plotted directly to indicate differences between similar or dissimilar target configurations.

All measures were originally taken in arbitrary units as used in the programs. For LUM experiments, these values transfer non-linearly into pixel luminance settings (Fig. 2). This transformation was also done for the analysis of the LUM popout data (Exp. 1) where all targets were presented on the same level of (virtual) background elements at target locations. However, the transformation was less obvious for the LUM grouping data, where the luminance contrast of bright and dark targets differed. For better readability, these data are therefore given in luminance steps, as plotted in Figure 2. For the DIS experiments, values give the disparity shift in the number of pixels on the screen, which, for the small angles used, transfers linearly into multiples of 1.8' disparity. This unit was then used in the presentation.

RESULTS

I. TARGET DETECTION EXPERIMENTS

On a background of identical or slightly modulated items, a target was presented at an increased contrast. The observer's task was to detect the target and indicate on which side from the fixation point it was located. Two parameters were varied in these tests, (a) the variation of background features from one item to the next in horizontal and vertical direction (background variation) and (b) the contrast of the target relative to the background value at this location (identical to that of two neighboring targets in an oblique direction).

Luminance variations

Experiment 1:
Detection of a salient luminance target

The experiment was performed by five subjects (three female) on regular arrangements of 12 x 12 circular blobs. Patterns were briefly presented (t=100ms) and masked afterwards by a pattern with random luminance settings (t=500ms). All targets (and background items) were brighter than the screen background. Typical stimulus patterns are shown in Figure 1.

Target detection varied between chance performance, for patterns with too low target contrast, and perfect detection rates, for patterns with large target contrast. This is illustrated in the data of subject GK (Fig. 3). Target detection rates were measured for different sets of background conditions, i.e. with different strengths of feature variations between neighboring items (different symbols and curves). From each data set, the target contrast with 75% correct performance was calculated and taken as the threshold contrast for salient target detection with this background condition. While on a uniform background (background variation 0.0), small differences were sufficient to let the target be reliably detected, thresholds increased when background items differed more strongly.

The 75% threshold contrast of all conditions is summarized in Figure 4 (squares) together with the data of the other four subjects. As is obvious, thresholds are small and almost constant when backgrounds are uniform or vary only little between neighboring items, but increase over-proportionally when background variation is further increased, until the available luminance range was finally not sufficient to make the targets be reliably detected when luminance differences between neighboring items in the background were large. Note that despite similar performance characteristics across subjects there were also notable differences. Thus, it is not quite clear why subject SL was so much better than all other subjects, at medium background variations. In fact, her performance in the first session of the experiment looked very similar to their performances. This may suggest either a "trick" this subject had used (see General Discussion) or strong training and learning effects, which were however not seen with the other four subjects.

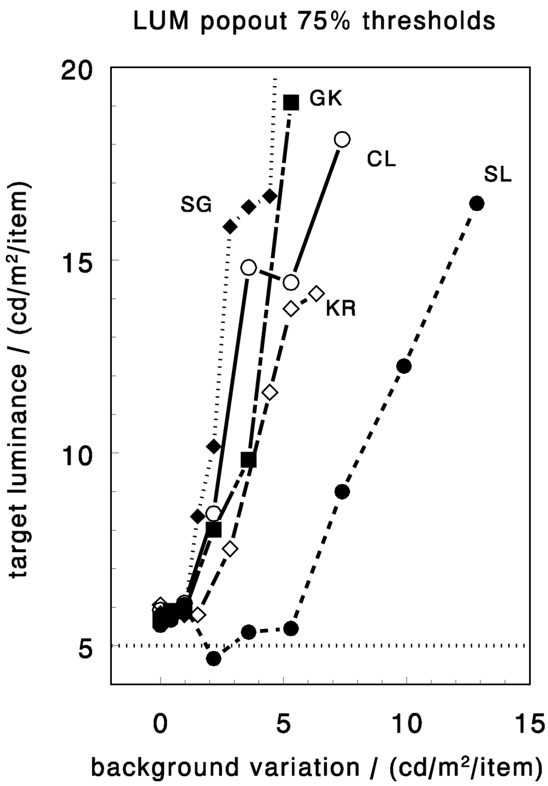

Figure 4. *Increasing 75% target luminance thresholds with increasing background luminance variations.* Data of 5 subjects (including subject GK from Fig. 3). For all subjects, the target contrast had to be strongly increased when the luminance gradient between items in the background was enlarged. Dotted horizontal line gives the luminance of (virtual) background elements at the target's location.

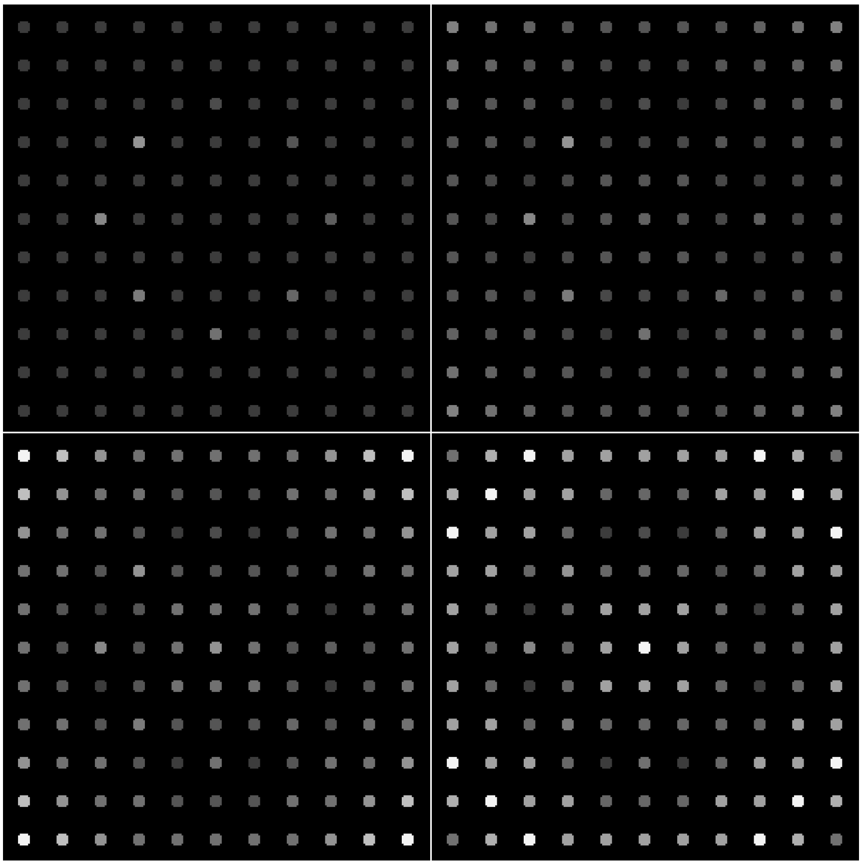

Figure 5. *Illustration of the main finding from Experiment 1.* The curves in Fig. 4 indicate that it becomes increasingly difficult to detect a luminance-defined target among items with an increased background luminance variation. This is illustrated here. Patterns show a circle of eight salient blobs with clockwise increasing luminance contrast. On a background of homogeneous items (top left; luminance gradient zero), almost all targets are seen. With an increasing luminance gradient of background items (top right, bottom left, bottom right) more and more targets loose their salience and can only be detected by careful screening of the according locations.

The findings of Experiment 1 can be visualized in Figure 5. Each pattern shows the same circular arrangement of 8 targets with clockwise increasing feature contrast to their immediate neighbors on the ring. While almost all targets can be detected when the background is constant (upper left), the visibility of the targets is quickly reduced when the luminance gradient across background items is continuously increased. With the largest background variation shown here (bottom right), almost none of the 8 targets is immediately detected.

Disparity variations

Experiment 2:
Detection of a salient disparity target

The corresponding experiment in the DIS domain was performed by five subjects (2 female); only two of them had also served as subjects in Experiment 1. Subjects wore glasses that were darkened in alternation to allow for different patterns through each eye. Stimuli displayed the same blobs as in Experiment 1 (now at 20.4 cd/m^2; but glasses had reduced all luminance settings to 10%) in an 8 x 8 slightly jittered arrangement; blobs could now occur at various disparities. All targets were "near" (convergent disparity), background items varied between "near" and "far" (divergent disparity). Patterns were shown for 160 ms and masked afterwards with an 8 x 8 pattern showing blobs at random disparities.

Performance (Fig. 6) was, in principle, similar to that in Experiment 1. On a homogeneous background (no disparity variations) subjects detected 75% of the targets already from small disparity shifts. But when background items themselves varied in disparity, much larger disparity shifts were needed to make the targets still be detected. Again, there was considerable variation between subjects.

The relatively fast increase of threshold contrast with increasing background variation of subjects GK and MI is perhaps not too surprising. Subject GK had already revealed a slightly reduced sensitivity in stereo vision in the initial tests (see Methods) but could reliably detect

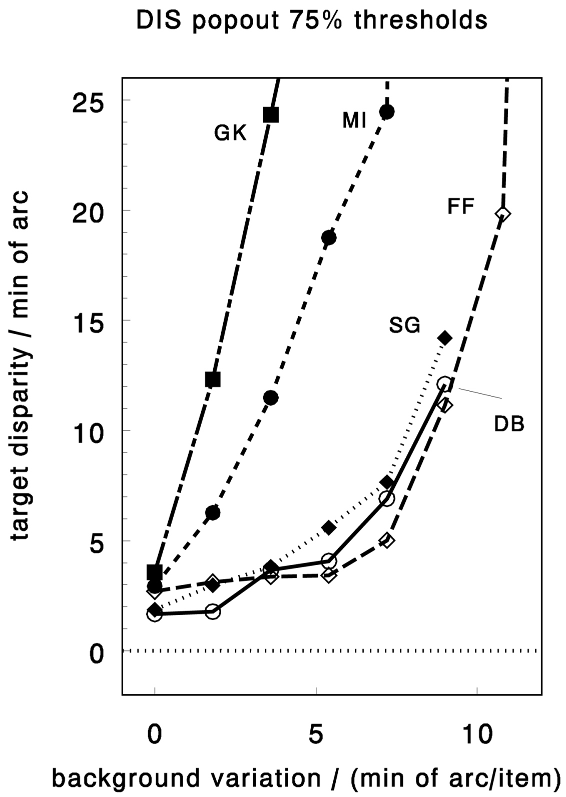

Figure 6. *Increasing 75% target disparity thresholds with increasing background disparity variations.* Data of 5 subjects from Experiment 2 (DIS popout). For fast and easy detection of the targets, their disparity contrast had to be strongly increased when the disparity gradient between surrounding items was enlarged. The general response pattern is similar to that for luminance variations in Fig. 5. Dotted horizontal line indicates the disparity of (virtual) background elements at target locations, relative to the monitor.

targets at 40" disparity which is well above the disparities tested here (> 1.8'). Subject MI could detect targets at 20" disparity in the initial tests but needed more time to perform this test than the other subjects (see below).

Note that the patterns in this experiment were comparable to the LUM patterns illustrated in Figure 1 except that all luminance variations should be imagined as variations in disparity. To illustrate the findings of Experiment 2, a demo similar to Figure 5 is given in Figure 7. The circle of salient blobs is far more compelling in the top than in the bottom stereogram.

II. ANALYSIS OF GROUPING PROCESSES

Is the grouping of salient items indeed based on the detection of feature similarities, as Gestalt psychology had suggested (e.g., Wertheimer, 1923)? To study this question I asked subjects to identify the configuration of three salient items which formed a global triangle (Fig. 8). In some conditions, these items were identical; in other conditions, two were identical and the third one was different. In the previously published experiments of the project (Nothdurft, 1993b) subjects did not make a difference between figures from similar or dissimilar salient items when salience was obtained from orientation or motion differences (and presentation time was short) but grouped all items just by their (similar) salience. In a third group of test conditions there were four salient items (Fig. 8c), three were identical and the fourth one was different. In the earlier study subjects failed to detect the triangle of similar items and performed at chance with these patterns when salience was generated from orientation or motion contrast. In the following experiments, the same paradigm was tested on luminance and disparity contrast, respectively.

Luminance variations

Experiment 3: Grouping of salient luminance targets

The role of luminance features in grouping was studied with configurations of bright and dark targets (Fig. 8). If target similarity is important for grouping, one should expect that subjects detect the triangles from three identical bright items (Fig. 8a) better than the triangles from bright and dark items (Fig. 8b) and also correctly perform the task with four items, three of which are identical (all bright; Fig. 8c). However, if target similarity and features are not important for grouping but only the increased salience of targets, one should expect subjects to perform the task similarly good for triangles from same or different targets (Fig. 8a and b) and fail completely in the four-target patterns (Fig. 8c).

To give all targets the same salience, they were presented with the same luminance difference (luminance steps) to the screen background and to virtual background items at these positions. Upon a small but efficient

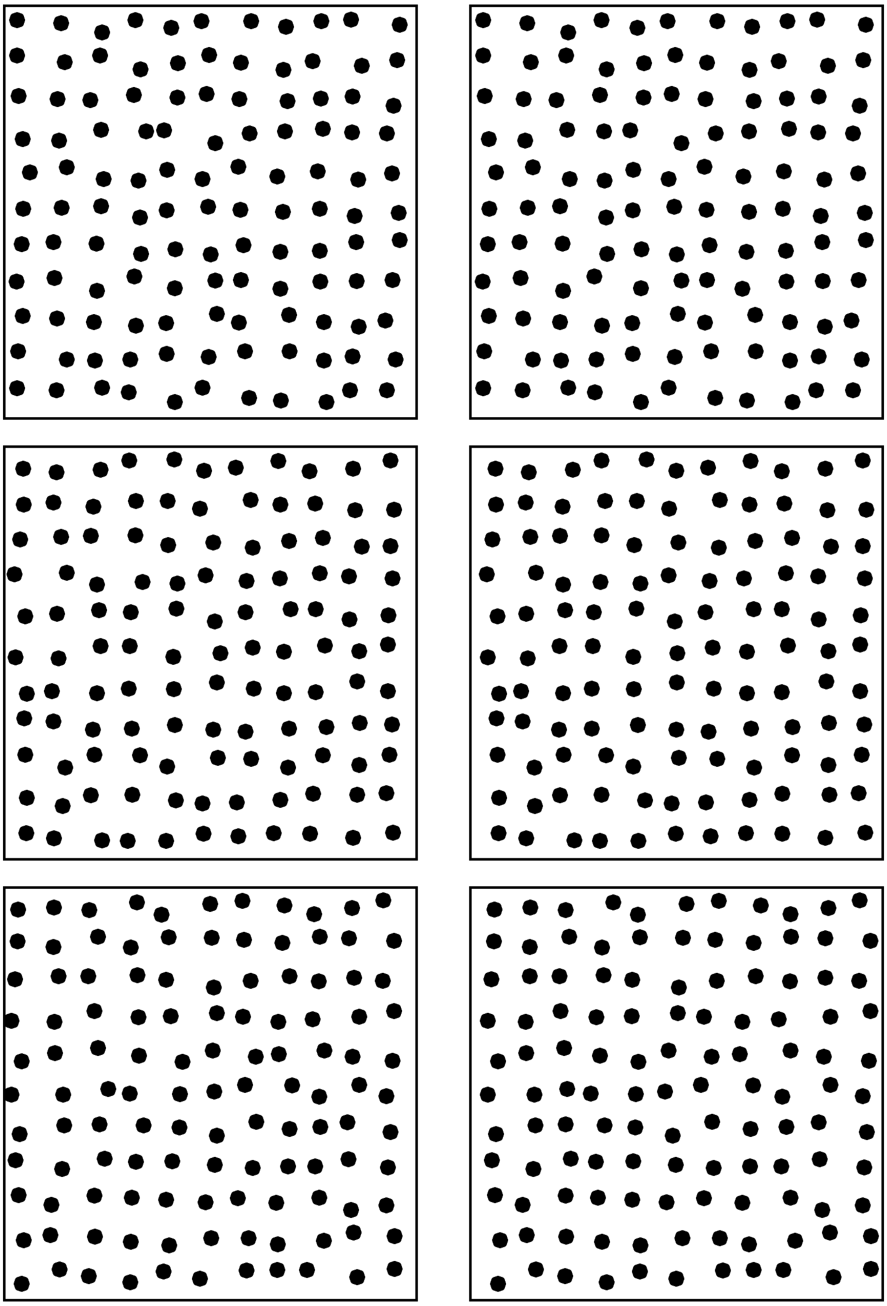

Figure 7. *Illustration of an increased difficulty to detect targets from disparity contrast when background variation is increased (Experiment 2).* The figure mimics the demo of Fig. 5, in depth. The stereo patterns are designed for free fusion (left pattern – left eye; right pattern – right eye). It may take time to see all target and background variations. The top pattern pair shows a circle of salient blobs with clockwise increasing disparity contrast to their neighbors on homogeneous ground. When the disparity gradient between background items is increased (middle and bottom stereograms) some blobs on the circle loose their salience and are more difficult to be detected. See Fig. 21 for an analogous illustration with random dot stereograms.

Published online: 15-Dec-2015 © *christoph.nothdurft@vpl-goettingen.de* ISSN:2364-3641

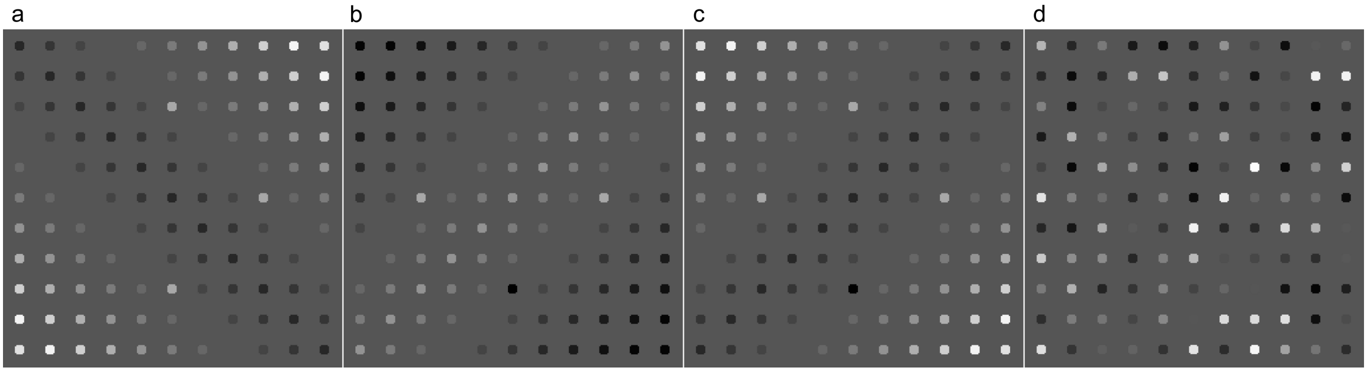

Figure 8. *Introduction to the grouping experiments (here LUM grouping).* **a.-d.** Example stimuli showing triangles of similar *(a)* and dissimilar targets *(b)*, and quadruples of three identical targets and a different one *(c)*. Patterns were shown for 100-200 ms and then masked *(d)*. In every condition, triangles were presented in all four orientations. Target luminance contrast was systematically varied. Purpose of the experiments was to find out if target similarity, and hence the identification of target features, is essential for perceptual grouping. If this were the case, subjects should perform better in *(a)* than in *(b)* and should also be able to detect the triangles composed of similar items in *(c)*. However, if target identities were not important and target features ignored, subjects should perform *(a)* and *(b)* similarly well and should be strongly confused in *(c)*. Note that the "similar" targets were always bright. For "dissimilar" target conditions one of the bright targets was replaced by a dark one; for the "square" condition, a dark target was added. Also note that short presentation times and a not too strong target contrast are essential in the experiment, as we all can tell the differences between *(a)*, *(b)*, and *(c)* under prolonged inspection. Luminance contrast of black targets is slightly increased here, for illustration.

luminance variation of background items (steps of 20), luminance contrast of the three or four targets was systematically increased, while the subjects' performance in detecting the global triangles was measured. The examples in Figure 8 represent a rather strong target luminance contrast to illustrate the various test conditions. In experiment, target contrast was usually smaller and patterns (11 x 11 blobs) were only briefly shown and masked afterwards. Five students (four female) served as subjects in this task. They were not asked to look for target identities but simply to identify the form of the global triangle, and they were not informed about the confusing fourth item in some configurations. Only two subjects, however, expressed their suspicion that there might have been four instead of three salient targets in a few trials. In these cases, their suspicion was not discussed but they were encouraged to continue to indicate the clearest and most obvious impression of the triangle form in every trial. (Fortunately, two thirds of the trials showed only three targets and in the majority of the remaining trials target contrast was small enough to prevent a clear and strong impression of four-target stimuli.)

The individual performances of three subjects are shown in Figure 9. The data are quite different from what has earlier been found with orientation or motion contrast (Nothdurft, 1992, 1993b). Instead, the data are little conclusive and partly support and partly reject the conclusion that feature identity is important. All subjects could better identify the triangles of similar targets than the triangles of dissimilar targets, and when patterns contained four targets, they still could quite often detect the triangles of similar items within. But on the other hand, the detection rates of figures from dissimilar targets were often rather large (and not very different from those of figures from similar targets), and the additional target in four-target patterns did notably (though not completely) disturb the detection of the same-target triangles in these samples.

The non-conclusive response characteristics are also seen when the data of all five subjects are averaged (Fig. 10a). But note that subjects were differently sensitive in the grouping task (cf. Fig. 9). For one subject (MI) presentation time had to be increased to 200ms to obtain an about similar range of performance variations; all other subjects were tested with 100ms.

How could the poorer performance with dissimilar targets and the reasonable but not perfect performance with square target configurations be explained? One immediate suspicion would be that different targets might not have displayed the same salience as the similar targets. If grouping would not care about features but only about the salience of targets (based on feature *differences*), then

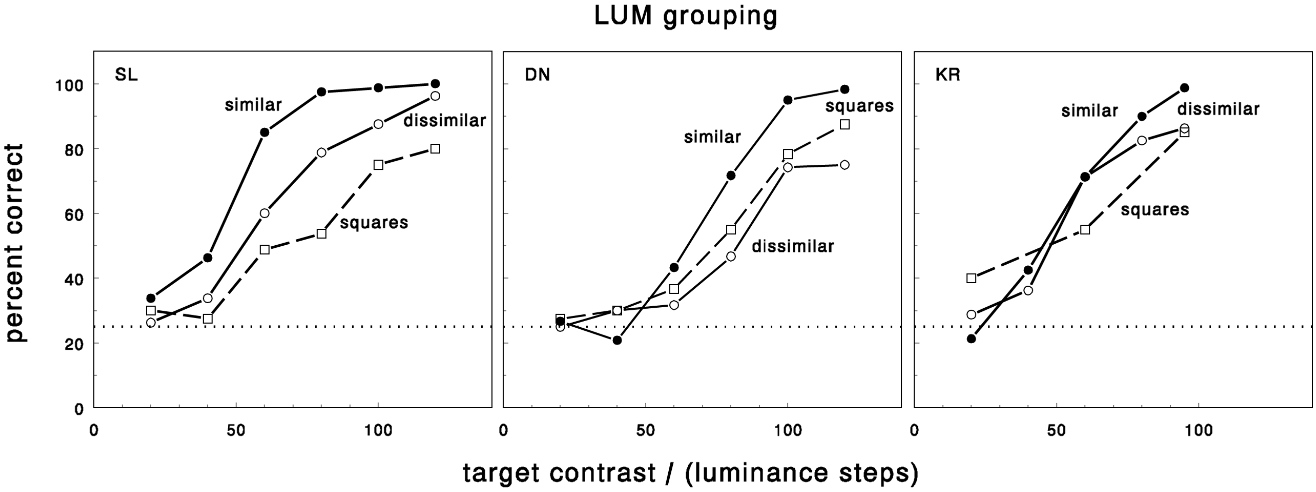

Figure 9. *Performance of three subjects in Experiment 3 (LUM grouping).* Results lie between the two predictions made in Fig. 8; triangles from similar targets were seen slightly better than triangles from different targets, but performance was only gradually deteriorated when a fourth different target was added ("squares"). Note that luminance scaling in this and the following figures refers to the computer luminance steps plotted in Fig. 2. Zero refers to the lookup-table value 130 (6.5 cd/m^2), to which target luminance contrast was added.

the combination of salient bright and less salient dark targets should produce exactly the response patterns seen. Performance should be better with similar targets (all bright and equally salient) than with dissimilar targets (with one dark and less salient target). And performance with four-target patterns should be gradually but not completely disturbed by the fourth (slightly less salient) target. Indeed there are good arguments to assume that bright and dark targets in Experiment 3 were not equally salient.

Although the monitor background was adjusted to the mid level of the used luminance scale (Fig. 2) und differences of bright and dark targets to their neighbors displayed the same luminance steps on that scale, the nonlinear transform to pixel luminance settings created differences in some contrast measures between dark and bright targets. The Weber contrast of bright targets was larger than that of dark targets. Thus, if salience is related to the Weber contrast, the bright targets should have been more salient than the dark targets.

There are different reports on the role of Weber contrast in salience from luminance (Dannemiller & Stephens, 2001; Nothdurft, 2015a, b). Dannemiller and Stephens found bright targets being generally less salient than dark targets with the same Weber contrast and concluded that the Michelson contrast might be a better measure of the salience of bright and dark targets. In that case, luminance differences of bright targets to background should be larger than luminance differences of equal-salient dark targets to the background, as in the present experiments. The authors did, however, report that equal salience followed the Weber contrast when luminance differences were not too large (p. 121). Also Nothdurft (2015a) found recently that equal-salience matches of bright and dark targets are obtained for the same Weber contrast. In that case, the dark targets in the present study were clearly less salient than the bright targets. When targets are shown together with other items and presented at different background levels (which was however not the case in the present experiments), salience computation may become even more complicated (Nothdurft, 2015b).

Experiment 4: Detection rates of bright and dark targets

To measure potential salience differences between the dark and bright targets in Experiment 3, a control experiment was performed in which single targets, instead of triangles were shown. Subjects were asked to indicate in which direction from the central fixation point the salient blob occurred, irrespectively of whether it was bright or dark. The principal design of Experiment 4 was similar to that of the triangle test before; targets occurred

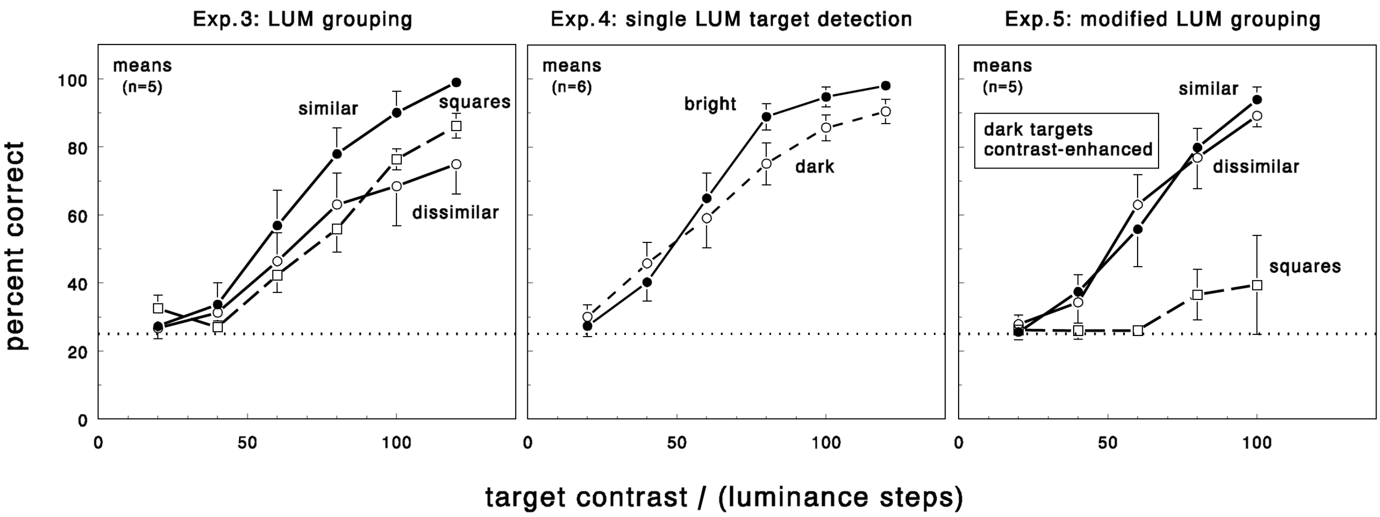

Figure 10. *Mean data of all subjects in Experiments 3-5.* **a.** Data of all five subjects. Curves confirm the uncertain role of target identity in Exp. 3; target features (bright or dark) were neither essential for grouping (then the performance on “squares” should have been better) nor fully ignored (then the performance on “squares” should have been at chance). One conjecture was that bright and dark targets might have differed in salience. **b.** Experiment 4 measured the detection rates of single bright and dark targets. Indeed, with increasing target contrast, bright targets were better detected than dark targets, which might have affected performance in Exp. 3 *(a)*. **c.** Experiment 5 was a repetition of the grouping experiment with dark targets presented at increased contrast (factor 1.3). Triangles of same or different targets were now detected at similar rates, and the detection of groups of similar targets was now strongly confused by the added dissimilar target. Vertical bars in this and the following figures plot the s.e.m. (if larger than symbols).

at the same locations, the same background patterns and luminance settings were used (including the systematic variation of target contrast), and subjects had the same four keys for responses. Six subjects (four female) performed the task, including all subjects who had performed Exp. 3. Presentation times were 100ms for the subjects already tested (200ms, for subject MI), and 160ms for the additional subject (RZ).

Five of the six subjects showed notable performance differences between the bright and dark blobs, which are also seen in the mean responses (Fig. 10b). With increasing target contrast, subjects detected more frequently the bright than the dark target. Only one subject revealed about similar detection rates with the two targets. As explained above, this salience difference between targets could, in principle, explain the performance variations seen in Experiment 3 (Fig. 10a). It would be interesting to see if the variations disappear when the experiment is repeated with targets that are better adjusted for salience.

Experiment 5:
Grouping of luminance targets with better adjusted salience

The experiment was identical to Experiment 3 except that *dark targets* were shown at an increased luminance contrast (factor 1.3 in the luminance steps) compared to bright targets. That is, a bright target with a luminance step of, e.g., 60 was combined with a dark target with a luminance step of 78 to background. This modification was only applied to the targets; background variations were identical to those in Experiment 3. Pattern presentation (100-200ms, for different subjects; masking) and task were the same as before. Five subjects (three female) performed the experiment; they all had also been tested in Experiment 4 but only four of them in Experiment 3.

The effect of presenting dark targets at an increased contrast is notable and can directly be seen when comparing Figure 10a and 10c. Triangles of similar and dissimilar targets were now almost equally detected, and performance with the square target configurations was now more often at chance (25%). Thus subjects did apparently not use target features, like their similar lightness, for grouping but looked at the target salience generated from feature *differences*. When targets differed in salience, as was obviously the case in Experiment 3 (cf. Fig. 10b), subjects could not detect triangles from same and different targets equally well and could, on the other hand, still detect quite a few salient triangles from the four-target configurations (Fig. 10a). But when bright and dark targets were better adjusted in salience, as in Experiment 5, they apparently looked more similar and could not be distinguished in grouping (Fig. 10c).

Note however that there still are deviations. At luminance steps 80 and 100, same and different target ratings began to be distinguished and (bright) same-target triangles in the four-target configurations were sometimes recognized. But this is not surprising. With a background luminance of 6.5 cd/m^2 as in the experiments, the Weber contrast of dark targets was limited to 4.5, a value that was reached with bright targets at a luminance step of 60 (cf. Fig. 2). Thus, for larger luminance contrast steps, the salience of bright and dark targets could not be adjusted in the experiments.

Disparity variations

Experiments 6–8 represent similar tests in the disparity domain; stimulus examples are illustrated in Figure 11. Experiment 9 is an additional experiment explained below. All experiments were performed by the same three subjects (one female). The standard disparity gradient in background items was 1 step per item corresponding to 1 pixel shift (1.8') between the two patterns. In Experiment 9, also 2 pixel shifts were used. Background disparity shifts varied between near and far and were constructed to display zero disparity (relative to the monitor) at target locations. Targets were presented at variable disparity contrast with triangles made of similar (all near) or dissimilar targets (near and far) or four-target patterns (three near, one far). The subjects' task was to identify the form of salient target configurations and indicate the orientation of the triangles, as in the LUM grouping experiments before. Subjects were not told about the occasional occurrence of four targets in some patterns.

It turned out in the early experiments that subjects had considerable problems to see and identify the global figures of salient targets in short presentations, when targets were too widely spaced. Therefore, the triangle configurations were modified and targets were displayed more closely to each other (as illustrated in Fig. 11). This new configuration was then used in all DIS grouping tests, and standard presentation time was increased to 200ms.

Even this duration was apparently far too short for one subject (MI) who could not perform the standard DIS grouping test under these conditions (no data in Exp. 6) but required much longer presentation times to reach similar performance levels as the other two subjects. This is documented with a *single target detection* task (from Exp. 7) in Figure 12. Detection rates of this subject differed considerably between 200ms presentations and 1s presentations. To make performance comparable between the three subjects, the longer presentation time of 1s was generally used for disparity tests with this subject, from Experiment 7 on.

Experiment 6:
Grouping of salient disparity targets

The experiment on disparity grouping (Fig. 13a) revealed qualitatively similar response variations as did Experiment 3 in the luminance domain. Triangles from similar items (all near) were better identified than triangles from dissimilar items (two near, one far), and triangles from similar items in square configurations (three targets near, one far) were often recognized. However, the same conjecture that had been made with bright and dark targets in Experiment 3 should also be applied here. Were the near and far targets indeed equally salient? Or might subjects have simply distinguished the targets for their different salience?

Experiment 7:
Detection rates of near and far targets

The control experiment was designed in analogy to Experiment 4. Subjects were asked to detect a single salient target that occurred under either convergent (near) or divergent disparity (far).

Over a large range of disparities, all subjects detected the near target much better than the far target (Fig. 13b). They thus may indeed have grouped the targets in Experiment 6 not for feature similarity (targets occurring near) but simply for their different salience. To overcome

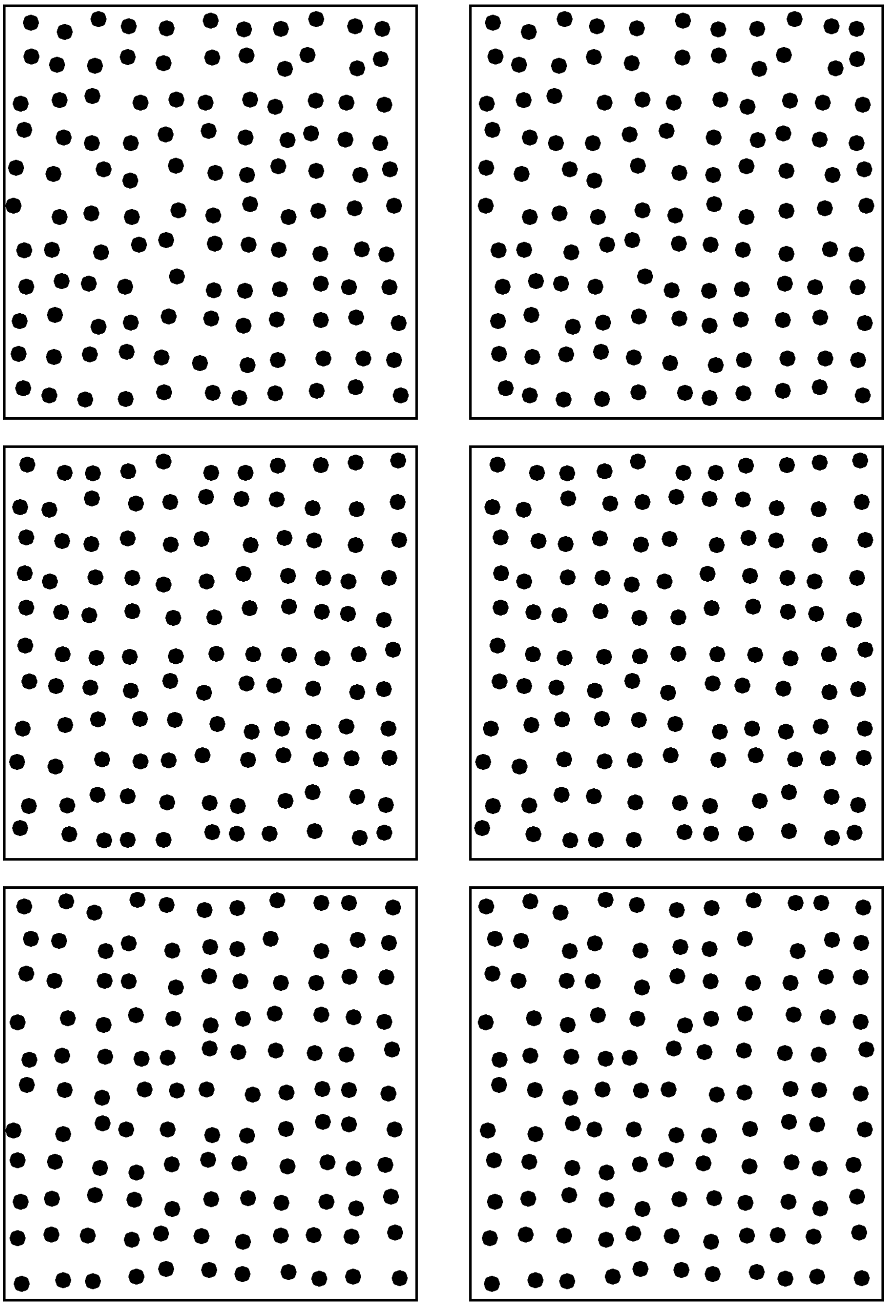

Figure 11. *Introduction to Experiment 6 (DIS grouping)*. Stereograms illustrate the three test conditions with binocular disparity; the mask is not shown. Examples show a triangle of similar targets (all near; *top)*, a triangle of dissimilar targets (two near, one far; *middle)*, and the square configuration of three near targets with a single far target *(bottom)*. Patterns were constructed in analogy to those used in the LUM grouping experiment (Fig. 8), with the exception that the disparity targets were located one row/column closer to the midpoint than the luminance targets in Exp. 3-5. Similar predictions for grouping performances can be made.

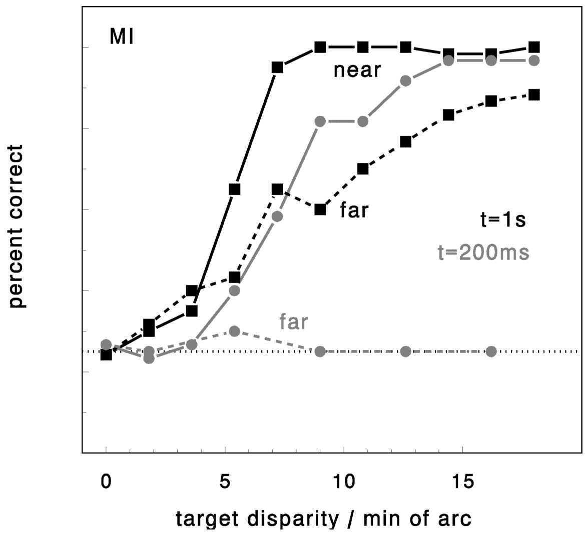

Figure 12. *Poor performance with too short presentation times.* The figure illustrates a problem that occurred with DIS grouping. While other subjects could perform the DIS grouping task with 200ms presentation time, this subject failed completely under this condition. The data displayed here are from a single target detection task. The target appeared at different disparities (abscissa) either near or far, and the subject had simply to detect and localize it. With 200ms presentation time, the subject did not detect any far target. When presentation time was increased to 1s, performance improved in the detection of both near and far targets, but a difference still remained.

this problem, I therefore have applied a similar modification in Experiment 8 as I did in Experiment 5 before; disparity contrast of the far target was increased to enhance its salience.

Experiment 8: Grouping of disparity targets with partly adjusted salience

Like in the luminance domain, grouping performance changed notably when different targets were better adjusted in salience. Figure 13c shows the data with patterns in which far targets had twice the disparity contrast of the accompanying near targets. Triangles from different targets (near and far) were then detected better and the addition of a far target to the triangle of near targets produced the expected confusion (reduced performance with "squares"). For larger disparities, however, the improvement was suboptimal. Triangles from same and different targets were still seen at different rates (though less different than in Fig. 13a) and same-target triangles in four-target configurations were still sometimes detected (less often than in Fig. 13a). But this is to be expected from Figure 13b, which shows that detection rates of near and far targets seem to settle at different levels, so that salience differences cannot be compensated

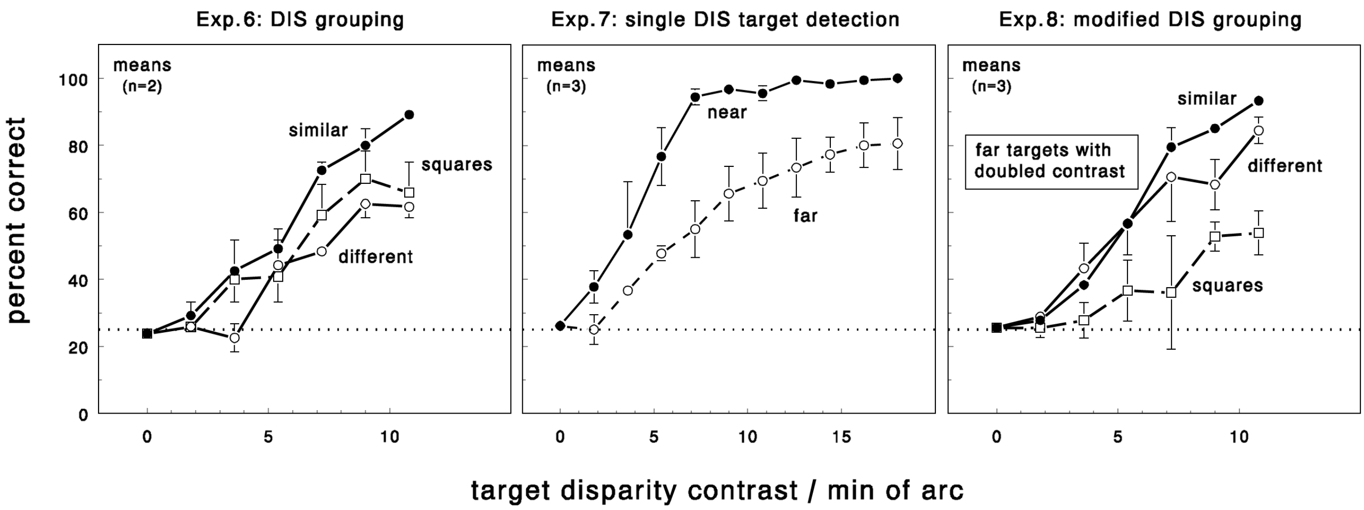

Figure 13. *Mean data from Experiments 6-8.* **a.** Data from the DIS grouping experiment (Exp. 6) reveal similar characteristics as the data from the LUM grouping Exp. 3 (Fig. 10a). Neither are targets at the same depth reliably distinguished from targets at another depth nor is the depth information fully ignored. Therefore, a similar conjecture was raised that near and far targets might have differed in salience. **b.** This was confirmed in Experiment 7, which measured the detection rates of single near and far targets at same disparity contrast. Near targets were detected much better than far targets. **c.** In analogy to Exp. 5 (modified LUM grouping) the disparity contrast of far targets was doubled in Experiment 8. Up to a disparity contrast of about 7' (14' for dark targets), detection rates of triangles from "similar" and "different" targets do now overlap and the detection of the similar-target triangle in "squares" is disturbed.

for any disparity contrast above 7 min of arc, consistent with the increasing deviations of curves in Fig. 13c from this level on. Up to this level, however, similarity grouping effects were reduced when target saliences were adjusted. In summary, thus, the evidence for feature-based grouping in disparity is week. Grouping seems to be primarily driven by target salience, and any apparent preference to group near targets better than mixtures of near and far targets can be changed by adjusting the relative salience of near and far targets.

Experiment 9: Grouping of same and different near targets

Given the different sensitivity to near and far targets, a different stimulus design was tested in Experiment 9. Same and different targets were now all presented at convergent ("near") disparities, with the same relative disparity contrast to adjacent neighbors (to guarantee that they are equal salient), but they still differed in absolute disparity (relative to the screen). This was achieved by presenting targets on different disparity levels of virtual background items, as illustrated in Figures 14 and 15. Same targets did then appear near, and different targets even nearer or less near than these. These variations resulted in a total of five test conditions for each tested target contrast; (i) the "same"-targets condition (all targets near) as in the previous experiments; (ii) and (iii) two "different"-targets conditions, in which one target occurred nearer or less near than the others, and (iv) and (v) two four-target conditions, in which the confusing fourth target was nearer or less near than the three targets that formed the triangle. The latter two cases are illustrated in Figures 14 and 15. Again, if target similarity was important for grouping, subjects should better detect triangles of similar than dissimilar targets, and should still detect the same-target configurations in four-target patterns. If grouping is not based on similarity, however, but only on salience from disparity *contrast*, performance should not differ between the various triangle configurations and stay at chance in all four-target conditions.

The stimulus design is somewhat tricky and therefore also visualized in luminance (Fig. 14). Note however that in the tests all blobs had same luminance and differed only

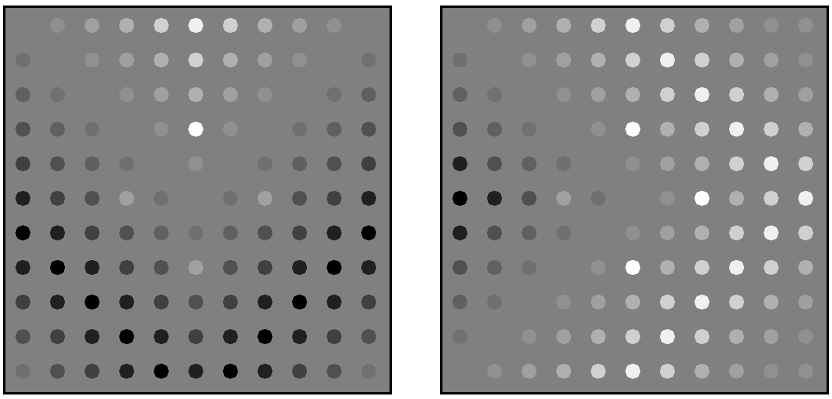

Figure 14. *Grouping within depth categories (Experiment 9).* The figure illustrates two square configurations with triangles of same targets and a fourth different target. Note that the experiment was performed on binocular disparity; this demo is only meant for illustration (a stereogram with binocular test conditions is shown in Fig. 15). All targets display identical feature contrast to their neighbors; other than in the previous grouping experiments, however, the "different" target was not dark (i.e. far) but also bright (near) like the three "same" targets. The subjects' task was identical: Find the triangle and indicate its orientation. This modification had two major effects. The similar disparity contrast of all targets in the same direction (all were brighter, that is nearer than their neighbors) made them all equally salient (not in the luminance demo here). In addition, grouping was now restricted to the same category (bright, i.e. near). Two different combinations of square target configurations were distinguished; near and nearer targets (bright and brighter; left-hand figure) and near and less near targets (bright and less bright; right-hand figure). The same distinction was made for triangles from different targets (not shown). Together with the one same-target condition, this gave a total of five test conditions in Experiment 9, which all were tested with varying target contrast. (Some targets are shown at enhanced contrast to illustrate the paradigm.)

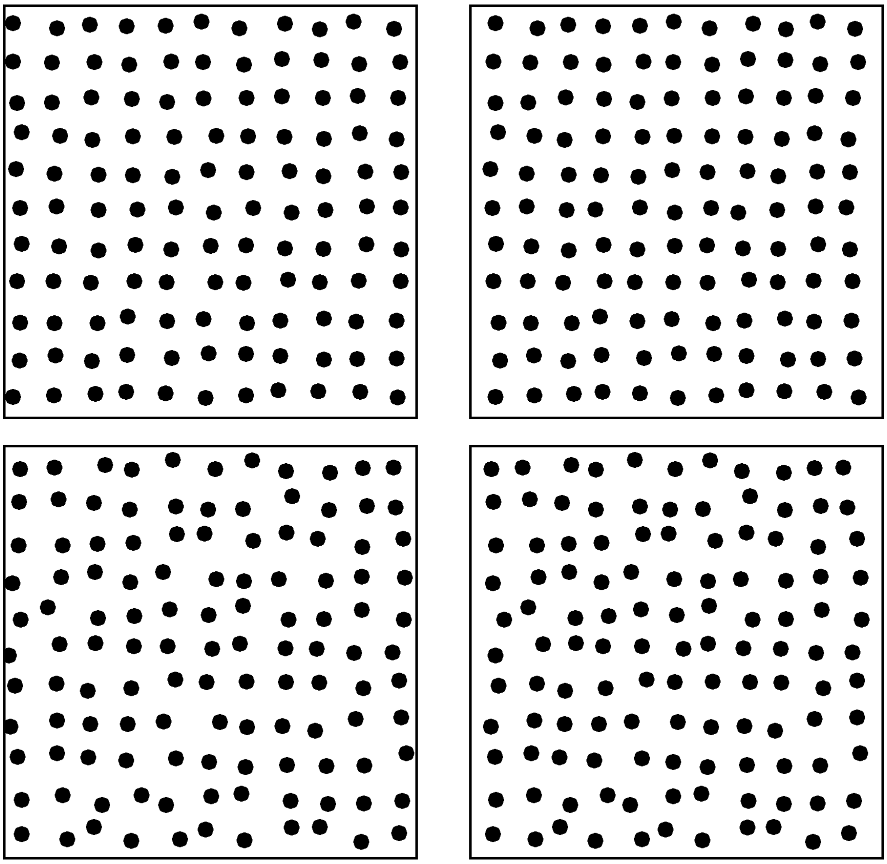

Figure 15. *Stereograms with two examples of square target configurations (as illustrated in Fig. 14).* The upper stereogram shows a triangle of "same" targets (same depth) pointing to the left plus a single (nearer) target on the right; all targets have the same relative disparity contrast to their neighbors and all appear "near"; the right-hand target is nearer in absolute terms. The lower stereogram shows a triangle of same targets pointing to the right plus a single (less near) target on the left; all targets have the same relative disparity contrast to their neighbors and all appear near; in absolute disparity, the left-hand target is less near than the other three targets.

in disparity. For readers with good stereo vision, the two "squares" conditions are shown in Figure 15. All three subjects who had served in Experiments 6-8 did also perform Experiment 9. Presentation time was 200ms (also for subject MI who was then considerably trained); the patterns were masked afterwards (not for subject MI, however). Background disparity gradients were 1.7 min of arc / item and increased to 3.4 min of arc / item for tests with larger target contrast.

The results were quite impressive (Fig. 16). Triangles from different targets were about as well detected as triangles from same targets, and the confusion from a fourth dissimilar target ("squares") was quite strong in all tests (Fig. 16a). There were small deviations between the two variants with different targets (combinations with nearer targets were slightly better than combinations with less near targets), which compensate however, so that the averages of the two variants represent exactly the data that should be expected for pure salience and absent similarity grouping (Fig. 16b).

The differences between Figures 13 and 16 are compelling. They indicate that targets are not grouped for similar disparity but rather for salience (as generated from disparity contrast); absolute feature properties (disparities) of the targets are then ignored. Thus, grouping processes in vision are similarly unaware of absolute features in disparity as they are unaware of absolute features in orientation or direction of motion (Nothdurft, 1993b).

There are two different explanations, however. If subjects could not even *see* the differences between near and nearer, or near and less near targets but had simply been blind to these variations, they should have produced exactly the same results. (In fact, the experiment was explicitly designed to reveal such a blindness for feature properties.) From the construction of stimulus patterns, however, the absolute disparity difference between same and different targets was four times the background variation, hence 7.2 and 14.4 min of arc, respectively, in different tests. These values were well above the visibility thresholds of the tested subjects for *relative* disparity (cf. the detection rates for "same" targets in Figs. 13 and 16). What the data in Fig. 16 do show, however, is a functional blindness to absolute disparity differences between targets in grouping, which thus cannot be due to a general insensitivity to disparity differences of this magnitude.

The other explanation is a merely hypothetical one. The observation that grouping was not entirely insensitive to target similarity in Experiment 8 (Fig. 13c) but was so in Experiment 9 (Fig. 16b) is not only explained by the better controlled equal salience in Experiment 9. It might also be that "near" and "far" represent different categories in disparity which can perhaps not be ignored in grouping. Only if grouping were restricted to either category alone (like in Exp. 9 to "near"), disparity features might then have been ignored. We cannot refuse this hypothesis on the basis of the present data. Further experiments would be nessary to study the grouping of near and far targets that are exactly matched in salience. The fact, however, that grouping of such targets does tend to ignore features even across categories if and when targets *are* better matched in salience (Fig. 13c; data up to disparity contrast 7.2 min of arc) indicates that this hypothesis is likely not to be verified.

Experiment 10: Grouping of depth from shadow

The last two experiments of the study were added partly for anecdotal reasons. It is well-known that the impression of depth may also occur from many monocular cues like perspective, texture, or occlusion, which also can be

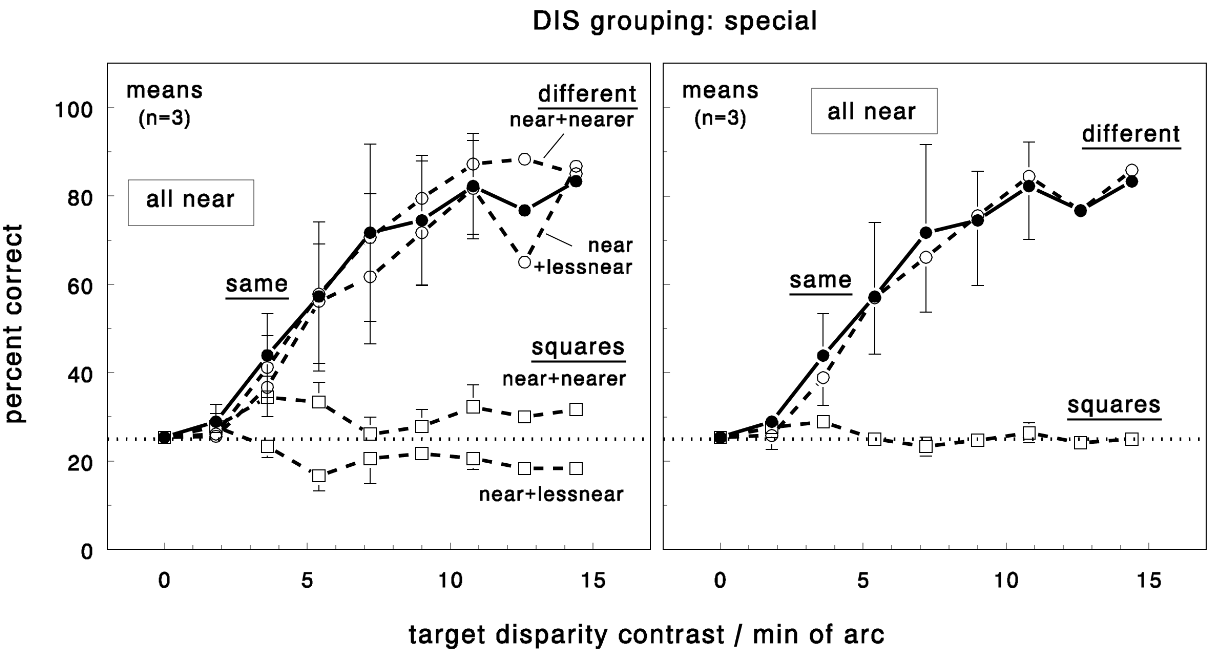

Figure 16. *Mean performance in Experiment 9.* Curves plot detection rates of target triangles against increasing target contrast. **a.** For "different" and "squares" target conditions two curves are distinguished in which the different or confusing target was either nearer or less near than the other targets. These two conditions are collapsed in **b.** Detection rates for triangles from same and different targets are almost identical, while same-target triangles in square configurations could only be detected at chance.

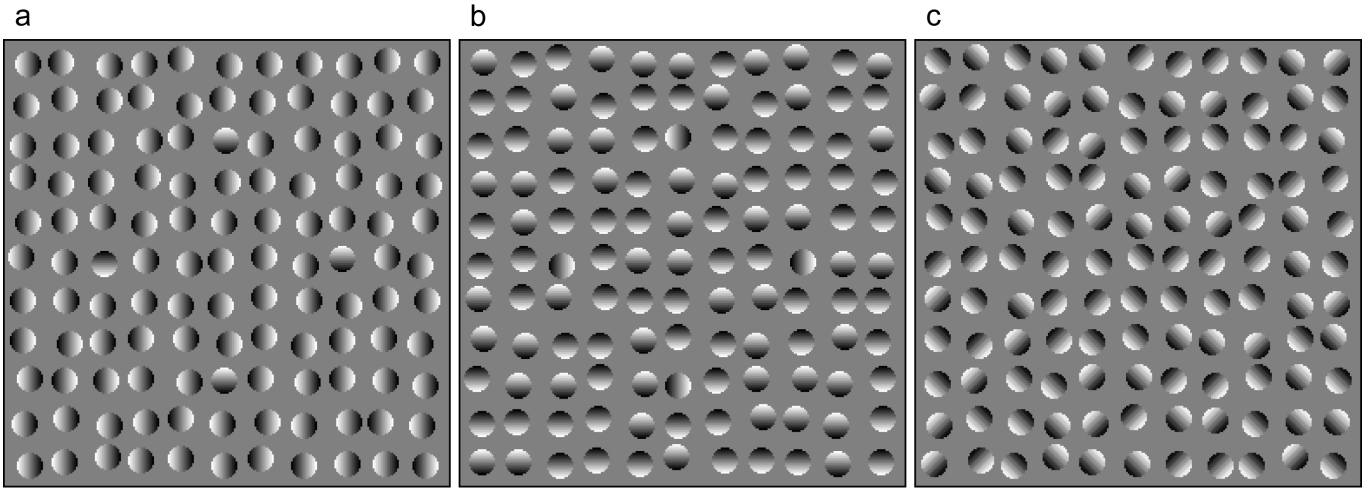

Figure 17. *Grouping of monocular depth cues based on shape-from-shadow (Experiment 10).* Only square target configurations were tested. **a., b.** Stimulus examples; **c.** mask. Patterns showed four salient blobs with different luminance gradients; three were identical and formed a global triangle, the fourth blob had the opposite luminance profile. Subjects had to detect and identify the global triangle, irrespective of the luminance gradients from which it was composed. Four luminance gradients of the blobs were distinguished, showing brighter upper or lower halves or brighter halves on the left or right-hand side. The remaining items in each pattern represented a random distribution of blobs at orthogonal gradient orientations. For an easy distinction, stimulus conditions were labeled according to the lighting direction under which three convex and one concave blobs would have produced the plotted luminance gradients. Examples show a triangle of identical bumps lighted from above (*a;* the triangle is pointing to the right) and a triangle of identical bumps illuminated from the right (*b*; the triangle is pointing downwards). In experiment, patterns were shown for variable presentation times and then masked with a pattern of random luminance gradients in oblique orientations *(c)*. Note that only lightings from above or below produce a strong impression of three-dimensional bumps and troughs *(a)*. The depth impression is absent with horizontal luminance gradients *(b)* but partly also seen with oblique gradients in the mask (c) if the upper halves of the blobs are brighter than the lower halves.

grouped perceptually or may affect grouping of other items (e.g., Aks & Enns, 1996; Enns & Rensink, 1990). One of these cues, depth from shadow, was implemented in Experiment 10 using circular balls with different luminance profiles (Fig. 17). With the implicit assumption that light is coming from above, these balls may give the vivid impression of bumps (if the upper part is brighter than the lower part) or troughs (if the lower part is brighter). The impression of apparent depth is much smaller or totally absent with luminance gradients in horizontal directions (resembling illumination from the left or right). Stimuli of this sort have been extensively studied (e.g., Braun, 1993; Kleffner & Ramachandran, 1992; Liu & Todd, 2004; Sun & Perona, 1998; Symons, Cuddy, & Humphrey, 2000).

I used these stimuli to study the influence of depth cues in grouping. If depth were an important cue for salience detection and grouping, so might be apparent depth. In that case we should expect balls with vertical luminance gradients (bumps or troughs) group easier than balls with horizontal luminance gradients (no depth impression).

The experiment was restricted to the four-target condition of the previous grouping experiments. Three of the four targets displayed the same feature, e.g. a blob with a vertical luminance gradient and a brighter upper half, which would be seen as peak when assumed to be illuminated from above (Fig. 17a). The fourth target displayed the "opposite" feature, in this case a blob with a vertical luminance gradient and a brighter *lower* half, which would be seen as trough when assumed to be illuminated from above. Thus, for this particular lighting condition, the four blobs should strongly differ in apparent depth. All other items in the pattern were blobs with randomly selected luminance gradients in the two orthogonal directions, i.e. horizontal luminance gradients and brighter halves on either the left or the right side. The observer's task was to identify the triangle of similar blobs, irrespective of which luminance gradient these blobs actually displayed. Note that, different to the previous experiments, subjects were now explicitly asked to look for target similarities and were aware that there were four salient targets on the display. Different trials in a

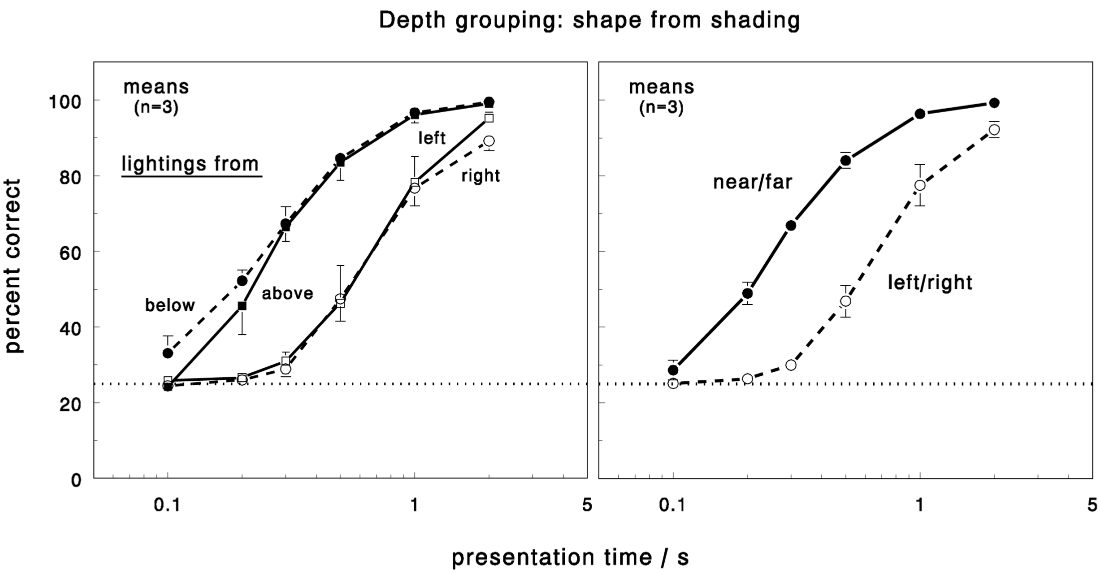

Figure 18. *Mean performance in depth-from-shadow grouping.* Triangles were much better detected and distinguished from the fourth disturbing target when luminance gradients in the blobs were associated with depth cues, i.e. with the percept of bumps or troughs under an assumed lighting from above or below. The grouping of targets which are illuminated from the left or right required more time. Given the similar performance in corresponding lighting conditions (from above or below vs. left or right), the two conditions are collapsed in the right-hand graph.

run showed patterns with different triangle configurations, over which performance (detection rates) was averaged, and triangles formed from different blob orientations (e.g., Fig. 17b), which were distinguished in analysis. The different blob orientations were labeled by the direction of light under which the triangles would appear as bumps (and the fourth item as trough), which you may confirm by rotating Figure 17 so that the according lighting direction were from above. Different to previous experiments, targets were not varied in salience. To measure performance variations, patterns were shown at various presentation times, between 100ms and 2s, and subsequently masked by a pattern with items at randomly selected oblique orientations (e.g., Fig. 17c). Within a given run, presentation time was held constant, but varied between runs. Experiment 10 was performed by three subjects (one female).

As you may visualize in Figure 17a and b, the triangles of similar items are well seen and their global form can easily be identified. However, some items tend to group faster and give a perhaps stronger impression of the global triangle than others. This is also found in the data (Fig. 18a). Performance was generally better (more correct identifications at shorter presentation times) for the two cases with assumed lighting from above or below (triangles made of bumps or troughs) than for the cases with lighting from the left or right, which do not give a strong percept of apparent depth (cf. Fig. 17a and b). For one subject, performance was nearly identical when bumps or troughs had to be grouped; the other two subjects showed small preferences for one or the other case. In the means, performances with lightings from above or from below fall close together, as also do, at a lower level, performances with lightings from the left or right. These curves are averaged in Figure 18b (continuous and dashed lines, respectively).

The description of blob orientations with the labels "lighting from top, bottom, left or right" should not confuse here. This nomenclature was only used to distinguish the four luminance gradients. In reality, however, we have no voluntary choice where to imagine the light is coming from and when to see bumps and troughs or not. The percept of apparent depth is closely linked to lightings from certain directions (Sun & Perona, 1998; Symons, Cuddy, & Humphrey, 2000). Only under these conditions are the triangle items seen as peaks (and the fourth item as trough). The two orthogonal luminance gradients do not even evoke the impression of an 3D object if you imagine new light sources at horizontal locations. But it would immediately when you rotate the figure.

Experiment 11: Grouping without depth from shadow

To test if the performance differences between vertical and horizontal luminance gradients were indeed related to the perceived depth, the stimuli of Experiment 10 were modified in a simple but very effective manner. The luminance *gradient* of each blob was replaced by a simple luminance *step* between two levels, bright and dark (Fig. 19). In this condition, there is no strong percept of peaks and troughs. All other aspects of the stimuli, like the orientation of target blobs and the arrangements of targets, were identical to those in Experiment 10. The same three subjects as in Experiment 10 also performed Experiment 11.

The intuitive assumption was that differences in grouping should now have disappeared since subjects could not further use apparent depth. But this was not the case. In fact, all subjects still revealed notable differences in the speed at which they could group the different blobs (Fig. 20). Subject RZ, for example, showed the same preference for vertical luminance steps (horizontal borders), independent of whether these were bright to dark or dark to bright, over horizontal luminance steps (vertical borders), as he did before in Experiment 10. His grouping performance with luminance steps was, in fact, even slightly faster than that with luminance gradients. Performances of the other two subjects were more variable. But all subjects produced much higher detection rates, at a given presentation time, with luminance steps that corresponded to the previous luminance gradients with lighting from above than with luminance steps that corresponded to the orthogonal luminance gradients with lightings from the left or right. However, the difference between these conditions became slightly smaller, since in the latter conditions (lightings from the left or right) performances with luminance steps had improved over those with luminance gradients.

Note that the successful grouping of targets in Experiments 10 and 11 itself is not surprising, given the long presentation time until grouping performance was perfect. All stimuli represent strong orientation cues (cf. Fig. 19) which should have helped subjects to group the triangle targets. Since mainly orientation differences and not differences in the direction of the luminance step would contribute to salience (Caelli, Hübner, & Rentschler, 1986; cf. Fig. 2 in Nothdurft, 2006), all four targets in the test patterns (but only these) were salient and should have been detected easily. Such targets can reliably

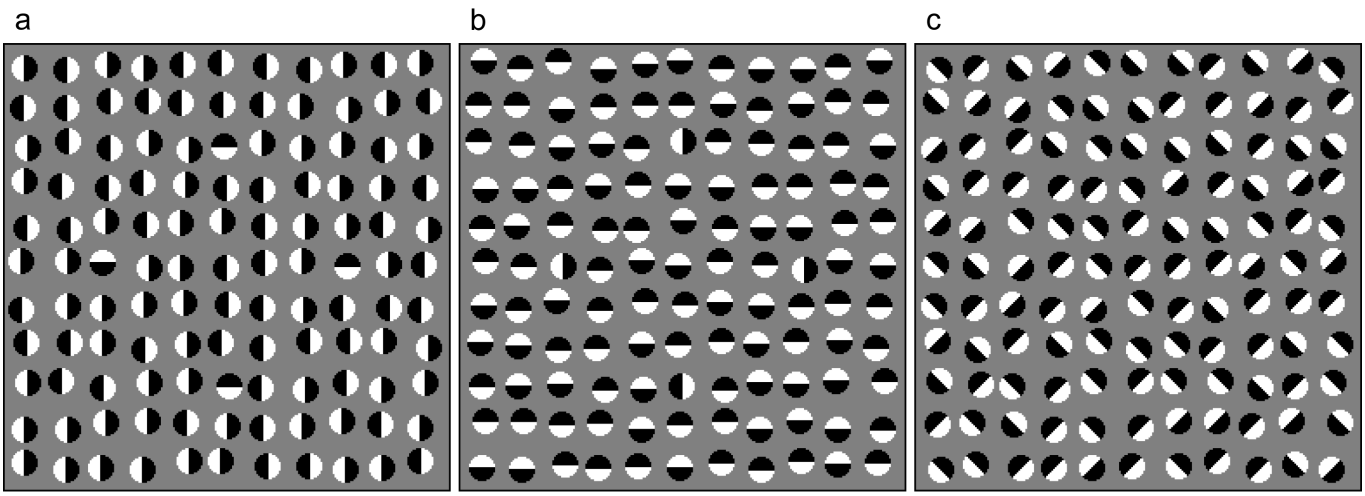

Figure 19. *Replication of Exp. 10 with depth cues removed (Experiment 11).* **a., b.** Stimulus examples; **c.** mask. Continuous luminance gradients in Exp. 10 are replaced by luminance steps which do not produce a similarly strong and immediate percept of depth. Examples show target configurations with vertical (*a)* and horizontal luminance steps *(b*; borders run orthogonal), and a masking pattern with luminance steps in oblique directions *(c).* Subjects were asked to detect the triangles of similar blobs and indicate their global orientation (i.e. pointing to the right or upwards, respectively, in the examples *(a)* and *(b)).*

  ISSN:2364-3641

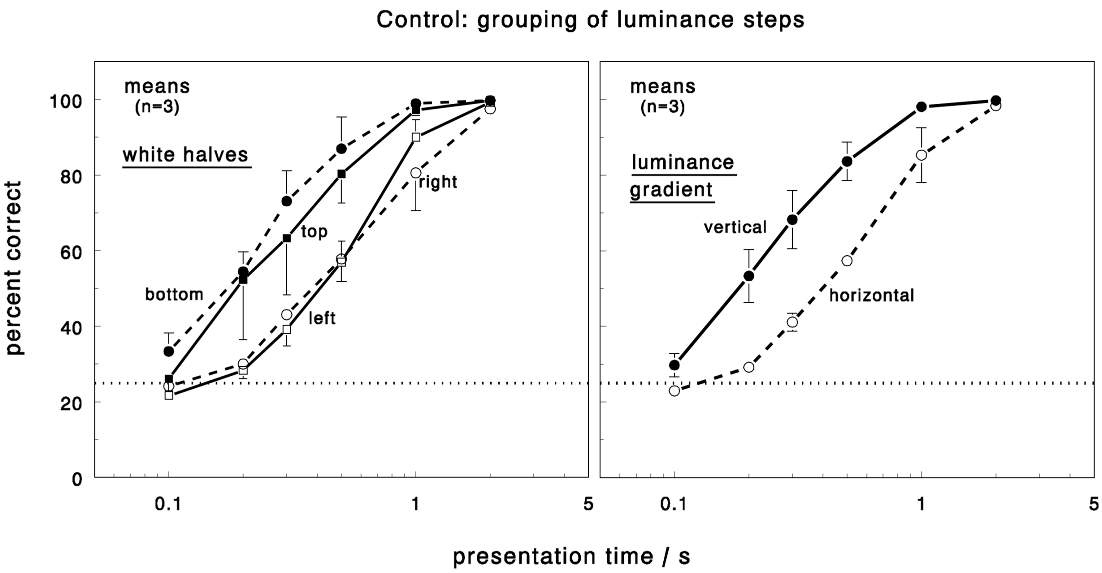

Figure 20. *Mean performance in Experiment 11 (control).* Even though targets cannot be distinguished for apparent depth effects, performance was similar to, and partly even better than that in Exp. 10. Triangles with vertical luminance variations were, in general, better detected and distinguished from the different target than targets with horizontal luminance variations. Larger s.e.m. bars than in Fig. 18 indicate that there was more variability between subjects as to which luminance border was grouped best.

be grouped for salience already at presentation times of 160 ms (Nothdurft, 1992). Presentation times of up to 1s, like in the present experiment, over which grouping (for similar features) did still improve (Fig. 20) should have given the subjects sufficient time to identify and distinguish the four salient blobs.

Experiments 10 and 11 thus do not provide evidence that fast (in earlier studies "pre-attentive") grouping is provided by depth cues and feature identity. They only indicate that salient targets (that are already grouped in much shorter presentations; Nothdurft, 1992) can be distinguished and perceptually re-grouped for certain features. Depth cues, even monocular, might be quite helpful in this additional sorting process, as they have been shown to provide fast discrimination and popout (cf. Enns & Rensink, 1990). In the present study, however, the better performance with certain targets was not unequivocally due to an apparent depth-from-shadow effect. Experiment 10 showed that targets with a luminance gradient that generate the percept of depth are faster sorted (grouped) than targets with orthogonal luminance gradients that do not generate this percept. Experiment 11, however, showed that a similar difference is obtained with vertical vs. horizontal luminance steps that both do not generate a depth percept.

DISCUSSION

The experiments of this study have revealed two major findings that are, in part, very similar to findings previously obtained with other visual features (Nothdurft, 1992, 1993b). They have shown that target salience is provided by an increased *feature contrast,* here tested in luminance and disparity (Exp. 1 and 2). And they have shown that perceptual grouping does not require the analysis of features and item similarities but is often faster and more easily obtained from target salience as produced by local feature *differences* (Exp. 3-9). Only in a second and typically more time consuming process may salient items be re-grouped for feature similarity (Exp. 10 and 11), which is however ignored in the first perceptual grouping process. The experiments have also shown that the strength of salience is important; even small salience differences between targets may affect this first grouping process (Exp. 3 and 6).

Salience from feature contrast

It is not unexpected that a single target in a homogeneous array of different items (distractors) appears salient and pops out. This has been frequently reported for many features in visual search (cf. Wolfe, 1998) and also for the feature dimensions studied here (*luminance*, e.g. Borji, Sihite, & Itti, 2013; Engel, 1974; Irwin et al., 2000; Proulx & Egeth, 2008; Theeuwes, 1995; Turatto & Galfano, 2000; *disparity*, e.g., Theeuwes, Atchley, & Kramer, 1998) including monocular depth cues (Enns & Rensink, 1990; *depth-from-shading*, Braun, 1993). However, the important role of feature *differences* rather than features themselves has not always been obvious. While for luminance it is textbook knowledge, that the detection of a target or border is provided from luminance contrast, not luminance *per se* (as nicely demonstrated, for example, by Cornsweet, 1970), less agreement apparently exists about this aspect in binocular disparity. Even the question whether luminance contrast does contribute to visual salience has been disputed in the literature. While it is rather obvious that a single bright or dark stimulus is more salient than no stimulus at all, the salience aspect of luminance in natural scenes was questioned (Einhäuser & König, 2003) but has meanwhile been confirmed in several studies (Borji, Sihite, & Itti, 2013; Liu, Cormack, & Bovic, 2010; Parkhust & Niebur, 2004). Salience from disparity has recently found increased interest, mainly in computer vision studies (e.g., Ciptadi, Hermans, & Rehg, 2013; Fang et al., 2014; Niu, Geng, Li, &Liu, 2012; Wang, Perreira Da Silva, Le Callet, & Ricordel, 2013; Zhang et al., 2008) and the analysis of natural viewing conditions (e.g., Jansen, Onat, & König, 2009; Lang et al., 2012; Liu, Bovic, & Cormack, 2008; Liu, Cormack, & Bovik, 2010; Wang et al., 2012;). Thus, it is important to document salience thresholds in these two domains, as was done in Experiments 1 and 2. It may be interesting in the context of the present study that the voluntary fixation behavior of human subjects seems to prefer slightly higher contrast and gradients in luminance, but slightly lower contrast and gradients in disparity, compared to random gaze shifts over a scene, as if "the binocular visual system, unless directed otherwise …, seeks fixations that simplify the computational process of disparity and depth calculation" (Liu, Cormack, & Bovic, 2010, p. 14).

What might be surprising, however, is the observed dependence of threshold feature contrast with feature variations in the pattern background (Figs. 4 and 6). Search performance with non-uniform distractors has been addressed in several studies, but always from a different perspective (e.g., Duncan & Humphreys, 1989; D'Zmura, 1991; Nagy & Sanchez, 1990; Wolfe et al., 1992). Here I showed that when the target is embedded in a background of items with a continuous feature gradient, its feature contrast has to be increased over the background variation, for an easy detection. Above a certain background gradient, the required target contrast grew over-proportionally and soon reached values that could not be realized in the experimental setup. Similar observations have been reported for variations in orientation, color, and the direction of motion (Nothdurft, 1992, 1993b). Experiments 1 and 2 show that it also holds for luminance and disparity. Even in Random Dot Stereograms, a strong increase of disparity modulation in the background may partly mask the visibility of targets at a given disparity contrast, as is illustrated in Figure 21.

There is a caveat in these experiments, however. The systematic and constant feature gradient in backgrounds made patterns look fairly regular, and subjects might have detected some targets on these backgrounds not only by looking for salience but perhaps also by searching for irregularities in the pattern. Such irregularities can be seen in the pictures with large background variations in Figures 5 and 7. Some targets there though not particularly salient may be found from irregularities in the pattern structure. Therefore, to measure true salience variations presentation time has generally been short and patterns were masked in the experiments.

There were notable performance differences between some subjects. While this is perhaps not unexpected for binocular disparity (Fig. 6) where large variations in sensitivity are known to exist between individuals, the very large differences seen with luminance (Fig. 4) were unexpected. Subject SL could still detect salient targets on backgrounds, on which other subjects failed to reach the threshold performance with even the maximally achievable luminance contrast. Although this subject was highly trained in this task, I cannot exclude that she might not have also looked for pattern irregularities instead of target salience. Note that she nevertheless failed in the majority of trials with still larger background variations.

It is tempting to speculate about the reason why target threshold contrast increases so dramatically when background variations are increased, or, equivalently, why targets on strong background variations cannot be made

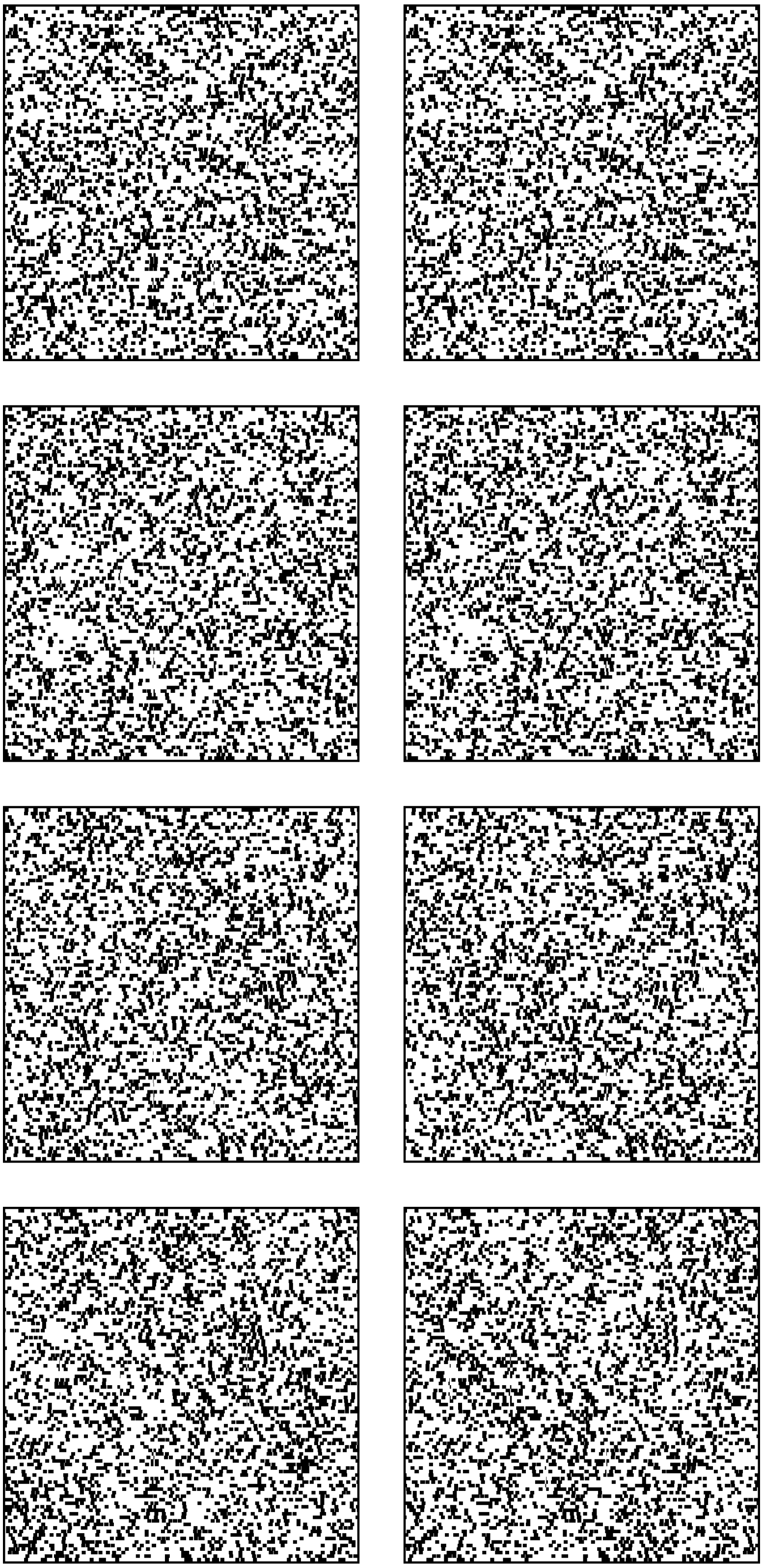

Figure 21. *Illustration of Experiment 2 (diminished salience of disparity contrast under large disparity variations in the background) with random dot stereograms.* Each stereogram shows four patches at clockwise increasing disparities (all near). The overall disparity variation across the pattern increases from top to bottom. While all patches should clearly be seen in the top stereogram (zero variation), their detection becomes increasingly difficult towards the bottom stereogram (large disparity gradient all over the pattern).

salient. As already mentioned, this was also observed in other feature dimensions and seems to be a general property of salience from feature contrast. In orientation, basically two models could account for this phenomenon (see also Discussion in Nothdurft, 1993c). One model would postulate that a small feature gradient in the background is still seen as a continuous orientation flow, from which local deviations or disruptions are then readily detected. But if background variations become too large, the continuous orientation flow will disappear, and hence no disruptions can be seen. This model is supported by numerous studies on the continuation of line paths in flanker regions around line segments ("association field"; e.g., Field, Hayes, & Hess, 1993). However, while this model may work for orientation, I do not see how it could be applied to luminance, color, or depth. What would be a good continuation of feature flows in these dimensions? An alternative explanation, still for the salience of orientation contrast, is given by the contextual modulation observed in certain neurons of the early visual cortex. Both in area V1 of macaques (Knierim & Van Essen, 1992; Nothdurft, Gallant, & Van Essen, 1999; Zipser, Lamme, & Schiller, 1996) and in area 17 of cats (Kastner, Nothdurft, & Pigarev, 1997; Li & Li, 1994; Nelson & Frost, 1978; Sillito et al., 1995) there are neurons the responses of which are strongly modulated by stimuli presented in the nearby non-responsive regions around the classical receptive field. In certain cells, the responses to a line stimulus in the classical receptive field are strongly suppressed by parallel lines outside, but less when the lines outside are oriented differently. The activity distribution of such cells over a stimulus pattern with a homogeneous background and one contrasting line should give a strong response to that line (little suppression from the surround) and small responses to all lines in the background (strong suppression from their surrounds). With an orientation gradient in the background, however, the suppression for background items should be reduced and, if the orientation gradient is strong, responses should become virtually indistinguishable from the response to the target line. In this model, the failure to detect an orthogonal target on a background with large orientation variations between items would not reflect the inability of the visual system to constitute a congruent orientation flow in the background but would rather reflect the increased salience of background items from which the salience of the target could not be distinguished. One advantage of this latter model is that it could easily be transferred to other features (cf. Gao, Mahedevan, & Vasconcelos, 2008). In monkey primary cortex, for example, contextual modulation of motion was earlier reported than contextual modulation for orientation in monkey area V1 (Allman, Miezin, & McGuinness, 1990). For luminance, the existence of feature-specific suppression from the receptive field surround is known about as long as properties of texture segmentation have been studied (e.g., Enroth-Cugell & Robson, 1966) and also color-specific contrast has meanwhile been demonstrated (see, e.g., Conway, Hubel, & Livingstone, 2002; Shapley & Hawken, 2011, for reviews). To my knowledge, only disparity contrast around the classical receptive field was not yet systematically studied, although interocular suppression was among the very early findings in cortical neurons (Barlow, Blakemore, & Pettigrew, 1967; Pettigrew, Nikara, & Bishop, 1968; Bishop, Henry, & Smith, 1971) and response variations with disparity have meanwhile extensively been studied in cat and monkey primary visual cortex (e.g., Ferster, 1981; Poggio & Fischer, 1977; Poggio & Talbot, 1981; for reviews see, e.g., Gonzalez & Perez, 1998, and Poggio, 1995).

Grouping for similarity?

In the grouping experiments of the present study, subjects were asked to identify the form of global figures that were combined from various salient targets in the pattern. Grouping of *similar* targets has been proposed by Gestalt psychologists (e.g., Wertheimer, 1923) as a major rule in perceptual organization. But the present experiments show that this proposal is not always correct. The grouping of similar targets was seriously disturbed by adding a fourth, different target and even dissimilar targets did perceptually group when they were equal-salient. The only similarity in grouping then was the *similar salience* of targets. But since salience is associated with feature differences, not features themselves, it is obvious that grouping for similar features did not work here. Indeed, bright and dark blobs did perceptually group together, as did blobs with convergent and blobs with divergent disparities. Grouping of dissimilar targets can also visualized in Figures 17 and 19 used to test grouping of depth-from-shadow effects. A quick look at the figures in a and b will show you squares of four salient targets; the triangles of similar targets can only be found from prolonged and careful inspection.

The observation that, in the present experiments, grouping of dissimilar items was not always as good as the grouping of similar items can be explained by differences in salience. In the used luminance scaling (Fig. 2), dark targets were less salient than bright targets with the same difference in luminance values, as was experimentally verified. If this salience difference was compensated (Exp. 5), the grouping for salience improved and figures from similar and dissimilar targets were equally well detected. The compensation done at the time of the experiments, however, was incomplete; meanwhile more details on salience matching of dark and bright targets are available (Nothdurft, 2015a, b). Salience differences were also seen with disparity; far targets generally appeared less salient to the observers than near targets. Whereas most early studies on the disparity sensitivity of cortical neurons did not mention such a bias (Ferster, 1981; Maunsell & Van Essen, 1983; Poggio & Talbot, 1981; see also Gonzaley & Perez, 1998, and Poggio, 1985), several newer studies have noticed similar differences between "near" and "far" disparities in various tasks (Calabro & Vaina, 2011; Kasai, Morotomi, Katayama, & Kumada, 2003; Trotter et al., 1996; Nienborg & Cumming, 2007; Hinkle & Connor, 2005; Jansen, Onat, & König, 2009; Tanimoto et al., 2004, Wang et al., 2012). In the present study, it was not possible to fully compensate for these differences in salience (cf. discussion of Exp. 8 above) but already a gradual compensation did strongly change the grouping performance (Fig. 13). When all targets were "near" at the same relative disparity contrast, hence equal-salient, (fast) grouping was entirely restricted to salience, i.e., feature *differences*, not to absolute disparity, i.e. feature similarity.

Several reports in literature document a special role of disparity in grouping; disparity helps to perceive *surfaces* at different depths, upon which various items can then be sorted and grouped (He & Nakayama, 1992, 1994). Surface integration affects search (Wheatley, Cook, & Vidyasagar, 2004); search for targets is easier within than across depth planes (Kim & Verghese, 2014). Attention can be selectively applied to certain depth planes (Bauer et al, 2012); switching attention between surfaces needs time (cf. Cobo, Pinilla, & Valdes-Sosa, 1999). Disturbing textures in a scene can be perceptually "removed" when appearing on a different depth plane (Zhaoping, Guyader, & Lewis, 2009). All these findings seem to suggest that in disparity indeed features, not feature differences might be important for grouping. This is an interesting issue that should be addressed in more detail in a future study. It seems that only smooth variations in disparity can be seen as a homogeneous surface whereas local discontinuities in disparity appear salient and pop out (cf. Wismeijer, Erkelens, van Ee, & Wexler, 2010). The present data show that grouping is also obtained across different depth planes, when targets are equally salient and/or occur at the same relative disparity (Exp. 9). However, depth is not unique and surface integration can be also obtained from color and texture (Yin, Kellman, & Shipley, 2000). Also motion differences may constitute the percept of different "surfaces" (Cobo, Pinilla, & Valdes-Sosa, 1999) upon which different items can then be sorted and better detected than when the surfaces "collapse" (Nakayama & Silverman, 1986). Nevertheless, items moving in different directions may also group for salience from motion *contrast* (Nothdurft, 1993b).

Depth from shading

The strong percept of depth when the luminance distribution of an item resembles the shadow of a three-dimensional object illuminated from above, was demonstrated many years ago (Kleffner & Ramachandran, 1992; Ramachandran, 1988; Todd & Mingolla, 1983) and shown to reveal texton quality (Braun, 1993). That is, targets can be seen to group and perceptually segment from other items, are fast detected even in large assemblies (thus, apparently "in parallel") and detection does not require focal attention (thus, is seemingly "pre-attentive"). This was also seen in the present data. However, since the time these experiments had been performed (1990/91), several new studies have elucidated important details of shape from shading. It appears that the 3D impression is partly learned and may slightly vary across cultures with different reading and writing preferences (Andrews, Aisenberg, d'Avossa, & Sapir, 2013). The association with convex or concave curvatures can be modified by opposite haptic sensations (Champion & Adams, 2007) or mental relocation of the imagined light source (Proulx, 2014), which both, however, does not change the asymmetry in fast visual search. Systematic orientation tuning studies have shown that for many observers the depth impression is particularly strong for lighting from the upper left side (Symons, Cuddy, & Humphrey, 2000, but see Sun & Perona, 1998). Despite all these minor variations, however, the general impression is, and also was for the subjects in the present study, that of three-dimensional objects illuminated from above. All subjects could group targets with these properties much faster, i.e. in shorter presentations, than targets with assumed horizontal lighting variations. Problems may arise when items are placed too closely together (Kawabe & Miura, 2004), which was not the case in the present study. Interestingly, several studies report a similar bias for vertical over horizontal luminance variations with luminance steps ("bipartite" luminance profiles; Kawabe & Miura, 2004; Kopp et al, 2010; Nadakumar, Torralba, & Malik, 2011; Symons, Cuddy, & Humphrey, 2000) that was also found in the present study (Exp. 11).

CONCLUSIONS

The present data confirm earlier proposals on the (nearly absent) role of features for salience and grouping, now also for luminance and disparity. Targets appear salient, if their feature contrast to neighbors is well above the feature contrast of other items in the pattern. Salient targets are detected fast and can be grouped to global figures, irrespective of the feature properties they display.

It is important to underline what this would mean for feature analysis. Salience is the result of feature differences, not features *per se*. Therefore, if items are perceptually grouped for similar salience, they are grouped for similar *differences*, not similar features. Several early grouping processes seem to be blind for features, and hence for similarities; the notion of objects being grouped for similarity is therefore wrong in this context. (Also from an economic or technical point of view, the detection of differences at any location should be much easier to implement than the fast ("in parallel") evaluation and integration of feature similarities over large distances in space.) Only in a second and maybe more time consuming process may salient items be evaluated and then perceptually grouped for feature similarity, as Experiments 10 and 11 have shown. But this later grouping can apparently only be performed on salient targets. Of particular interest may here be the grouping of items in different surfaces or depth planes (not studied in the present paper) which appears to provide a fast accessible structure in perceptual organization.

ACKNOWLEDGEMENT

I like to thank Jutta Tolksdorf for help in running the experiments, and all subjects for patient collaboration. Last not least, I am grateful to authors who still seem to appreciate the idea of similarity grouping in perceptual organization, which forced me to stress the contrary results in my experiments.

REFERENCES

Aks, D.J., & Enns, J.T. (1996). Visual search for size is influenced by a background texture gradient. *Journal of Experimental Psychology: Human Perception and Performance, 22* (6), 1467-1481.

Allman, J., Miezin, F., & McGuinness, E.L. (1990). Effects of background motion on the responses of neurons in the first and second cortical visual areas. In *Signal and Sense: Local and Global Order in Perceptual Maps* (Edited by Edelman, G.M., Gall, W.E., & Cowan, M.W.) pp. 131-142. Wiley-Liss, New York.

Andrews, B., Aisenberg , D., d'Avossa, G., & Sapir, A. (2013). Cross-cultural effects on the assumed light source direction: Evidence from English and Hebrew readers. *Journal of Vision, 13* (13):2, 1-7, http://www.journalofvision.org/content/13/13/2, doi: 10.1167/13.13.2.

Barlow, H.B., Blakemore, C., & Pettigrew, J.D. (1967). The neuronal basis of binocular depth discrimination. *Journal of Physiology (London), 193*, 327-342.

Bauer, D., Plinge, A., Ehrenstein, W.H., Rinkenauer, G., & Grosjean, M. (2012). Spatial orienting of attention in stereo depth. *Psychological Research-Psychologische Forschung, 76* (6), 730-735.

Beck, J. (1966). Effect of orientation and of shape similarity on perceptual grouping. *Perception & Psychophysics, 1,* 300-302.

Beck J. (1982). Textural segmentation. In *Organization and Representation in Perception* (Edited by Beck J.), pp. 285-317. Erlbaum, Hillsdale, NJ, London.

Ben-Shahar, O. (2006). Visual saliency and texture segregation without feature gradient. *Proceedings of the National Academy of Sciences of the USA, 103* (42), 15704-15709.

Bishop, P.O., Henry, G.H., & Smith, C.J. (1971). Binocular interaction fields of single units in the cat striate cortex. *Journal of Physiology (London), 216,* 39-68

Borji, A., Sihite, D.N., & Itti, L. (2013). What stands out in a scene? A study of human explicit saliency judgment. *Vision Research, 91,* 62-77. http://dx.doi.org/10.1016/j.visres.2013.07.016

Braun, J. (1993). Shape-from-shading is independent of visual attention and may be a texton. *Spatial Vision, 7* (4), 311-322.

Caelli, T., Hübner, M., and Rentschler, I. (1986). On the discrimination of micropatterns and textures. *Human Neurobiology, 5,* 129-136.

Calabro, F.J., & Vaina, L.M. (2011). Population anisotropy in area MT explains a perceptual difference between near and far disparity motion segmentation. *Journal of Neurophysiology, 105* (1), 200-208.

Champion, R.A., & Adams, W.J. (2007). Modification of the convexity prior but not the light-from-above prior in visual search with shaded objects. *Journal of Vision, 7* (13):10, 1-10, http://journalofvision.org/7/13/10/, doi: 10.1167/7.13.10.

Ciptadi, A., Hermans, T., & Rehg, J.M. (2013). An in depth view of saliency. *Proceedings of the British Machine Vision Conference 2013,* 1-11.

Cobo, A., Pinilla, T., & Valdes-Sosa, M. (1999). Attention to surfaces defined by transparent motion: Measuring dwell time. *Brain and Cognition, 40* (1), 85-90.

Conway, B.R., Hubel, D.H., & Livingstone, M.S. (2002). Color contrast in macaque V1. *Cerebral Cortex, 12* (9), 915-925. doi: 10.1093/cercor/12.9.915

Cornsweet, T.N. (1970). *Visual Perception.* Academic Press, New York.

Dannemiller, J.L., & Stephens, B.R. (2001). Asymmetries in contrast polarity processing in young human infants. *Journal of Vision 1(2):5,* 112-125. http://journalofvision.org/1/2/5/

Duncan, J., & Humphreys, G.W. (1989). Visual search and stimulus similarity. Psychological Review, 96 (3), 433-458.

D'Zmura, M. (1991). Color in visual search. Vision Research, 31(6),951-966.

Einhäuser, W., & König, P. (2003). Does luminance-contrast contribute to a saliency map for overt visual attention? *European Journal of Neuroscience, 18,* 1089-1097.

Engel, F.L. (1974). Visual conspicuity and selective background interference in eccentric vision. *Vision Research, 14,* 459-471.

Enns, J.T., & Rensink, R.A. (1990). Influence of scene-based properties on visual search. *Science, 247*, 721-723.

Enroth-Cugell, C., & Robson, J. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology (London), 187,* 517-552.

Fang, Y., Wang, J., Narwaria, M., Le Callet, P., & Lin, W. (2014). Saliency detection for stereoscopic images. *Ieee Transactions on Image Processing, 23* (6), 2625-2636.

Ferster, D. (1981). A comparison of binocular depth mechanisms in areas 17 and 18 of the cat visual cortex. *Journal of Physiology (London), 311,* 623-655.

Field, D.J., Hayes, A., & Hess, R.F. (1993). Contur integration by the human visual system: Evidence for a local 'association field'. *Vision Research, 33,* 173-193.

Foster, D.H., & Westland, S (1995). Orientation contrast vs orientation in line-target detection. *Vision Research, 35* (6), 733-738.

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.

Gonzalez, F., & Perez, R. (1998). Neural mechanisms underlying stereoscopic vision. *Progress in Neurobiology, 55* (3), 191-224.

He, Z.J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature, 359*, 231-233.

He, Z.J.J., & Nakayama, K. (1994). Perceiving textures – beyond filtering. *Vision Research, 34* (2), 151-162.

Hinkle, D.A., & Connor, C.E. (2005). Quantitative characterization of disparity tuning in ventral pathway area V4. *Journal of Neurophysiology, 94* (4), 2726-2737.

Irwin, D.E., Colcombe, A.M., Kramer, A.F., Hahn, S. (2000). Attentional and oculomotor capture by onset, luminance and color singletons. *Vision Research, 40,* 1443-1458.

Jansen, L., Onat, S., & König, P. (2009). Influence of disparity on fixation and saccades in free viewing of natural scenes. *Journal of Vision, 9* (1):29, 1-19, http://journalofvision.org/9/1/29/, doi:10.1167/9.1.29.

Julesz, B. (1962). Visual pattern discrimination. *IRE Transactions on Information Theory, 8,* 84-92.

Julesz, B. (1975). Experiments in the visual perception of texture. *Scientific American, 232* (4), 34-43.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature 290,* 91-97.

Julesz, B. (1986). Texton gradients: The texton theory revisited. *Biological Cybernetics, 54,* 245-251.

Kasai, T., Morotomi, T., Katayama, J., & Kumada, T. (2003). Attending to a location in three-dimensional space modulates early ERPs. *Cognitive Brain Research, 17* (2), 273-285.

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

Kawabe, T., & Miura, K. (2004). Perceptual grouping in shape from shading. *Perception, 33* (5), 601-614.

Kim, Y.J., & Verghese, P. (2014). The influence of segmentation and uncertainty on target selection. *Journal of Vision, 14* (3):3, 1-11, http://www.journalofvision.org/content/14/3/3, doi:10.1167/14.3.3.

Kleffner, D.A., & Ramachandran, V.S. (1992). On the perception of shape from shading. *Perception & Psychophysics, 52* (1), 18-36.

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.

Kopp, B., Kizilirmak, J., Liebscher, C., Runge, J., & Wessel, K. (2010). Event-related brain potentials and the efficiency of visual search for vertically and horizontally oriented stimuli. *Cognitive Affective & Behavioral Neuroscience, 10* (4), 523-540.

Landy, M.S., & Bergen, J.R. (1991). Texture segregation and orientation gradient. *Vision Research, 31,* 679-691.

Lang, C., Nguyen, T.V., Katti, H., Yadati, K., Kankanhalli, M., & Yan, S. (2012). Depth matters: influence of depth cues on visual saliency. *Computer Vision - ECCV 2012, Part II* (Edited by Fitzgibbon, A. et al.), pp. 101-115. Springer, Berlin, Heidelberg.

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.

Li, Z. (1999). Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proceedings of the National Academy of Sciences of the USA, 96,* 10530-10535

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences, 6* (1), 9-16.

Liu, B., & Todd, J.T. (2004). Perceptual biases in the interpretation of 3D shape from shading. *Vision Research, 44,* 2135–2145.

Liu, Y., Bovik, A. C., & Cormack, L. K. (2008). Disparity statistics in natural scenes. *Journal of Vision, 8* (11):19, 1-14, http://journalofvision.org/8/11/19/, doi:10.1167/8.11.19.

Liu, Y., Cormack, L.K., & Bovik, A.C. (2010). Dichotomy between luminance and disparity features at binocular fixations. *Journal of Vision, 10* (12):23, 1-17, http://www.journalofvision.org/content/10/12/23, doi:10.1167/10.12.23.

Maunsell, J.H.R., & Van Essen, D.C. (1983). Functional properties of neurones in middle temporal visual area of the macaque monkey. 2. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology, 49* (5), 1148-1167.

Motoyoshi, I., & Nishida, S. (2001). Visual response saturation to orientation contrast in the perception of texture boundary. *Journal of the Optical Society of America. A, 18* (9), 2209-2219.

Nandakumar, C., Torralba, A., & Malik, J. (2011). How little do we need for 3-D shape perception? *Perception, 40,* 257-271. doi:10.1068/p 6762

Nagy, A.L., & Sanchez, R.R. (1990). Critical color differences determined with a visual search task. *Journal of the Optical Society of America. A, 7* (7), 1209-1217.

Nakayama, K., & Silverman, G.H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320*, 264-265.

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

Nienborg, H., & Cumming, B.G. (2007). Psychophysically measured task strategy for disparity discrimination is reflected in V2 neurons. *Nature Neuroscience, 10* (12), 1608-1614.

Niu, Y., Geng, Y., Li, X., Liu, F., & Ieee (2012). Leveraging Stereopsis for Saliency Analysis. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. *454-461.* Providence, RI.

Nothdurft, H.C. (1985). Sensitivity for structure gradient in texture discrimination tasks. *Vision Research, 25,* 1957-1968.

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. (1993a). Saliency effects across dimensions in visual search. *Vision Research, 33* (5/6), 839-844.

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

Nothdurft, H.C. (1993c). The conspicuousness of orientation and motion contrast. *Spatial Vision, 7*, 341-363.

Nothdurft, H.C. (1995). Generalized feature contrast in preattentive vision. *Perception, 24,* S22.

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

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

Nothdurft, H.C. (2015). Luminance-defined salience of homogeneous blob arrays. *VPL-reports, 1,* 1-38, www.vpl-reports.de/1/

Nothdurft, H.C. (2015). Luminance-defined salience – targets among distractors. *VPL-reports, 2,* 1-97. www.vpl-reports.de/2/

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.

Olson, R.K., and Attneave, F. (1970). What variables produce similarity grouping? *American Journal of Psychology, 83,* 1-21.

Parkhurst, D.J., & Niebur, E. (2004). Texture contrast attracts overt visual attention in natural scenes. *European Journal of Neuroscience, 19* (3), 783-789.

Pettigrew, J.D., Nikara, T., & Bishop, P.O. (1968). Binocular interaction on single units in cat striate cortex: simultaneous stimulation by single moving slit with receptive fields in correspondence. *Experimental Brain Research, 6,* 391-410.

Poggio, G.F. (1995). Mechanisms of stereopsis on monkey visual cortex. *Cerebral Cortex, 5* (3), 193-204.

Poggio, G.F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving Rhesus monkey. *Journal of Neurophysiology, 40* (6), 1392-1405.

Poggio, G.F., & Talbot, W.H. (1981). Mechanisms of static and dynamic stereopsis in foveal cortex of the Rhesus monkey. *Journal of Physiology (London), 315,* 469-492.

Proulx, M.J. (2014). The perception of shape from shading in a new light. *PeerJ, 2:*e363; doi: 10.7717/peerj.363.

Proulx, M.J., & Egeth, H.E. (2008). Biased competition and visual search: the role of luminance and size contrast. *Psychological Research-Psychologische Forschung, 72* (1), 106-113.

Ramachandran, V.S. (1988). Perceived shape from shading. *Scientific American 259* (2), 76-83.

Sagi, D. & Julesz, B. (1987). Short-range limitations on detection of feature differences. *Spatial Vision, 2,* 39-49.

Shapley, R., & Hawken, M.J. (2011). Color in the cortex: single- and double-opponent cells. *Vision Research, 51,* 701-717.

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

Sun, J., & Perona, P. (1998). Where is the sun? *Nature Neuroscience, 1* (3), 183-184.

Symons, L.A., Cuddy, F., & Humphrey, K. (2000). Orientation tuning of shape from shading. *Perception & Psychophysics, 62* (3), 557-568.

Tanimoto, N., Takagi, M., Bando, T., Abe, H., Hasegawa, S., Usui, T., Miki, A., & Zee, D.S. (2004). Central and peripheral visual interactions in disparity-induced vergence eye movements: I. Spatial interaction. *Investigative Ophthalmology & Visual Science, 45* (4), 1132-1138.

Theeuwes, J. (1995). Abrupt luminance change pops out; abrupt color change does not. *Perception & Psychophysics, 57* (5), 637-44.

Theeuwes, J., Atchley, P., & Kramer, A.F. (1998). Attentional control within 3-D space. *Journal of Experimental Psychology - Human Perception and Performance, 24* (5), 1476-1485.

Todd, J.T., & Mingolla, E. (1983). Perception of surface curvature and direction of illumination from patterns of shading. *Journal of Experimental Psychology-Human Perception and Performance, 9* (4), 583-595.

Treisman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing, 31,* 156-177.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology,* 12,97-136.

Trotter, Y., Celebrini, S., Stricanne, B., Thorpe, S., & Imbert, M. (1996). Neural processing of stereopsis as a function of viewing distance in primate visual cortical area V1. *Journal of Neurophysiology, 76* (5), 2872-2885.

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

Wang, J., Le Callet, P., Tourancheau, S., Ricordel, V., Perreira Da Silva, M.(2012). Depth bias of observers in free viewing of still stereoscopic synthetic stimuli. *Journal of Eye Movement Research 5* (5):1, 1-11.

Wang, J., Perreira Da Silva, M., Le Callet, P., & Ricordel, V. (2013). Computational Model of Stereoscopic 3D Visual Saliency. *Ieee Transactions on Image Processing, 22* (6), 2151-2165.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychologische Forschung, 4,* 301-350.

Wheatley, C., Cook, M.L., & Vidyasagar, T.R. (2004). Surface segregation influences pre-attentive search in depth. *Neuroreport, 15* (2), 303-305.

Wismeijer, D. A., Erkelens, C. J., van Ee, R., & Wexler, M. (2010). Depth cue combination in spontaneous eye movements. *Journal of Vision, 10* (6):25, 1-15, http://www.journalofvision.org/content/10/6/25, doi:10.1167/10.6.25.

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

Wolfe, J.M., Cave, K.R., & Franzel, S.L. (1989). Guided Search: an alternative to the Feature Integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 419-433.

Wolfe, J.M., Friedman-Hill, S.R, Stewart, M.I., & O'Connell, K.M. (1992). The role of categorization in visual search for orientation. *Journal ol Experimental Psychology: Human Perception and Performance, 9 (1),* 14-49.

Yin, C., Kellman, P.J., & Shipley, T.F. (2000). Surface integration influences depth discrimination. *Vision Research, 40* (15), 1969-1978.

Zhang, L., Tong, M. H., Marks, T. K., Shan, H., & Cottrell, G. W. (2008). SUN: a Bayesian framework for saliency using natural statistics. *Journal of Vision, 8* (7):32, 1-20, http://journalofvision.org/8/7/32/, doi:10.1167/8.7.32.

Zhaoping, L., Guyader, N., & Lewis, A. (2009). Relative contributions of 2D and 3D cues in a texture segmentation task, implications for the roles of striate and extrastriate cortex in attentional selection. *Journal of Vision, 9(11):20, 1-22,* http://journalofvision.org/9/11/20/, doi:10.1167/9.11.20

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

Published online: 15-Dec-2015

For later additions and for comments see http://www.vpl-reports/3/