Cued visual selection of conjunction targets – no evidence of additional attentional requirements for the binding of color and orientation

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

The technique of cued visual selection (CVS) was used to measure dynamic processes in the identification of combined color and orientation targets. It has been proposed that the different features in such items must be attentively linked together for correct identification. In arrays of red and green lines at different orientations, one line (which thus became the target) was cued and had to be identified. Like with one-feature identification tasks in CVS, in which color is generally faster identified than orientation, observers also identified the color of combined targets faster than their orientation. Even in conjunction targets thus, features are identified largely independent from each other. False conjunctions were not obtained from a lack of attention but because one or the other feature was not yet correctly identified. When the performance in (separate) one-feature identification tasks was taken to predict the performance in the (combined) conjunction task, orientation identification was found to be slightly accelerated compared to the predictions. An analogue effect in color was not seen or notably smaller and in the opposite direction. Detailed analysis however showed that the improvement of orientation identification in conjunction tasks was not achieved on the cost of simultaneous color identification, nor was iFigs.t explained by learning effects or possible luminance differences in the tasks. It rather seems to reflect a better encoding of orientation signals in color channels or a better utilization of attentional resources in conjunction than in pure orientation tasks. Altogether there is no evidence that the attentional resources needed for target identification were also used for the binding of target feature components. *© Author*

Citation: Nothdurft, H.C. (2020). Cued visual selection of conjunction targets – no evidence of additional attentional requirements for the binding of color and orientation. VPL-reports, 11, 1-14, www.vpl-reports.de/11/

INTRODUCTION

It is a compelling observation in so-called *conjunction search* that an odd-colored item among differently colored items or a single line among orthogonal lines are both quickly detected but an item with a singular combination of both is not (Fig. 1). Apparently, the link between features is not immediately (and not preattentively) available to the visual system (Treisman & Gelade, 1980). Instead, when patterns are briefly presented, observers may see "illusory" (i.e., false) conjunctions between features presented at different locations (Treisman & Schmidt, 1982). These (and other) observations have led Ann Treisman to formulate her influential "feature integration theory" (FIT): Features are assumed to be represented in independent modules from which they can instantaneously be detected, but it would require focal attention to *bind* a certain feature in one module with a certain feature in another module (Treisman & Gelade, 1980; Treisman, 1985). Thus, for example, the single red line among green lines (Fig. 1) would be detected in the presumed feature module *red*, and the single right-tilted line among left-tilted lines in the feature module *oblique orientations tilted to the right*; but to detect the single left-tilted green line among items that partly share the target color or orientation (but not both) focal attention is needed

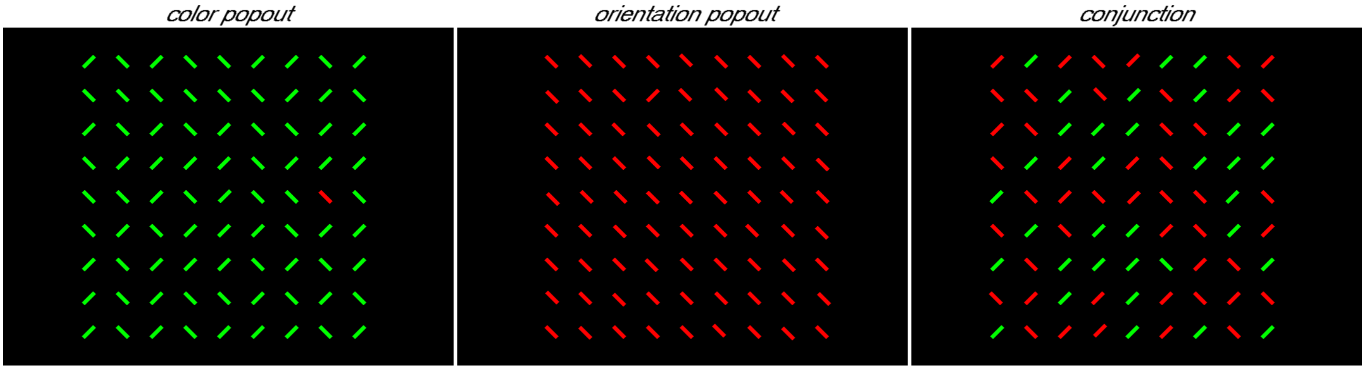


Figure 1. *Conjunction search of orientation and color.* Find in each pattern the unique item that differs from the rest. The pure color (left) or orientation singularities (middle) are quickly found; the combination singularity is not (right). Feature integration theory *(FIT)* claims that to find the odd item there, features must be attentively linked to each item.

to bind the different colors and orientations in both modules to individual items.

Since the original proposal of FIT there have been several attempts to confirm or disprove the postulate (see Wolfe & Cave, 1999, for an overview). Beyond an apparent need for modifications (Treisman & Sato, 1990) or alternative models (Wolfe, Cave, & Franzel, 1989; Wolfe, 1994) it turned out that the perceptual modules of *certain* feature combinations do *not* have to be attentively linked for the detection of conjunction targets. In some dimensions, instead, items perceptually segregate into sub-patterns with only one feature of that dimension. Items at different binocular disparity, for example, may appear to segregate into surfaces at different depth, in each of which an item with a singular feature in another dimension is then easily detected (Nakayama & Silverman, 1986). Independent perceptual filtering has also been reported for motion where the binding with color has revealed surprising dynamic variations (Vigano, Maloney, & Clifford, 2017) and perceptual asynchrony (Moutoussis & Zeki, 1997a, b; Viviani & Aymoz, 2001; Holcombe & Cavanagh, 2008; Rangelov & Zeki, 2014; McIntyre & Arnold, 2018). Also conjunctions of form and certain motion features (e.g., opposite directions, moving vs. stationary) are immediately and apparently pre-attentively detected (McLeod, Driver, Dienes, & Crisp, 1991). One of the few apparently still "valid", i.e. non-rejected FIT pairings is orientation and color (see Fig. 1). It is apparently not easy to perceptually segregate features in these dimensions so that a singular combination will be immediately detected[1]. Even with color and orientation, however, search strategies may quickly make use of inhomogeneities or unequal feature distributions in the pattern and thus optimize the detection of a singular conjunction item (Zohary & Hochstein, 1989).

With the technique of *cued visual selection* (CVS, Nothdurft, 2017a, see also Nothdurft, 2002, 2006) it is possible to direct attention to a target and measure attention effects as the time required to identify that target. Can this technique help us to test the role of feature binding in target identification? Obviously, if FIT would require attention to bind the different features to one object, directing attention to that object as with CVS cannot prove or disprove the model. But it would be interesting to see whether there are special attentional requirements for feature binding, on top of the attentional requirements needed for target identification. In particular, one should conclude from FIT that a conjunction target when focal attention is directed to it should be correctly

---

[1] It is not impossible to achieve such a segregation, though. If you are able to concentrate on the green lines in the *conjunction* pattern in Figure 1 and ignore all red lines, the green left-tilted line may (mildly) pop out. Without tricks (and unless you are suffering from astigmatism in one or both eyes) it is more difficult to "concentrate" on one line orientation, e.g. on all lines tilted to the left. One such trick would be to change the apparent line contrast, e.g., by looking at the (flat) Figure 1 from the lower right corner so that left-tilted lines appear short and more contrasting than right-tilted lines. Among the more contrasting lines you may then immediately see the single green line popping out.

identified and should not generate the percept of illusory conjunctions (Treisman & Schmidt, 1982) which are assumed to occur when different features are *not* appropriately linked together under focal attention. This conclusion was, however, not verified in the present study. Even in conjunction targets, different feature components were processed independently and at their own dynamics; the necessary presentation time for correct target identification did strongly vary between features. The target color, for example, was often already identified when target presentation was yet too short to let observers also discriminate target orientation. In such cases, observers might have indicated a false conjunction even though attention was directed to the target.

While the faster identification of color than orientation corresponds well with pure-feature studies (Nothdurft, 2020; see also Cheal & Lyon, 1992), it would be interesting to see if both identification processes interact when performed in parallel. For independent processes, performance in the conjunction task can be predicted from the identification probabilities of each feature alone. It should be interesting to see whether these predictions are met in conjunction tests, or whether the identification of combined targets would suffer from capacity limitations in attention, in particular if additional attentional resources were needed to bind these features together. That the identification of conjunction targets takes longer than the identification of one-feature patterns is suggested from studies which have reported different processing times per item between such patterns (Dugué, Xue, & Carrasco, 2017). The opposite was found in the present study. Conjunction targets were generally identified a little faster than predicted. Altogether, performance accuracy with conjunction targets was not limited by a (common) attentional mechanism that had to be shared in the identification of color and orientation features. There was no obvious deployment of attentional resources to integrate both features and no indication of any extra cost for feature binding.

METHODS

Overview

The experiment reported here was part of a larger series of experiments designed to study CVS dynamics with different features, including orientation and color. The present paper studied the *combination* of these latter two features in "conjunction" targets, i.e. red or green lines at one of two (oblique) orientations (see Fig. 2). In a large array of 80 items with random attributes, one item *(the target*) was cued and thus selected; observers then had to report the identity of this target. Shortly after cue presentation, all items (including the target) were masked by bright (white) crosses composed of the two possible line orientations. Two parameters were systematically varied in the course of the experiment; the *cue delay* (after stimulus onset) and the stimulus presentation time (*target duration*) between the cue and the mask onsets. Performance accuracy was averaged over similar trials with newly randomized patterns; all four possible feature combinations served as potential targets.

Accuracy in performing the conjunction task (measured here) was compared with the accuracy in detecting one or the other feature and, in addition, with the performance in separate one-feature identification tasks (measured in the accompanying study; Nothdurft, 2020). These comparisons revealed different dynamics of individual feature components and allowed me to compare the analysis of single features with that of combined features.

Stimuli

Stimuli were generated with DOS VGA techniques on a 17" color monitor (Sony Trinitron multiscan 17se II) in front of the observer. Monitor frame rate was 60 Hz. (In the two test series of the parallel study, the data of which will be used for comparison, different frame rates were used; 60 Hz for orientation, and 100 Hz for color.) Viewing distance was 73 ±1.5 cm, with small variations due to head size differences between observers (who had their heads conveniently leaned against the wall).

Patterns displayed red or green lines in a rectangular 9 x 9 raster, with a raster width of 1.8 deg; full stimulus patterns covered an area of approximately 15 deg x 15 deg. The center element of the raster was spared and instead a white fixation cross (0.25 deg x 0.25 deg) was shown. Lines were 0.8 deg x 0.2 deg and red or green; colors were matched for equal luminance by means of heterochromatic flicker photometry. At various delays after stimulus onset, one of these lines was cued (50 ms), and after a variable presentation time thereafter all items were masked. (Similar test patterns were used in the accompanying study except that lines (orientation) were white and color items were non-oriented squares.)

Cues were made of four little squares around the target *(four-dot cues),* each 0.2 deg x 0.2 deg, which were

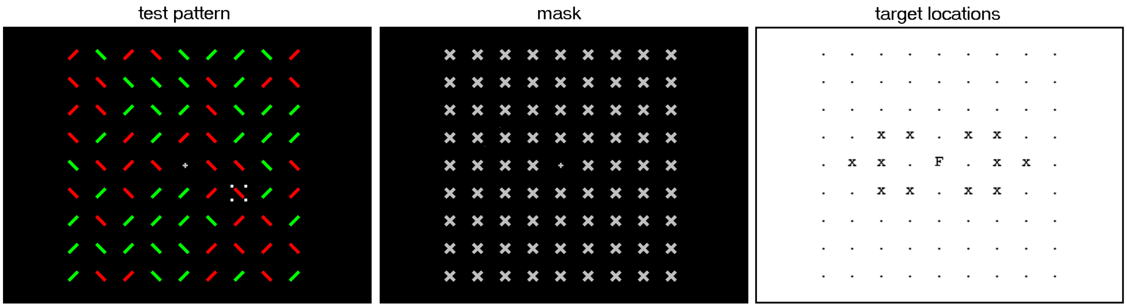


Figure 2. *Test patterns used in experiment.* Patterns showed regular arrays of 80 lines randomly assigned to one of two orientations and one of two colors, plus a central fixation cross (*F* in the right-hand scheme). At various delays after stimulus onset, one line was selected with a 50 ms *four-dot cue* (as shown) and observers had to identify this target. After a short presentation time, *Δt,* measured from the cue onset the whole pattern was masked. To avoid too large performance variations from crowding and limited attentional resolution, targets were selected at fovea-near positions (*x* in the right-hand scheme). The full stimulus pattern covered an area of about 15 deg by 15 deg.

located 0.6 deg from the target center in the four oblique directions (Fig. 2). To reduce the large performance variations from crowding and limited attentional resolution (Intriligator & Cavanagh, 2001; Nothdurft, 2017a) but still keep the uncertainty of cued locations large enough, possible *target locations* were restricted to fovea-near raster positions as indicated in Figure 2. Subjects were not informed about this restrain.

Luminance settings of colored lines and white masks were 21 cd/m^2 and 28 cd/m^2, respectively; for some observers, however, the individual luminance matches of colored and equiluminant white lines (using heterochromatic flicker photometry) revealed slightly larger values (22-27 cd/m^2). White lines in the accompanying study were 20 cd/m^2. Four-dot cues (68 cd/m^2) and the fixation marker (about 44 cd/m^2) were notably brighter when measured with an extended test stimulus. All stimuli were shown on a dark background (about 10 cd/m^2).

Procedures

Patterns were viewed binocularly. Trials started with a 1s presentation of the fixation point before the stimulus pattern was shown; cues (50 ms) were superimposed upon the pattern at various delays (0 ms, 50 ms, 100 ms, 150 ms, 200 ms, 300 ms, 400 ms, 500 ms, 750 ms, 1000 ms, 1500 ms, 2000 ms, 5000 ms). After the presentation time (measured from the cue onset), the pattern was masked (500 ms). Thereafter the screen was blanked and only the fixation point remained visible, until subjects had entered their responses. After a short moment with a blank screen, a new trial began with the 1s presentation of the fixation point.

Responses were made in a modified four-alternative forced choice task (Nothdurft, 2017b) on a computer keyboard. The modifications were: (i) Subjects could reject, and later repeat, a trial if they felt they had been inattentive during the presentation or had lost fixation during long delays; and (ii) they could change their last response immediately after the trial if they had erroneously pressed the wrong key. Response keys were assigned to finger key presses with both hands; left-hand, "<" and "y" keys for red targets tilted to the left and right, respectively; right-hand, "." and "-" keys for green targets tilted to the left and right, respectively (German keyboard layout). The key assignments were quickly learned and memorized; in addition, sketches of these assignments were shown below the monitor. There was no time pressure in responding and observers could take any time to assign the identified target to the appropriate fingers and keyboard keys.

Tests were blocked for cuing delays; only target duration was varied within a block. The test range of target durations in each block was individually adjusted for each observer to provide a good coverage between chance performance (25% accuracy) and perfect target identification (100% accuracy). This has led to slightly different test ranges and resolutions for different observers. Within blocks, test conditions (target durations) were randomly intermixed, with 5-10 repetitions each. In the course of the experiment, the different blocks (delays) were repeated in interleaved sequence, to generate a final data base with usually 50 repetitions of every test condition. Experiments were carried out in sessions of 2h, each covering several test runs. Subjects could pause whenever they wanted. Usually, tested features (here, conjunctions) remained the same between subsequent blocks, and only the delay was changed. In the course of a full session, however, if there was time left, there might have been a switch from the conjunction task reported here to a feature studied in the accompanying paper.

All tasks were performed *under fixation*. Good fixation performance was checked with a video camera above the monitor which was focused upon the observer's eyes. All subjects had quickly learned to perform the CVS identification task without moving their gaze. Most targets were identified from short presentation times after the cue (target durations < 200 ms), for which there should have been no benefit from moving the eyes (Fischer *et al*., 1993).

Analysis and comparison with feature identification in the accompanying study

Data analysis was made offline after completion of the full test series. First, data in the (measured) conjunction task were analyzed for the correct identification of target components, i.e. either orientation or color, irrespective of whether the second feature in the conjunction target was correctly identified or not. Second, performance in the (current) conjunction task was compared with accuracy measurements on each feature alone obtained in the accompanying study (Nothdurft, 2020). These data were collected in separate, usually interleaved test sessions with nearly identical procedures in the same setup. There was one major difference, though; the identification of color targets had been measured with a monitor frame rate of 100 Hz for a better temporal resolution; all other tests were made with 60 Hz frame rate. For comparison with the current data, therefore, accuracy data from the 100 Hz color task in Nothdurft (2020) were transformed to 60 Hz resolution by linear interpolation (and extrapolation, if 100% values had already been reached).

Subjects

The same five observers who served as subjects in the accompanying study, also participated in the current experiment. Four of them (20-23 years old) were students at the Göttingen University and were paid for the time they spent in the experiment. The fifth observer was the author (68 years when the experiment began). All subjects had normal or corrected-to normal visual acuity on both eyes and, except the author, were naive as to the aim of the experiment. All subjects had carried out other experiments with cued target identification before.

RESULTS

The general data analysis in the presentation is based on the mean performance data of all observers. The various observers revealed very similar results, with similar psychometric curves, and individual performances are resolved in the presentation when necessary.

Independent feature analysis in conjunctions

Figure 3 shows the mean performance data in the conjunction task (red data points). Observers had four keys to identify the four possible targets (red or green lines tilted to the left or right); chance accuracy was 25% (red dashed line). Obviously, accuracy increased the longer the target was visible after the cue. Of particular interest are the responses in which either the target color (blue data points) or the target orientation were correctly identified (black data points), irrespective of whether the second feature was correctly seen or not. These rates lie well above the rates of fully identified targets (red data points) but show a strong difference between the features. Color identifications were much better, and reached a much higher accuracy at short presentation times than orientation identifications. In these one-feature analyses, chance performance was 50% (black dotted lines in Fig. 3). Note that the observers' task was not to identify the color of the target and then eventually its orientation or *vice versa*, but observers were asked to identify one of four possible targets (displaying one of two colors and one of two orientations). Still, the chance for an observer to correctly identify the color of a target but not yet its orientation was high at short target durations. The correct identification of the conjunction targets can be predicted by multiplying the probabilities of single-feature identification rates (gray lines). When both features were poorly identified (at short target durations), predictions lay well below the individual feature identification rates (e.g., for chance performance rates, $0.5 \cdot 0.5 = 0.25$). However, when one feature of the target (usually its color) was correctly seen (accuracy near 100%, i.e. identification

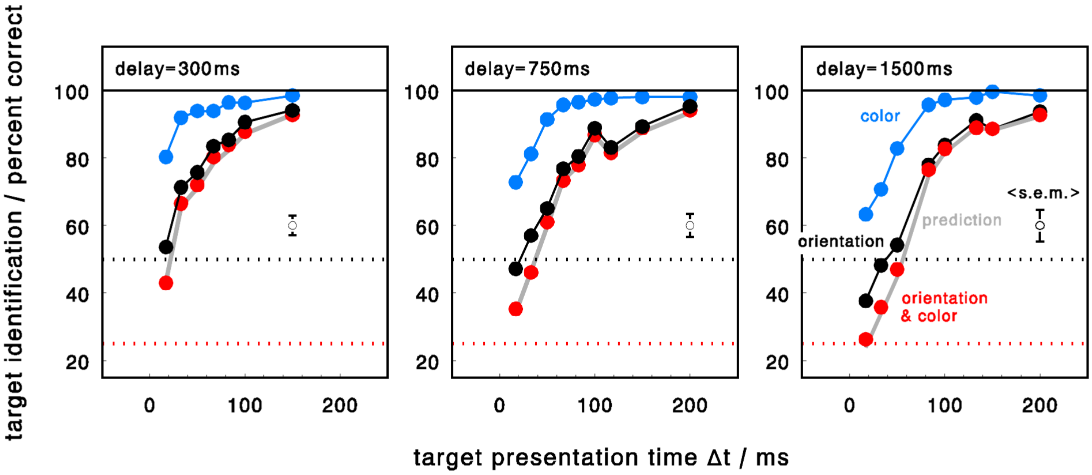


Figure 3. *Accuracy variations with increasing presentation time, at different delays.* Graphs show the correct identification of single features in combination targets (blue, color; black, orientation), irrespective of how the second feature was identified. Fully correct identifications (color and orientation) are shown in red. With all curves, accuracy increases with presentation time. Even though four individual targets had to be distinguished in the task, target color was always faster identified than target orientation. Data plot the mean accuracy of all five observers, plus averaged s.e.m., at three different cue delays. Gray curves show the presumed performance with combined features predicted from the independent performances with either feature alone. Predictions are here identical with the measured data, since all data were obtained in the same task.

rates close to 1), the identification accuracy of conjunction targets was close to that of the other feature component alone.

The fact that the measured identification rates with conjunction targets (red data points) fall exactly upon the probability predictions from independent feature components is not surprising. All data points in Figure 3 were, in fact, taken from the same task in which observers had to identify targets with two feature components. Correct color and correct orientation data should then necessarily predict the performance in the (measured) conjunction task. What is interesting (and important) in these data, however, is the large difference between the single-feature curves. It indicates that target features were processed and identified independent of each other. The much faster and better performance with color than with orientation was also seen in the accompanying study in which the same features were tested separately (Nothdurft, 2020). This difference between features was thus maintained in the conjunction task, in which the features were combined.

It should then be interesting to compare the present conjunction task directly with the measurements of identification rates on separate features, which had been made in the accompanying study. While the data in Figure 3 replicate the finding that color and orientation are identified at different speeds, it would be particularly interesting to see whether or not performance in the conjunction task can be exactly predicted by the performance in separate tasks. This comparison is made in Figure 4.

The identification of color targets (blue) and the identification of orientation targets (black) show similar differences as those in Figure 3. But in Figure 4, data were obtained from separate one-feature tasks, a pure-color and a pure-orientation discrimination task, while in Figure 3 all data were obtained from the (same) conjunction task. Again, we may predict the identification of combined targets by multiplication of the probabilities obtained in the two independent single-feature tasks and compare these predictions with the true performance in the (present) conjunction task (gray curves vs. red data points). In this comparison, we now make use of three separate test series, and real data might differ from the predictions. Predictions show the same properties as in Figure 3; when identification rates were low for both features, predicted performance in the conjunction task should even be lower. When one feature identification rate was nearly 100% (color), performance in the conjunction task should be almost identical to performance in the other

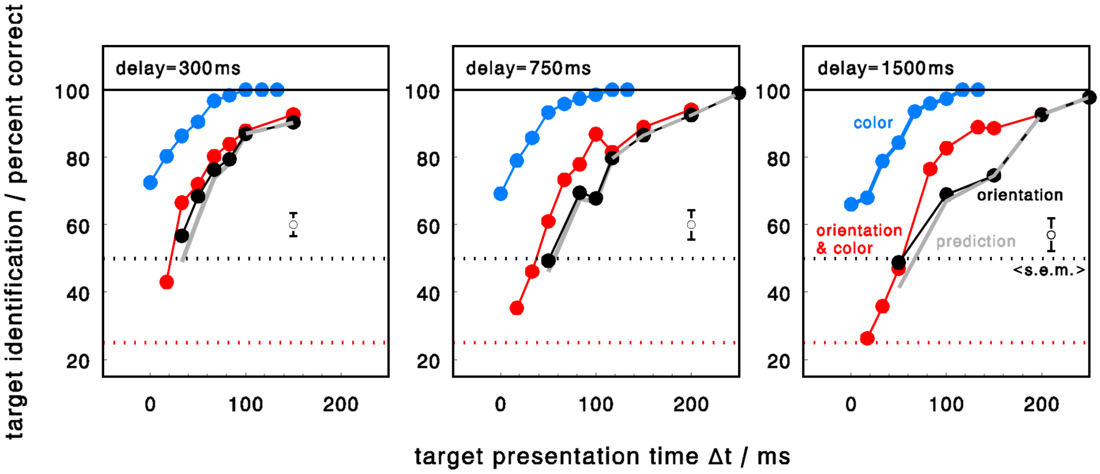


Figure 4. *Accuracy variations in the same conjunction task as in Figure 3 now compared with separate tests on either feature alone (blue, color; black, orientation).* Data presentation as in Figure 3 except that color and orientation discrimination was now measured in single-feature tasks (data adapted from Nothdurft, 2020; see text for details. Only test conditions are shown that were tested with all observers.) Performance in the conjunction task (red data points) was now sometimes better than predicted (gray). These deviations are analyzed in the following figures.

feature task (orientation). Different to Figure 3, however, the true measures from the conjunction task in Figure 4 (red data points) sometimes deviated from these predictions. Targets in the conjunction task were sometimes better identified than expected from the identification of independent feature components. This finding is remarkable and will be further analyzed below.

Statistics. Before that, however, it would be adequate to look at the "reliability" of these effects. The differences between features in the conjunction task (Fig. 3) were highly significant. The Wilcoxon signed-rank test revealed $|z| = 18.92$ ($p<0.001$) for all pairs of color and orientation identification, at same delays and with same target durations. (Note that for large N, here $N = 597$, the sampling distribution of the Wilcoxon test approaches a normal distribution and the test statistic W can be transformed into a z-score.) This is already obvious in Figure 3 where data points are separated by multiples of the standard errors of the mean. Differences were significant not only in the full data set from all observers but also in the data of each individual observer ($N \geq 111$; $|z| \geq 7.83$; $p<0.001$). The according differences in Figure 4 were already analyzed in Nothdurft (2020) and are highly significant, too ($|z| = 9.69$; $p<0.001$ for all observers, and $|z| \geq 3.82$; $p<0.001$, for the individual observers).

Deviations between feature identification rates in single and conjunction tasks

In this section, the apparently better feature identification of conjunction targets than predicted from single-feature tasks (Fig. 4) is further analyzed. For each cue delay and target presentation time, accuracy measures of one feature identification in the conjunction task were compared with the same accuracy measures in the according single-feature task, and deviations between the measures were computed. Positive values indicate that the feature was better identified in the conjunction task, negative values that it was better identified in the single-feature task. Deviations were calculated for both orientation and color, on the data of each individual observer, and are pairwise (same delay, same presentation time) plotted in Figure 5. The cloud distribution already shows that orientation identification was notably better, and color identification slightly worse in the conjunction tasks. Over all observers, the mean deviation in all conditions was 0.070 for orientation (corresponding to an overall accuracy improvement of 7 % in the conjunction task), and -0.016 for color (corresponding to an overall accuracy loss of 1.6 % in the conjunction task). For individual observers, the improvements in orientation

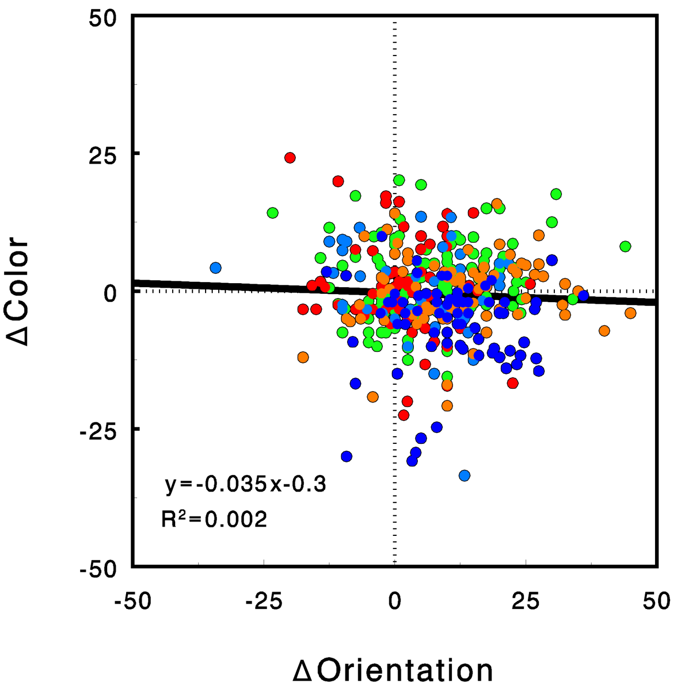


Figure 5. *Deviation scatter showing performance differences between conjunction and single-feature tasks for orientation and color.* Data from different observers are color-coded. Each data point is based on accuracy measures at the same delay and the same presentation time; a positive deviation indicates better performance in the conjunction task, for this particular test condition. On average, deviations in orientation were larger than deviations in color but deviations in both features were not correlated (flat regression line; R^2 near 0).

identification was stronger, with accuracy differences of up to almost 11 %.

Statistics. For statistical analysis all pairs with 100% accuracy in both tests were removed, because no deviations can be measured between perfectly seen targets. For the Wilcoxon signed-rank test, also all zero deviations must be removed. Statistical analysis was then performed on the remaining 508 (orientation) and 481 deviations (color) from all observers. Deviations were significantly positive for orientation ($|z| = 12.93$; $p<0.001$) and negative for color ($|z| = 3.66$; $p<0.001$). In the data of individual observers, however, the pattern is more variable. For orientation, overall deviations were significant in four observers ($N \geq 97$; $|z| \geq 5.50$; $p<0.001$) and not significant in observer HCN ($|z| = 0.54$). For color, the overall deviations were significant in only one observer (OC; $N = 101$; $|z| \geq 7.64$; $p<0.001$) and not significant in anyone of the other four observers ($N \leq 128$; $|z| \leq 1.47$).

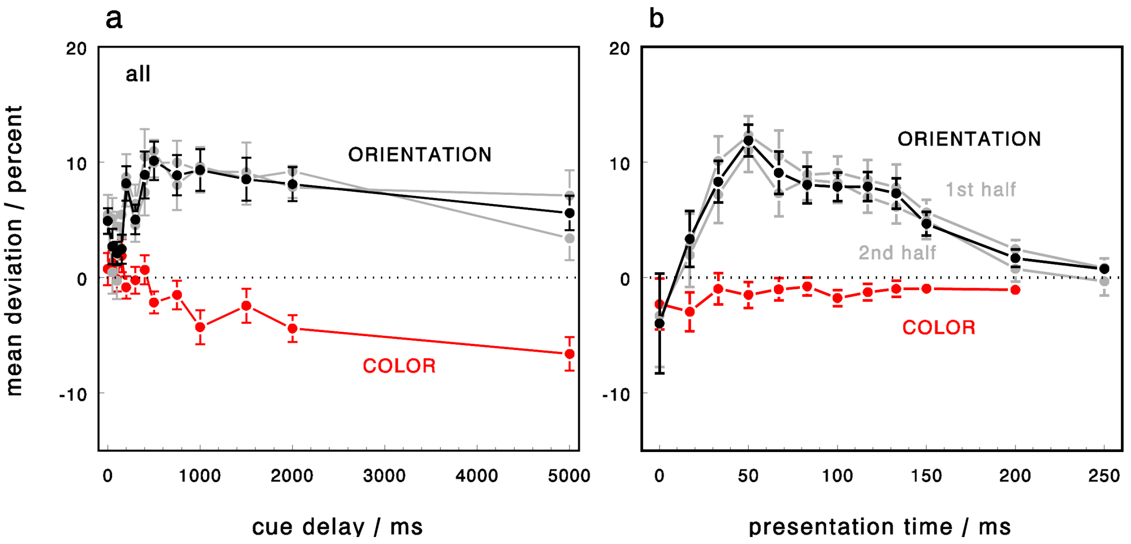
In the following I will look at certain properties of these distributions. First, it would be interesting to see if there were any compensation effects. The overall data suggest that the better orientation identification in the conjunction task might have been achieved from a poorer performance in simultaneous color identification. In Figure 5, however, there is no strong anti-correlation in the data. The regression line is almost flat and the coefficient of determination, R^2, is almost zero (R^2=0.002). For some observers, regression lines were slightly steeper (slopes down to -0.159) but coefficients of determination were not notably increased (maximum $R^2 = 0.045$). Thus, the better identification of target orientations in conjunction tasks was not associated with a poorer identification of target color.

Second, I wondered whether the deviations might have varied with either the cue delay (i.e., the time passed since stimulus onset) or target duration (i.e., the time during which the target remained visible after the cue). There was a slow overall modulation with the cue delay and a dynamic modulation with the target duration after presentation of the cue (Fig. 6). To visualize these variations, all deviations for a given cue delay (Fig. 6a) or target duration (Fig. 6b) were averaged, irrespective of the duration and the delay, respectively, and means across all observers were plotted when based on more than 10 entries. Variations with the cue delay (Fig. 6a) show that the mean deviations increased during the first 500 ms after stimulus onset and then diminished slowly (orientation) or continuously diminished to increasingly negative values (color) indicating that color discrimination deteriorated with increasing delays in the conjunction task. With target duration (Fig. 6b) mean deviations for orientation show a clear modulation with a maximum at 50 ms, while the deviations for color remain small. It is likely, however, that this modulation is not a direct but an indirect effect from target duration. Improvements in target identification will only be pronounced when the target was not yet perfectly seen, i.e. at accuracies below 100%. The longer the target was shown, however, the better it was discriminated so that targets shown for 200-250 ms or more were almost always perfectly identified in both conjunction and single-feature tasks. With these durations then, an improvement of target identification in the conjunction task could not be measured. In summary thus, the overall improvement of orientation identification in conjunction tasks, compared to single-feature tasks, was best at short durations and at delays from 500 ms onwards, while there was initially only little effect on the identification of target colors. But with increasing delays color identification in conjunction patterns was increasingly (slightly) deteriorated compared to that in pure-color tests. Since target color was almost perfectly discriminated with presentation times above 100 ms, the increasing deterioration must occur predominantly at short target durations. Contrary to the absent short-term correlation in Figure 5, thus, color identification in conjunction tasks deteriorated during long presentations while orientation discrimination improved and remained increased at long delays. Note however, that color deviations over all test conditions were not significant for all but one observer (see *Statistics* above).

Performance deviations with target durations were seen in all observers (Fig. 7) although with different strength. While in four observers deviations in orientation were predominantly positive up to target durations of 200 ms, one observer (red data curves) showed positive and negative deviations which resulted in a small and non-significant overall effect (see *Statistics* above). In contrast, color deviations were relatively small in most observers, vanished at about 100 ms target duration and were neither predominantly positive nor negative. Only one observer revealed a stronger (and significant) deterioration of color identification in the conjunction task (blue data points).

Altogether, the analysis revealed that the identification of target features was notably improved, or weakly deteriorated, in conjunction compared to single-feature tasks. For orientation, there was a strong improvement at target durations up to 200 ms, with a maximum at 50 ms; for color, there was a small deterioration which has almost vanished at target durations of 100 ms and more. The shorter time course of color deviations is, of course, due to the fact that color was much faster identified than orientation and accuracy was already nearly perfect at target durations from 100 ms onward.

There are two caveats with this finding, however. First, since the idea to test conjunctions was developed during the project, conjunction and single-feature experiments were not strictly run in parallel, i.e. in interleaved sessions,



and continuously grow in amplitude up to the longest delay tested (indicating that target color after long delays was less accurately identified in conjunction than in pure-color tasks). In averages from same presentation times *(b),* deviations in color are only little modulated and deviations in orientation reveal a strong peak with a maximum at 50 ms; for longer presentation times, mean deviations diminish again. Note that this latter decay may reflect the overall performance in the tasks; if targets were correctly identified in both conjunction and single-feature tasks, which was usually the case with target durations of 200-250 ms, no deviations in performance could be measured. Gray curves plot the deviation analysis for orientation in different subsamples of the test series; see text for details.

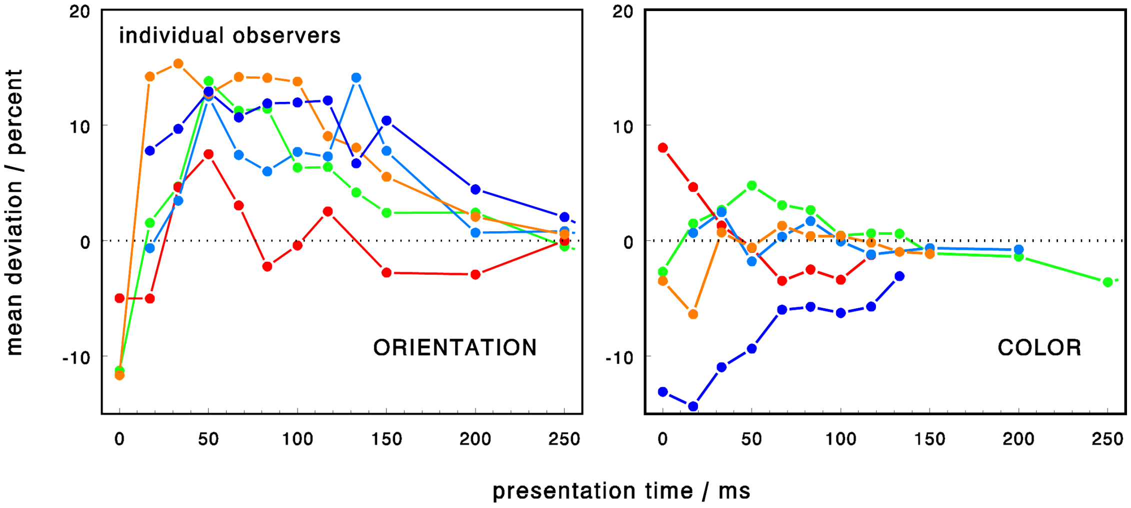


Figure 7. *Deviation analysis as in Figure 6b for individual observers.* Deviations in orientation were mainly positive but varied in amplitude. Deviations in color were smaller and differed between observers.

on all observers. Some observers had already performed several pure orientation tests when the conjunction tests began. Thus, if there had been training effects and subjects had improved their speed in orientation identification during the project, later performance (in the conjunction task) might have been better than the very early performance (in the pure orientation task). Second, when observers were asked to match the luminance of color targets (in the conjunction task) with the luminance of white targets (as used in the pure orientation task of the accompanying study), some of them matched the color targets brighter than the white lines used there. This apparent change of luminance contrast alone could have affected performance and might have let observers identify the orientation of colored lines better (faster) than that of the apparently dimmer white lines in the standard orientation task. The following analyses controlled for these two artifacts.

No learning effects at this stage. To prove the hypothesis that orientation discrimination might have improved over time so that, by chance, the better performance in later conjunction tests was compared with perhaps a much poorer performance in early orientation tests, the series of pure orientation tests in the accompanying study were split in two halves, early and late tests. For each of these halves, performance in orientation identification was analyzed and compared with the orientation identification in the conjunction task, as before. Should identification speed and accuracy have improved in the course of the experiment, deviations should have been much stronger in the first than in the second half of the test series. This was, however, not the case (gray data curves in Fig. 6). Differences between the data sets are rather small, and deviations in orientation identification between conjunction and pure orientation tasks were similarly strong in the early and late test sessions. It should also be mentioned that all observers had carried out other orientation identification tasks before so that a strong improvement during these series of experiments might have been unlikely. Thus, while learning effects were indeed found in early runs of each observer (see, e.g., Nothdurft, 2017a), they were not prominent in the present analysis and certainly not strong enough to explain the observed differences in orientation identification between pure orientation and conjunction tasks.

Potential but non-correlated performance variations with perceived line contrast. There is no easy proof of the second suspicion, that observers might have better performed in conjunction tasks because these lines appeared brighter to them. It is obvious that better visible stimuli should also be better (and faster) identified. But when the apparent pattern contrast of individual observers is compared with the strength of deviations in their

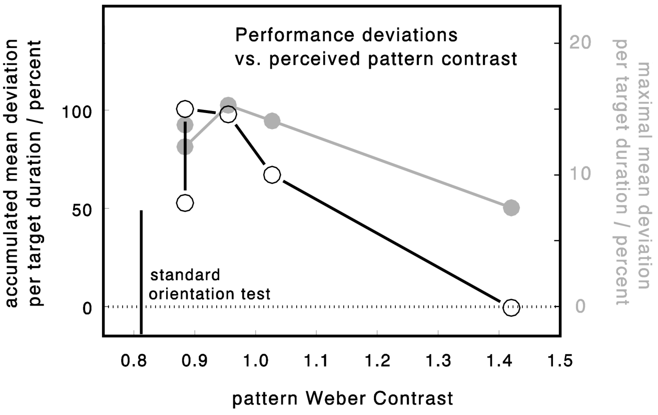


Figure 8. *Better orientation discrimination in the conjunction task was not related to an increased perceived pattern contrast.* When observers were asked to match colored and white targets for luminance, some adjusted the white stimulus brighter than the lines that were used in the single-orientation task. To see if these variations might explain the better orientation discrimination in the conjunction task (with the apparently enhanced pattern contrast), the relevant data are summarized in this graph. The Weber contrast of the standard orientation test (in the pure-orientation task) and the apparent Weber contrast of colored lines (in the conjunction task) as perceived by different observers are related to two measures of performance deviations in orientation, the accumulated difference for presentation times up to 250 ms (i.e., the integral of the individual deviation curves plotted in Figure 7), and the individual deviation maximum in this range. If the perceptually enhanced contrast had accelerated orientation identification in the conjunction task, deviations should *increase* with apparent contrast. The data do not show that.

(individual) performances (Fig. 8), the results are contrary. Across observers, the strength of orientation deviations was seemingly *anti*-correlated with their individual sensations of pattern contrast. Two observers who had matched the luminance of white lines (so that these appeared equiluminant with the color lines in the present experiment) close to that of test lines in the pure orientation experiment revealed the strongest orientation deviations among all observers (yellow and dark blue data curves in Fig. 7). On the other hand, one observer who had perceived the colored lines as particularly bright and had matched equiluminant white lines notably brighter than the standard lines in the pure orientation task (Weber contrast of 1.42 vs. 0.81) had produced particularly small orientation deviations between the conjunction and the pure orientation tasks (red data curve in Fig. 7). This indicates that deviations cannot be explained by an apparently larger pattern contrast of stimuli in the conjunction task. Altogether, the findings thus indicate that orientation (but not color) identification is improved in the conjunction task compared to that in the single feature identification tasks.

DISCUSSION

The study has analyzed the speed of target identification when the target was selected by cued attention from an array of many similar items. It was shown that for targets made of two independent features (color and orientation), these features were split in the analysis and independently processed. Target colors (red vs. green) were faster discriminated than target orientations (left- vs. right-tilted lines). With short presentation times, this has led to partly incorrect target identifications, in which the correct color was associated with the wrong orientation. Such percepts have been described as "illusory" conjunctions (e.g., Treisman & Schmidt, 1982) and are explained in FIT by missing or yet insufficient attentional linking processes. In the present study, however, attention was already cued to these locations, and the comparison with performance variations at slightly shorter or longer presentation times clearly shows that the perceptual failure is not due to a lack of attentional feature integration but to the inability of observers to correctly identify all target features in a too short presentation time. Note that this finding and interpretation already question the need of linking or binding processes in conjunction tasks (Pelli, Palomares, & Majaj, 2004).

The particularly fast processing of color, faster than orientation, is in agreement with other studies (e.g., Cheal & Lyon, 1992; Moutoussis & Zeki, 1997a, b; Zhuang & Papathomas, 2011; Rangelov & Zeki 2014; McIntyre & Arnold, 2018; Hansmann-Roth, Chetverikov, & Kristjánsson, 2019) and closely replicates the findings obtained with pure-feature identification tasks in the accompanying study (Nothdurft, 2020). The observation that the different features of a target are processed independently and with different dynamics seems to be a strong argument against perceptual models in which objects are processed as entities with the different features being linked together (Treisman, 1996, 1999). But this conclusion is not necessary. Binding may also occur between features that are presented in sequence, not simultaneously (Schneegans & Bays, 2017), which should cause serious problems for binding models based on the synchronization of responses (e.g., von der Malsburg, 1981; Eckhorn et al., 1988; Engel, Roelfsema, Fries, Brecht, & Singer, 1997). General problems with asynchronous binding may occur when objects are moving. Color changes, for example, are earlier perceived than changes in the direction of movement (Moutoussis & Zeki, 1997a, b; Viviani & Aymoz, 2001; Holcombe & Cavanagh 2008; Rangelov & Zeki 2014; McIntyre & Arnold, 2018; Nothdurft, 2018). The temporal differences in the encoding of individual features may then become critical for the correct perception of moving colored objects (Moutoussis & Zeki, 1997a, b; Viviani & Aymoz, 2001). But this does not seem to be a lack (or the slowness) of attentional feature integration but rather a lack of strictly synchronous processing and perception.

The independence of feature analysis in the present study was not complete, however, as orientation was identified slightly faster in targets with feature combinations (conjunction targets) than in targets in which only the feature orientation was to be distinguished (pure-feature targets). Possible reasons for this phenomenon shall now be discussed.

No evidence of shared attentional resources

If attention is required to identify and discriminate certain target features, one might have assumed that the simultaneous identification of two features could exceed the capacity limitations of spatial attention so that performance in conjunction tasks should have been slower than performance in a single-feature task. But this was not the case. Color and orientation were not identified more slowly in the combined-feature task but, in fact, even faster than in the single-feature tasks. There is evidence in the literature that orientation and, in particular, color identifications do not require attention (Braun & Julesz, 1998) but this observation was not generally confirmed. In certain tasks, attention is indeed required for (Joseph, Chun, & Nakayama, 1997), or at least associated with the identification of certain target properties (Joseph & Optican, 1996; Nothdurft, 1999). In the special CVS task of the present study, attention was automatically drawn to the target by means of a salient cue. This does not proof that attention was indeed necessary for discrimination but it obviously was necessary for target *selection*. The majority of tested observers had made use of this attentional selection; in only one observer (and only with color) the cued selection did not seem to have started a localized identification process. This observer could identify most target colors even later from the memory (see Nothdurft, 2020, for details).

But how could performance in the identification of one feature then be *improved* by the presumably simultaneous identification of another feature? One obvious possibility would be that improvement was not achieved on both features simultaneously but that identification was improved on only one feature on the cost of identification performance on the other feature. The better orientation identification in conjunction tasks should then have been accompanied be a deterioration of color identification, and *vice versa*. But this was not observed (Fig. 5). The small deterioration in color identification in Figure 6 does not nearly account for the much stronger (simultaneous) improvement in the identification of target orientation. Thus, there is no evidence that attention was split between features and performance had suffered from capacity limitations. But this apparently clear result must be interpreted with some care. Color identification improved very fast with increasing target presentation time and quickly reached 100% accuracy, above which performance variations (and hence deviations between the tasks) could not be measured and compensation effects between color and orientation might have been hidden.

Improved performance in conjunction tasks

If we assume that orientation discrimination was indeed better in conjunction than in pure-orientation tasks, it would be interesting to speculate about possible causes of this difference. The simplest case of perceptual learning can be excluded from the data (Fig. 6), but I will discuss three other possible sources.

Salience. One possibility would be that the salience cue was somehow strengthened in conjunction patterns and attention was attracted faster or with more strength to the target location. It is not obvious how this might have happened with the orientation stimulus (which was, for the feature orientation, the same as in the single-feature task). A feasible alternative might have been color. While the distribution of local color contrast across the pattern (red vs. green neighboring lines) was random and should *per se* not have generated local hot spots of increased salience, the general rules of color attraction might be different and less uniform than those for orientation. For example, color cues are very effective in attracting and guiding attention (Dunai, Castiello, & Rossetti, 2001; Zhuang &

Papathomas, 2011; Kibbe, Kàldy, & Blaser, 2017; Kasten & Navon 2018) and in particular red items may facilitate certain attention effects (Xia, Qi, Shi, Zhang, & Luo, 2018). Since salience attraction in one feature domain also improves identification in other feature domains (Nothdurft, 1993) and since salience effects may add across dimensions (Nothdurft, 2000) it might have been that (cued) red targets in the conjunction task had attracted "more" attention to the target than the overall white and only cued targets in the pure-orientation task.

Orientation signals. Another possibility might be that the orientation signals themselves have differed, on which the observers have based their reactions. The simplest and also plausible assumption is that the apparently slightly increased pattern contrast in the conjunction task might have facilitated orientation discrimination. An argument against this explanation is that the individually perceived pattern contrast does not correlate with the strength of this facilitation, as analyzed above (Fig. 8). An improvement of the underlying neural signal should also be obtained if more orientation sensitive neurons would, perhaps even more strongly, respond to colored than to white lines. This should then have enlarged the target signal and might have helped observers to accelerate their decisions. While there are many color-sensitive neurons in area V1 some of which might indeed have better responded to an appropriate color stimulus than to a white line, the resulting net effect of the population response to oriented lines is to my knowledge not known.

Alertness. Last but not least, it might have been that target identification in conjunction tasks was generally more challenging for observers than target identification in single-feature identification tasks. The need to distinguish four templates, with four different keys, instead of two templates with only two keys, might have increased the general alertness of observers. This might have improved performance in the conjunction task. There is no strong support of this model from the data, however. If alertness was generally increased, we should have noticed an improvement in color identification, too, not only orientation. While this was indeed seen for short target durations in some but not all observers (Fig. 7), deviations were small and tended to average out. This is quite in contrast to the much stronger improvements in orientation seen in all observers with many target durations.

CONCLUSIONS

The study showed that combined features in a target are processed independently and each at its own dynamics. While this itself is a good argument against attentive binding processes, data revealed that false conjunctions do not necessarily indicate a failure of binding but may simply reflect the yet incomplete processing of certain feature components. Furthermore, there was no evidence that feature integration or binding processes might have withdrawn attentional resources from target identification processes. Feature identification in combined targets was not slowed down compared to that in single-feature targets but was, in fact, accelerated.

REFERENCES

Braun, J., & Julesz, B. (1998). Withdrawing attention at little or no cost: Detection and discrimination tasks. *Perception & Psychophysics, 60 (1),* 1-23.

Cheal, M. L., & Lyon, D. R. (1992). Benefits from attention depend on the target type in location-precued discrimination. *Acta Psychologica, 81,* 243-267.

Dugué, L., Xue, A. M., & Carrasco, M. (2017). Distinct perceptual rhythms for feature and conjunction searches. *Journal of Vision, 17(3):22,* 1-15, doi:10.1167/17.3.22.

Dunai, J., Castiello, U., & Rossetti, Y. (2001). Attentional processing of colour and location cues. *Experimental Brain Research, 138,* 520-526. doi: 10.1007/s002210100740.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., and Reitboeck, H. J. (1988). Coherent oscillations: a mechanism for feature linking in the visual cortex? *Biological Cybernetics, 60*, 121–130.

Engel, A. K., Roelfsema, P. R., Fries, P., Brecht, M., and Singer, W. (1997). Role of the temporal domain for response selection and perceptual binding. *Cerebral Cortex, 7*, 571-582.

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

Hansmann-Roth, S., Chetverikov, A., & Kristjánsson, Á. (2019). Representing color and orientation ensembles: Can observers learn multiple feature distributions? *Journal of Vision, 19(9):2,* 1-17, https://doi.org/10.1167/19.9.2.

Holcombe, A. O., & Cavanagh, P. (2008) Independent, synchronous access to color and motion features. *Cognition, 107,* 552-580.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43,* 171-216. doi: 10.1006/cogp.2001.0755.

Joseph, J. S., Chun, M. M., & Nakayama, K. (1997). Attentional requirements in a 'preattentive' feature search task. *Nature, 387,* 805-807.

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

Kasten, R., & Navon, D. (2008). Is location cueing inherently superior to color cueing? Not if color is presented early enough. *Acta Psychologica, 127,* 89-102.

Kibbe, M. M., Kàldy, Z., & Blaser, E. (2017). Rules infants look by: testing the assumption of transitivity in visual salience. *Infancy,* 1-17. International Congress of Infant Studies (ICIS), Wiley Blackwell. doi: 10.1111/infa.12219.

McLeod, P., Driver, J., Dienes, Z., and Crisp, J. (1991). Filtering by movement in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 55-64.

McIntyre, M. E., & Arnold, D. H. (2018). Synchronous and asynchronous perceptual bindings of colour and motion following identical stimulations. *Vision Research, 146-147,* 41-47.

Moutoussis, K., & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proceedings of the Royal Society of London B, 264,* 393-399.

Moutoussis, K., & Zeki, S. (1997b). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proceedings of the Royal Society of London B, 264,* 1407-1414.

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

Nothdurft, H. C. (1993). Saliency effects across dimensions in visual search. *Vision Research, 33,* 839-844.

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

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

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

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

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

Nothdurft, H. C. (2018). Time window(s) of cued visual selection. *VPL-reports, 9,* 1-31. www.vpl-reports.de/9/

Nothdurft, H. C. (2020). Dynamic differences in the cued identification of orientation, color, and facial expressions. *VPL-reports, 10,* 1-27. www.vpl-reports.de/10/

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4(12),* 1136-1169, http://journalofvision.org/4/12/12/. doi: 10.1167/4.12.12

Rangelov, D., & Zeki, S. (2014). Non-binding relationship between visual features. *Frontiers in Human Neuroscience, 8: 749,* 1-11. doi: 10.3389/fnhum.2014.00749

Schneegans, S., & Bays, P. M. (2017) Neural architecture for feature binding in visual working memory. *Journal of Neuroscience, 37(14),* 3913-3925.

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

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

Treisman, A. (1999). Solutions to the binding problem: Progress through controversy and convergence. *Neuron, 24,* 105–110.

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

Treisman, A., and Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance,* 16, 459-478.

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14*, 107-141.

Vigano, G. J., Maloney, R. T., & Clifford, C.W. G. (2015). Transparent surface segregation enables visual feature binding in rapidly alternating displays. *Journal of Vision, 15(9):14,* 1-17, doi:10.1167/15.9.14.

Viviani, P., & Aymoz, C., (2001). Colour, form, and movement are not perceived simultaneously. *Vision Research, 41,* 2909-2918.

von der Malsburg, C. (1981). The correlation theory of brain function. MPI Biophysical Chemistry, Internal Report 81–2. Reprinted in *Models of Neural Networks II* (1994), E. Domany, J. L. van Hemmen, and K. Schulten, eds. (Berlin: Springer).

Wolfe, J. M. (1994). Guided Search 2.0: a revised model of visual search. *Psychonomic Bulletin & Review, 1,* 202-238.

Wolfe, J. M., & Cave, K. R. (1999). The psychophysical evidence for a binding problem in human vision. *Neuron, 24,* 11-17.

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.

Xia, T., Qi, Z., Shi, J., Zhang, M, & Luo, W. (2018). The early facilitative and late contextual specific effect of the color red on attentional processing. Frontiers in Human Neuroscience, 12, 224. doi: 10.3389/fnhum.2018.00224

Zhuang, X., & Papathomas, T. V. (2011). Cue relevance effects in conjunctive visual search: Cueing for location, color, and orientation. *Journal of Vision, 11(7):6,* 1-13, http://www.journalofvision.org/content/11/7/6, doi: 10.1167/11.7.6.

Zohary, E., and Hochstein, S. (1989). How serial is serial processing in vision? *Perception, 18,* 191-200.