

Uppsala University

This is an accepted version of a paper published in *Attention, Perception & Psychophysics*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the published paper:

Poom, L. (2011)

"Motion and color generate coactivation at postgrouping identification stages"
Attention, Perception & Psychophysics, 73(6): 1833-1842

URL: <http://dx.doi.org/10.3758/s13414-011-0132-8>

Access to the published version may require subscription.

The final publication is available at www.springerlink.com.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-158321>

DiVA 

<http://uu.diva-portal.org>

Running head: Motion and color generate coactivation

Motion and color generate coactivation at post-grouping
identification stages

Author:

Leo Poom

Address:

Uppsala University

Department of Psychology

Box 1225

SE-751 41 Uppsala

Sweden.

Email: Leo.Poom@psyk.uu.se.

Phone: +46 18 4712127

Fax: +46 18 4712000

Response times (RT) were measured in a post-grouping visual identification task. Shapes composed of multiple elements were distinguished by color, motion, orientation and spatial frequency alone or in pair wise conjunctions. A race model analysis on these results revealed largest amount of coactivation with color-motion conjunctions. In contrast, RTs for a pre-grouping detection task using the same target shape as in the identification task, distinguished by color, motion, or a conjunction of these features, showed no evidence for coactivation. The results provide psychophysical evidence for coactivation of color and motion signals in cortical regions specialized for grouping and object identification as opposed to separate processing of these features in cortical area V1 believed to limit performance in visual search and pre-grouping detection.

Keywords: Race model, identification, feature integration, response time, grouping

Introduction

Objects can be distinguished and identified visually by contrasts in any of several features. For example the same object can be identified whether it has a different color, brightness or texture than the surrounding scene, and when camouflaged against a textured background it immediately reveals itself when moving and can be identified by relative motion alone. Identification is possible even when constituent parts of an object are spatially separated, requiring grouping, and the parts are distinguished by different features (Poom, 2001; 2002). Although it has been debated to what extent features are segregated in feature specific channels, results providing evidence for segregation of color and motion have been used as a cardinal example to argue in favor for separate feature specific processing. For example, color and motion are to a high degree segregated at first stages of visual processing involved in visual search tasks, where odd targets embedded in distractors has to be detected, whereas other feature conjunctions produce coactivation (Koene & Zhaoping, 2007; Poom, 2009). Here, coactivation refers to summing of activity, or saliency, on a common neural population.

It has been known for long that response times (RT) are faster in redundant conditions when a target is specified by two constituent parts than if a single part is used (Miller, 1982; Raab, 1962; Todd, 1912). Also, in simple visual search tasks, where a target is embedded in distractor elements, redundancy gain typically occurs when two features distinguish a target, for example orientation and spatial frequency (Koene & Zhaoping, 2007; Poom, 2009). It is possible to use experimental psychological methods to investigate whether the output from the decision stage can be accounted for by separate feature specific processing or if it relies on coactivation. In the race model (Miller, 1982; Raab, 1962), formulated to

explain this redundancy gain, different features are processed in separate channels and a response is made as soon as the activity in either channel reaches a specific threshold. Then, on average if the RT distributions overlap, faster response times will occur if decisions are based on activities in two channels than if decisions are based on the activity in only one channel. Coactivation models, on the other hand may produce redundancy gain above the gain predicted by two separate channels, (i.e. Miller, 1982; Mordkoff & Yantis, 1993).

RT analysis and violations of the race model are frequently used as evidence for the influential saliency map models of visual search where feature contrast signals are summed on a master saliency map producing RTs that are faster than predicted by separate activities (Zehetleitner, Krummenacher, & Müller, 2009). The redundancy gain in a visual search task obtained for color-orientation and orientation-motion combinations (Koene & Zhaoping, 2007), and spatial frequency combined with orientation, color and motion respectively Poom (2009) is too large to be explained by statistical facilitation from separate channels and requires coactivation. Neither study, however, found any evidence of coactivation for color-motion combinations. In line with a saliency map summation model where feature contrasts on the same location add activity on the same neural population, coactivation was location specific since it was found only when both features distinguished a single target and not separate targets (Poom, 2009).

Whereas the above mentioned detection tasks relies on pre-grouping stages, in post-grouping task multiple elements are grouped into a larger shape. Field, Hayes and Hess (1993) used a set of oriented Gabor elements that created a contour to investigate post grouping stages and Nothdurft (1993) demonstrated that grouping by common motion and common color are two strong grouping prin-

ciples. In the above mentioned studies grouping principles were investigated using detection tasks primarily driven by bottom-up processes. It is largely unknown, however, how features combine in a post-grouping identification tasks, involving processes believed to occur hierarchically subsequent to processes crucial for detection and visual search tasks, and driven by top-down processes.

In Experiments 1 and 2 RTs were measured for identification of a well known target (an S-shape) composed of multiple elements, each distinguished by a single feature or two features in conjunction. In Experiment 2 a detection task was included. In brief, the results show that, contrary to detection, color-motion conjunctions in the identification task produce redundancy gain too large to be accounted for by separate processing. Further, the amount of coactivation for identification is larger for the color-motion conjunction than any other feature conjunctions tested.

The race model

If decisions are based on activities in separate channels and their time-to-reach-threshold distributions overlap then a statistical redundancy gain is possible. On average then, the time for the winner is less than the time for either channel alone to reach criterion (Miller, 1982; Raab, 1962). The general race model embraces several models with varying degree of channel independence including those where the activities of separate channels are correlated, and those where times to reach criterion is independent between channels, and models where the time to reach criterion for each channel is independent of stimulation levels of the other channel (Hughes & Townsend, 1998).

Suppose that two feature specific channels are sensitive to features f_A and f_B respectively, and $f_{A\&B}$ indicates that both features are present. Then the race model inequality originally formulated by Miller (1982) can be written:

$$P(RT < t | f_{A\&B}) \leq P(RT < t | f_A) + P(RT < t | f_B) \quad (1)$$

The left side of the inequality is the empirically obtained cumulative distribution functions (CDFs) of response times less than t from redundant feature conditions. The right side is the sum of corresponding CDFs from single feature conditions. That is, cumulative distributions for response times at each quantile (that is, points taken at regular intervals from the CDF) when two features are combined never reaches beyond the sum obtained from conditions where each feature is used alone. If the race model inequality is violated, at any quantile, then race models must be rejected in favor of coactivation models. Coactivation, however, cannot be rejected if the race model boundary is not violated. Equivalently, inequality (1) may be expressed as the minimum RT allowed at each percentile which then sets a race model RT boundary, evidence for coactivation occurs when RTs are faster than this boundary.

In neural coactivation models it is assumed that separate features produce activity in a common channel before the decision stage (Miller, 1982). This coactivation may lead to a activity boost that satisfies the criterion faster than a statistical redundancy gain as expressed by inequality (1).

There is a possibility that redundancy gain, measured as RTs, may occur at motor level with faster motor activity using redundant targets (Diedrich & Colo-

nius, 1987) but investigations addressing this issue suggest that this is not the case (Mordcoff, Miller & Roch, 1996).

Experiment 1

Post-grouping identification RTs were measured with shapes distinguished by color, motion, orientation or spatial frequency (single feature conditions), or by two of these features in conjunction (double feature conditions). RTs were measured for identifying an “S” target shape composed of 11 elements embedded in 89 distractor elements in 75 % of trials. In the remaining trials one of 11 non-target shapes were randomly presented, also composed of 11 elements, and the participants were asked to withhold their response (Figure 1A). An example of displays with the S-shaped target distinguished by orientation alone is shown in Figure 1B; spatial frequency alone in Figure 1C; and orientation and spatial frequency in conjunction in Figure 1D. The features, display, stimulus parameters, apparatus, design and procedures are the same as used in the detection study (Poom, 2009).

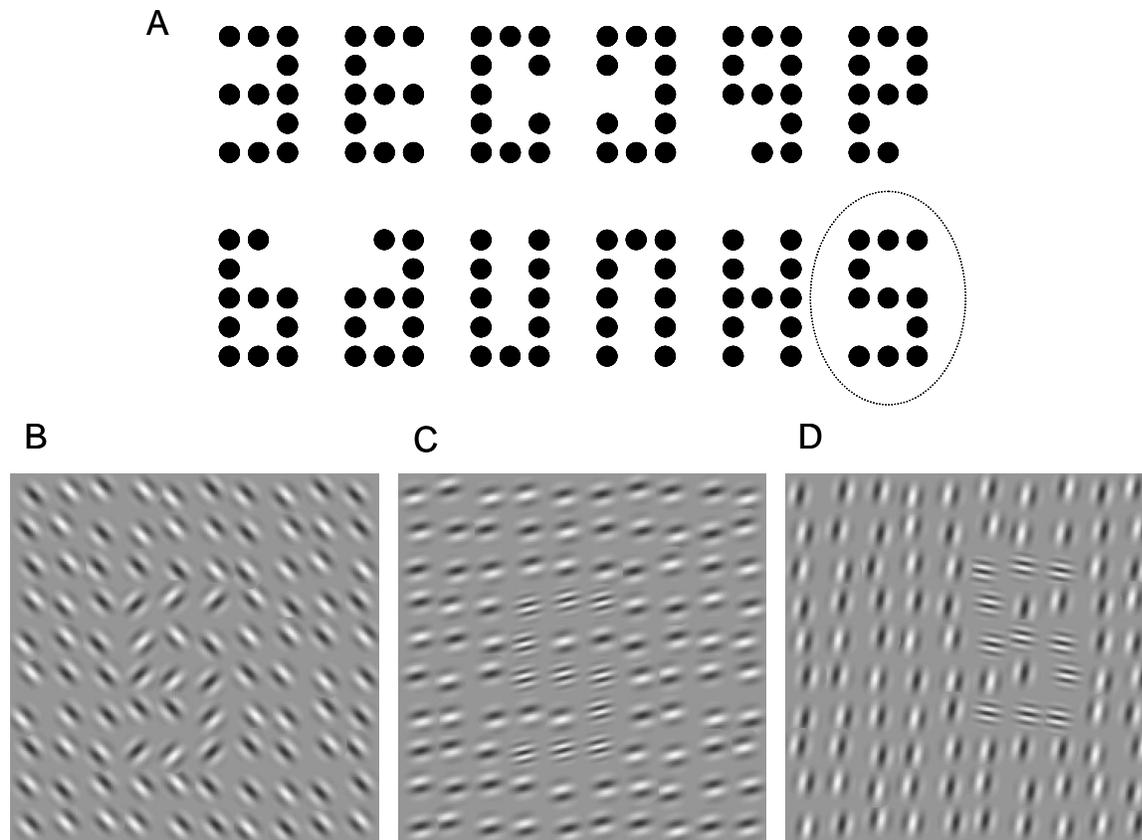


Figure 1. Stimuli. (A) Schematic view of the twelve shapes composed of 11 elements used in the experiment. The S-shaped target is enclosed. In the real display shape elements were distinguished from distractor elements by color, orientation, spatial frequency, and motion either alone or in pair wise conjunctions. Examples of stimulus displays are shown with the target S-shapes distinguished by (B) orientation and (C) spatial frequency in single-feature conditions, and (D) spatial frequency and orientation in conjunction in a double-feature display.

Methods

Participants

Nineteen participants (19-45 years old) with no color vision deficits were recruited to participate in the experiment. All had normal or corrected to normal vision. Informed consent was obtained from all participants after the nature of the

experiment was explained and they were rewarded by cinema tickets for their participation.

Apparatus and Stimuli

Stimuli were displayed on a 17 inch 1280 x 960 resolution CRT monitor with 24 bit color resolution and 100 Hz refresh rate run by a PC equipped with an Intel Pentium 4 processor. The mouse track ball was removed to increase the accuracy of timing (Chambers & Brown, 2003; Segalowitz & Graves, 1990).

The same stimulus parameters were used as for the detection task (Poom, 2009). The stimulus area occupied $13^\circ \times 13^\circ$ of visual angle from a viewpoint 60 cm from the screen and consisted of a regular array of 10 rows and 10 columns of cells. Each cell was occupied by a Gabor element randomly scattered up to 0.25° around the cell midpoint. The brightness of each element, appearing with about 0.8° diameter, was modulated by a Gabor function (a spatial two-dimensional Gaussian blob, with size $\sigma = 0.22^\circ$, multiplied with a sine function) according to the equation $L = L_0 + c \cdot G(x, y, a, f)$, where L_0 is background luminance set by the RGB values, c is luminance contrast, $G(x, y, a, f)$ is the Gabor function, x and y are spatial positions, a is the element orientation, and f is spatial frequency. The luminance contrasts of all shape- and background elements were assigned random values. The Michelson contrast had a rectangular distribution between 0.38 ($I_{\max} = 45$, $I_{\min} = 20$ cd/m²) and 0.72 ($I_{\max} = 60$, $I_{\min} = 10$ cd/m²) against the 30 cd/m² background with L_0 set to 60. This contrast noise minimized any possible luminance based saliency artifacts when color was used to distinguish shape elements. The element spatial frequency was 1.5 cycles/deg corresponding to a spatial period of 0.67° . A common randomly determined orientation was given to the

all elements at each trial but each element was given a random spatial phase.

Shapes were presented at randomly selected locations at the inner 8 x 8 array of cells.

All shapes were composed of 11 elements distinguished from the background elements by motion, orientation, spatial frequency, or color used in isolation or pair wise combined. Except when shapes were distinguished by color all elements were achromatic. These shape elements were located at selected positions of 13 elements arranged in an “8” figure within an area three columns wide and five columns height. A total of 12 shapes were used in the experiment and the target presented on “go” trials was an “S” figure whereas the other shapes presented on “no-go” trials were distractor shapes (Figure 1A). Shapes were defined either by one feature in isolation (single-feature conditions, Figure 1B,C) or by two features (double-feature conditions, Figure 1D).

When color was used to distinguish the shape elements the color saturation, expressed in RGB values, was spatially modulated between background grey and purple-red in phase with the luminance as stated by the following equations $R = L / 3 + S \cdot \text{Abs}[G(x, y, a, f)]$; $G = L / 3 - S \cdot \text{Abs}[G(x, y, a, f)]$; $B = L / 3$, where S is max color saturation (set to 50, and L_0 was set to 60), consequently $R + G + B = L$. As a result the shape elements distinguished by color appeared with successive bright and dark purple-red stripes while the background elements had bright and dark achromatic stripes. Possible differences in apparent color induced luminance contrasts between background and shape elements were efficiently camouflaged by the random contrast variation between all elements.

When motion was used to distinguish the target the Gaussian envelope of each element was stationary but the spatial frequency pattern was phase shifted

90° between the frames, i.e. the spatial frequency pattern was moved $\frac{1}{4}$ period between frames. Motion sequences consisted of four frames. Each frame was presented for 50 ms or five monitor frames. The sequence of four frames was repeated, creating the appearance of unidirectional motion with temporal frequency of 5 cycles/sec while the background elements were static. Although 50 ms motion frames were used, the spatial phase shift of the gabor patterns between frames were small, resulting in the appearance of quite smooth motion. The first frame after initializing a trial was visible for only 20 msec to minimize RT delays.

Shape elements distinguished by spatial frequency had a frequency of 2.6 cycles/deg, corresponding to a spatial period of 0.38° which was 0.56 times the period of the background elements (Figure 1C). Shape elements distinguished by orientation were oriented orthogonally to the distractor elements, i.e. 90° orientation contrast (Figure 1B).

Procedure

Different pairs of features were run in separate sessions. Six sessions were run with pair wise combination of motion, color, spatial frequency, and orientation. Each session included three blocks: two single-feature blocks where each surface feature in isolation distinguished the shape and one double-feature block where both surface features distinguished the shape. The three blocks were presented in random order within each session. On average 25 % of randomly scattered trials were no-go trials with distractor shapes. The blocked session procedure is equivalent to the procedure used in the detection task (Poom, 2009).

Before each session participants performed 20 practice trials with the same features as used in the session. The participants initiated each session by pressing

the space bar. When the S-shaped target appeared they responded by clicking the mouse left button with their right hand index finger (go-trial). If a distractor shape appeared they were asked to withhold any response (no-go trial). After a response or after 1.5 sec if no response was given the stimulus array disappeared and after 2.5-3 sec the next stimulus appeared. This was repeated until 15 successful go-trials were collected in each block. Participants were asked to be as accurate and as fast as possible. After each trial the number of correct trials, within the session so far, was displayed in a rectangular area in the center of the display which turned green after correct trials and red after incorrect trials to provide feedback.

The number of trials varied in each session since the 25% chance of no-go trials was independently assigned at each trial and every wrong response resulted in an extra trial. A total of six sessions, one for each possible feature combination, were run with an average of about 60 trials each, resulting in a total of about 360 trials to complete one experiment for each participant. RTs shorter than 150 ms were considered as anticipation and were excluded from the analysis. Short breaks were allowed between sessions and the experiment took about 40 minutes to complete.

Data analysis

An algorithm developed by Ulrich, Miller and Schrötter (2007), that does not require any large number of observations and allows the results to be aggregated across participants, was used to perform the race model evaluation. The algorithm estimates empirical CDFs for each participant and every stimulus condition and calculates the race model boundary B by adding the corresponding single feature CDFs (right side of equation 2). The CDF for double feature condition D is also

calculated. The D and B RTs are then compared at 10 percentiles (5, 15, 25, ..., 95 %) for each participant. Race model violations occur if the difference $B - D$ is significantly positive which is evaluated by computing separate paired t-test at each percentile across participants and type I errors are avoided by using Bonferroni corrections (Ulrich et al., 2007). The race model inequality is rejected and coactivation must be accepted if there is a significant violation of the inequality at any percentile.

Results

Figure 2 shows the total mean RTs for each feature combination and corresponding single features. Overall RT for identification in this task is about 500 ms.

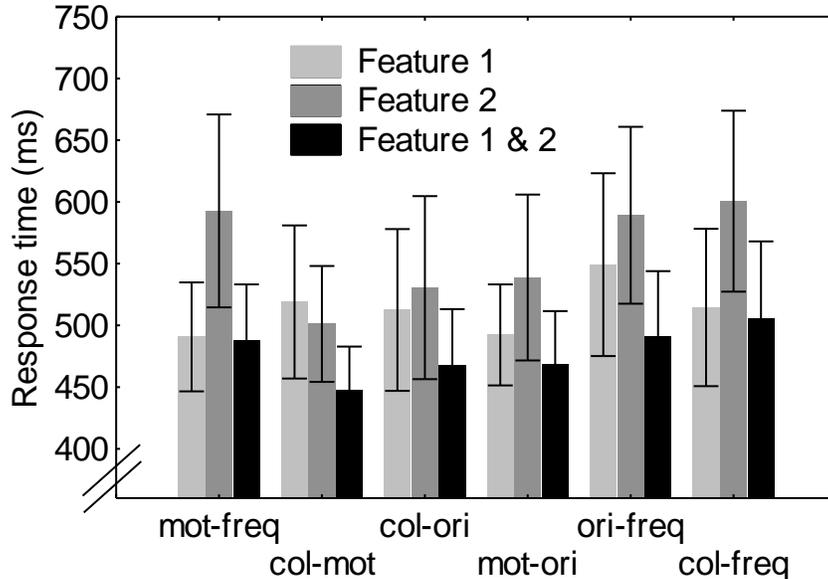


Figure 2. Average RTs across participants are shown for single and double features for each feature combination. The light grey and dark grey bars show RTs for feature 1 and feature 2 respectively in the same order as indicated on the abscissa. The whiskers show 95 % confidence intervals.

It can be noted from inequality (1) that race model violations during coactivation are bound to occur at lower cumulative probabilities since the left side of the inequality asymptotes at 1 and the right side asymptotes at 2. This is clearly revealed in Figure 3 where the double feature RT at each percentile is subtracted from the race model RT boundary (Poom, 2009). Race model violations are clearly indicated by positive values and significant violations are marked with crosses.

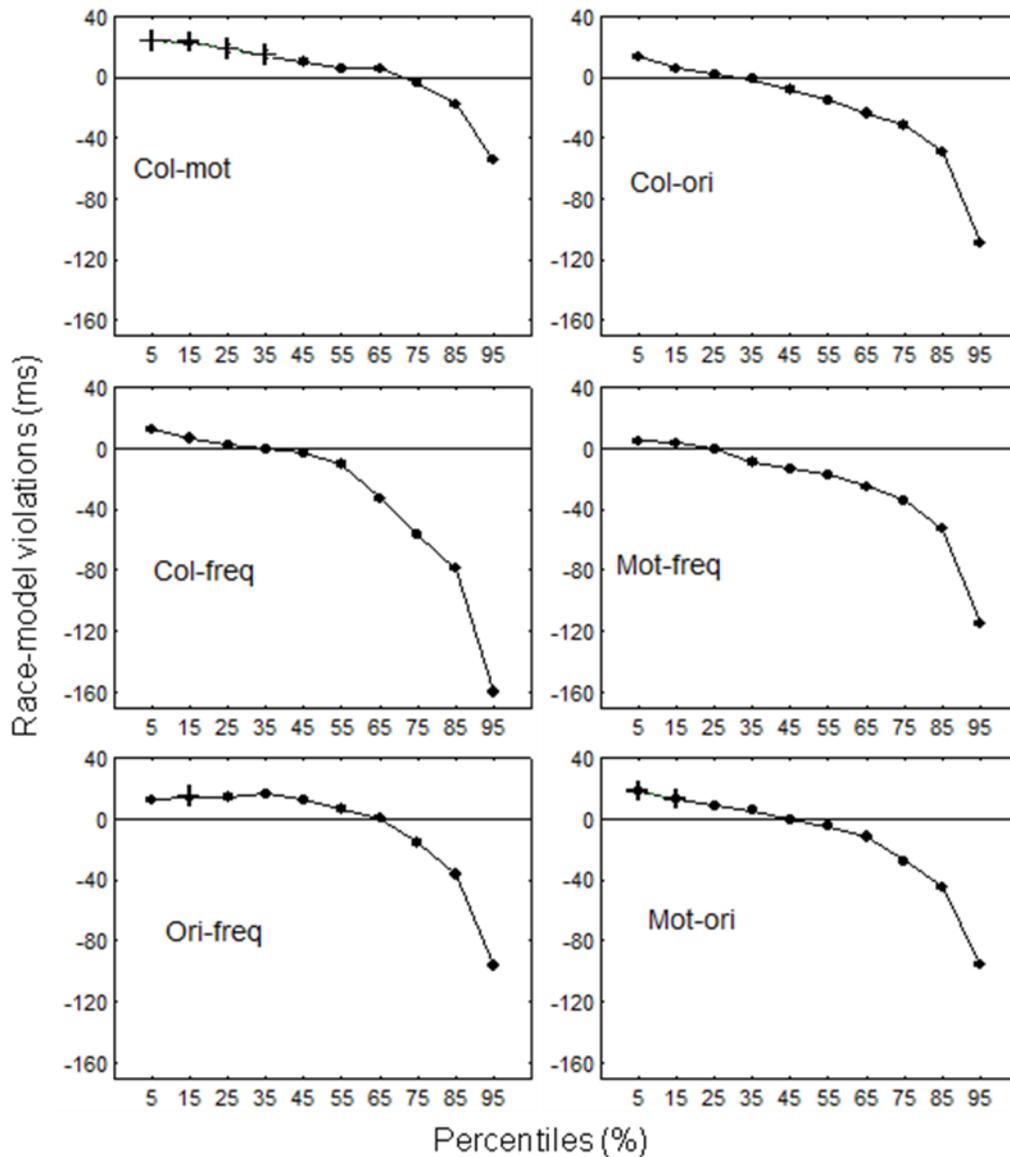


Figure 3. RT deviations from the race model boundary (line at zero) from Experiment 1, separately presented at each percentile. The differences between the boundary and RTs from the double feature conditions are displayed with filled discs and significant violations (positive signs) are displayed with crosses.

Redundancy gain with largest race model violations occurred for color-motion conjunctions providing strong evidence for coactivation. These violations, evaluated separately at 10 percentiles across 19 participant, were significant at

four percentiles between the fifth and 35th percentiles: $t_s(18) > 2.5$, $p_s < 0.05$.

Motion-orientation produced significant violations at the fifth and 15th percentile, $t_s(18) > 2.3$, $p_s < 0.05$ and orientation-spatial frequency produced a significant violation at the 15th percentile only, $t(18) = 2.1$, $p < 0.05$. The color-orientation, color-spatial frequency, and motion-spatial frequency conjunctions produced no statistically significant violations at any percentile although at lowest percentiles the race boundary appeared slightly trespassed.

FA rates varied between 19 % for single feature conditions for color defined shapes, and 22 % for orientation defined shapes. Corresponding values for double feature conditions was 20 % for motion-orientation conjunctions and 22 % for orientation-color conjunctions. For all conditions the number of misses, failure to press the key on go-trials, was below 1 % with an average of 0.5 %. The number of anticipations ($RT < 150$ ms) was very low.

Discussion

The most interesting result from Experiment 1 is that race model violation occurred with the motion-color conjunction which is contrary to results obtained in a similar detection task (Poom, 2009). Decisions based on activities in separate channels can be excluded and coactivation is demonstrated by the race model violations. Failure to find such violations, however, can be accounted for by statistical facilitation from separate activities.

Besides color-motion conjunctions, the only feature combinations that gave significant violations of the race model were orientation-motion and orientation-spatial frequency. The detection of motion and orientation combined could result from differential velocity signals due to the orientation difference of the

Gabor elements (von Mühlenen & Müller, 2001). Mühlenen and Müller showed in their experiments that the detection of a moving target line, embedded in moving and stationary distractors of various orientations, was facilitated due to differential motion signals. The orientation difference of the target line changes the motion signal relative to the other elements.

Here, all background elements were always stationary so the motion difference was the same in both motion alone condition and in orientation-motion conjunction, although the orientation was the same in the first condition and orthogonal in the second condition. Maybe, the orientation contrast produces greater relative motion signal so the orientation difference may have caused a motion signal that is greater than when orientation is the same among all elements. So, the faster RTs could have been caused by a change in the motion signal itself, due to orientation difference. Similarly, the faster RT obtained with combination of orientation and spatial frequency may result from that the orientation signal is stronger with higher spatial frequencies, and not resulting from coactivation of different features. I thank von Mühlenen for suggesting that in some cases the change in one feature may change the saliency of other features.

In sessions with color-frequency and motion-frequency, the single feature RTs differ with about 70 - 100 msec (Figure 2). This may reduce redundancy gain by statistical facilitation as described by the race model so the low redundancy gain in these cases should be interpreted with care.

The amount of violation can be indexed by the summed violations across percentiles where violation occurred (Colonius & Diederich, 2006). In Fig. 4 the amount of race model violations for each feature combination are compared for the present identification data from Experiment 1 and the detection data obtained

from the visual search task (Poom, 2009). Positive violations are summed across percentiles for each participant and then averaged across participants separately for each feature combination.

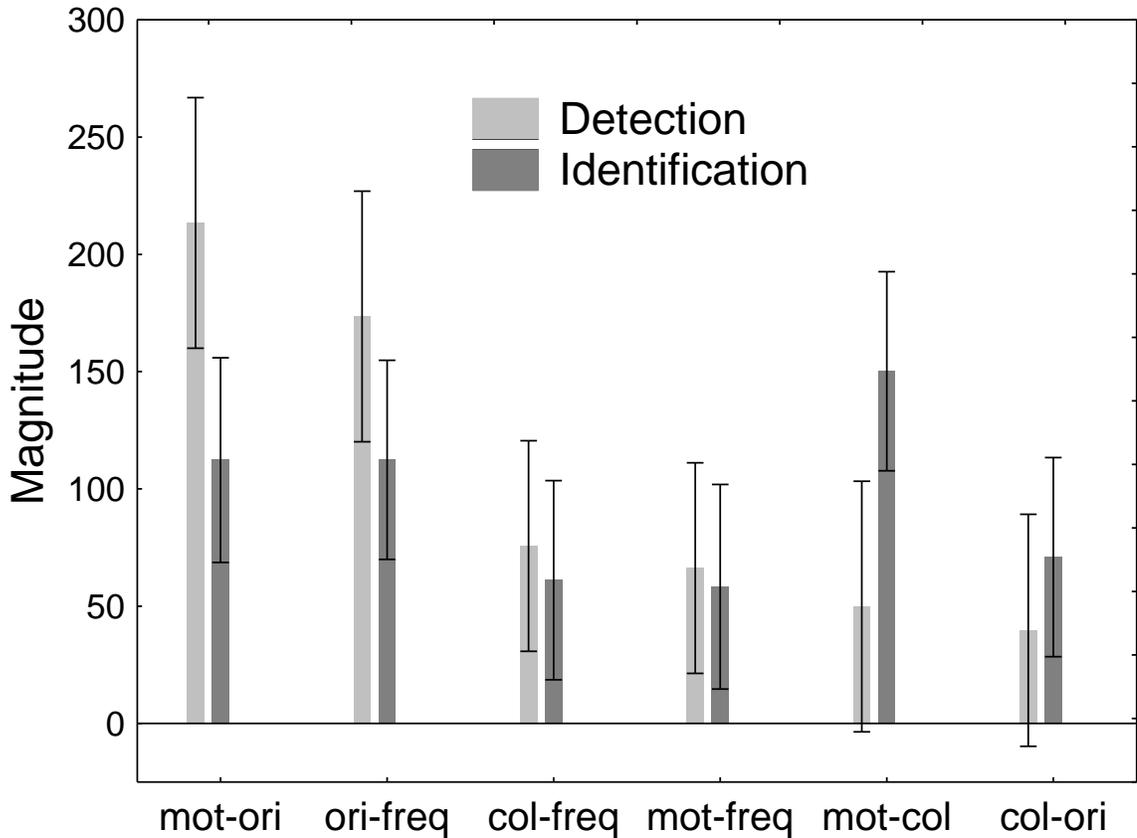


Figure 4. Numerical index of the amount of race model violation as described in the text for the different feature combinations in the identification task from Experiment 1 and the detection task from Poom (2009). Zero magnitude indicates no violation. The whiskers show 95 % confidence intervals.

Inspection of Figure 4 shows that for the motion-color combinations larger violations occurred in the identification task of Experiment 1 than the detection task, whereas the opposite pattern was obtained for the motion-orientation combinations as revealed by the non-overlapping 95 % confidence intervals. Orienta-

tion-frequency also shows quite large opposite pattern although confidence intervals overlap.

Double feature conditions may have led to increased failure to inhibit responses until the shape was correctly identified which would be revealed by higher false alarm (FA) rates in double-feature than in single-feature conditions. FA rates, however, were similar in both conditions. Since no differences in FA rates were obtained for single and double feature conditions, the redundancy gain, and race model violations obtained with double features cannot be blamed on selective failures to inhibit responses in this condition.

There is a possibility that the different pattern of results might result from the different number of target elements. In the detection task only one odd element had to be detected whereas in the identification task the target (and distracters) consisted of 11 elements. Maybe the limited color and motion signal in the detection task fail to coactivate whereas the stronger signal conveyed by 11 elements do coactivate.

Experiment 2

In Experiment 2 the detection task, using color-motion from Poom (2009) was replicated with targets composed of multiple elements. The same target was used in the detection and identification task, an S-shaped target as used in Experiment 1. Further, Experiment 2 tested the identification task and the detection task with a single group of observers and to minimize the false alarm rate, which was quite high in Experiment 1. Observers were encouraged to withhold responses in no-go trials to avoid false alarms, and it was emphasized that false alarms added extra trials.

Participants

Fourteen participants (20-47 years old) with no color vision deficits were recruited to participate in the experiment. All had normal or corrected to normal vision. Informed consent was obtained from all participants after the nature of the experiment was explained and they were rewarded with cinema tickets for their participation.

Stimuli

Only color, motion and color-motion conjunctions were used here, otherwise the stimulus parameters were the same as described in Experiment 1. In the detection task the S-shaped target was presented on go-trials and background noise elements alone were presented in no-go trials. The identification task was the same as in Experiment 1.

Procedure

The target occurred at 75 % of trials randomly interleaved with no-go trials. The detection and identification task were carried out in separate sessions, randomized between observers. Observers were encouraged to withhold responses in no-go trials to avoid false alarms, and it was emphasized that false alarms added extra trials.

Results

Figure 5 shows the mean RTs for single and double feature conditions in the detection and identification task. RTs from single feature conditions overlap in both

the detection and identification task and fastest RTs seems to be obtained when the target was specified by both color and motion. Figure 6 shows that again, the identification task, with motion-color conjunctions lead to race model violations requiring coactivation of color and motion signals. In the detection task, however, motion-color conjunctions failed to provide any evidence for coactivation, which replicate results obtained with a single element target (Koene & Zhaoping, 2007; Poom, 2009). The proportion of false alarm rate was reduced compared to Experiment 1, in the detection task it was about 3 % and in the identification task it was 7 %.

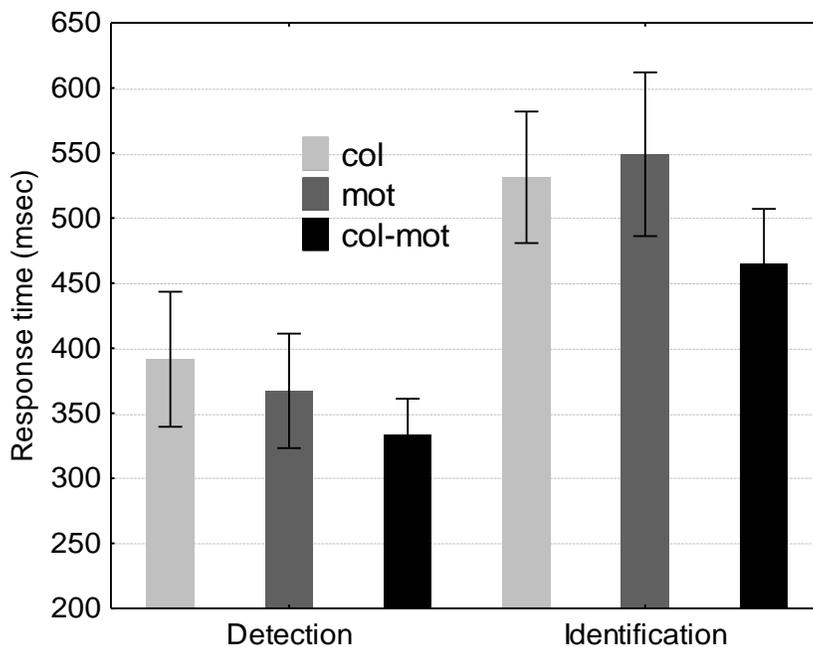


Figure 5. Average RTs across participants are shown for color alone (col) and motion alone (mot) and color-motion used in combination (col-mot) in the detection and identification task of Experiment 2. The whiskers show 95 % confidence intervals.

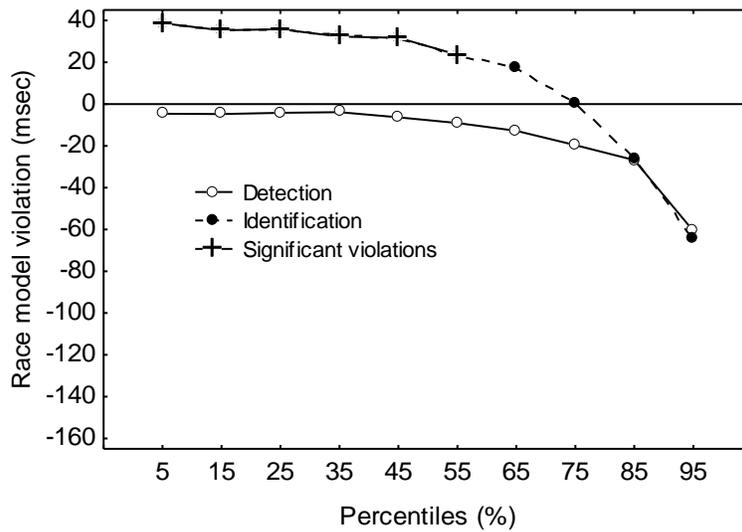


Figure 6. The results from Experiment 2, using color and motion as shape distinguishing features. RT deviations from the race model boundary (line at zero) separately presented at each percentile demonstrates clear evidence for coactivation in the identification task (positive race model violations). The same features and shape failed to provide any evidence for coactivation in the detection task.

Discussion

As in Experiment 1 coactivation was found in the identification task for color-motion conjunction in Experiment 2. The results from the detection task show no evidence for color-motion coactivation even though the target was composed of the same number of elements and had the same shape as the target used in the identification task, replicating previous results obtained with single target elements (Koene & Zhaoping, 2007; Poom, 2009). In the identification task the target shape had to be identified and distinguished from distractor shapes embed-

ded in noise elements, whereas in the detection task the target just had to be detected and distinguished from trials containing only noise elements.

The FA rates from Experiment 1 were severely reduced from about 20 % to 7 % in the identification task, and from 13 % in the identification task with a single target element (Poom, 2009) to about 3 % here using a target composed of multiple elements. This provides additional support that the results in the color-motion condition in Experiment 1 were not disturbed by high guessing rates.

General discussion

The segregation of color and motion is traditionally used as a cardinal example of feature specific processing in separate pathways as exemplified by visual search tasks (Koene & Zhaoping, 2007; Poom, 2009). In the identification task of Experiment 1 the greatest race model violation across features combinations occurred for color-motion conjunctions, revealing clear evidence for coactivation. In Experiment 2, using the same target in detection and identification task, color-motion conjunctions produced RTs in the detection task accounted for by the race model, whereas the identification task again produced race model violations. It should be noted that the luminance contrast for each element was assigned a random value which minimized any possible luminance based saliency artifacts when color was used to distinguish target elements.

Intuitively, one has to do two things to accomplish correct identification, first detect and locate the target elements among the distractor elements, and thereafter group the target elements to be able to identify the object (although the reverse is possible where some of the target elements may be detected after the object has been identified). It is possible that detection, grouping and identification

are sequential processes within the same pathway, or that separate routes of processing handles these tasks. In the first case some kind of dependency between results from detection and identification is expected. In the second case coactivation is independent between these tasks. Color-motion conjunctions seem to coactivate only when the task is identification at a post-grouping stage, not when the task is detection at a pre-grouping stage, whereas other feature combinations indicate opposite patterns. The opposite results obtained here indicate a dependency between processes. If features are separately processed at the pre-grouping stages then there is still a possibility of coactivation at the post-grouping stages. If features coactivate at the pre-grouping stage then no further gain may be possible at later stages due to saturation. Although admittedly loose, this may be an outline for a model explaining the opposite pattern of results in these two tasks.

So, the psychophysical evidence for separate processing of color and motion in detection tasks as opposed to coactivation for corresponding identification tasks may be due to hierarchically segregated processing. In this view detection is determined by activity in V1 where color and motion are separately processed but other features activate common neural populations (Koene & Zhaoping, 2007; Poom, 2009). In contrast, in area V2 there are cells tuned to all feature conjunctions inclusive color-motion conjunctions (Gegenfurtner, Kiper & Fenstemaker, 1996), which argues against V2 as locus selectively engaged in the visual search detection task. Identification, on the other hand, is crucially dependent on object selective neurons in lateral occipital cortex (LOC) and inferior temporal (IT) cortex (Gulyas, Heywood, Popplewell, Roland & Cowey, 1994; Sary, Vogels & Orban, 1993; Self & Zeki, 2004). One may conceptualize these brain areas as maps

of frequently encountered objects, as opposed to the location specific saliency maps in cortical area V1.

Gulyas et al. (1994) showed that color and motion specified objects activate different anatomical regions in the brain suggesting segregation whereas Self and Zeki (2004) found that motion and color used together to define objects activated LOC more than if either feature was used alone and found cross adaptation effects suggesting integration on common neural populations. Whether convergence occurs at single neuron level as shown by single unit measurements in macaque IT cortex (Sary et al., 1993) remains unclear. These areas, however, are known to receive major contributions from both the color specialized parvocellular and the motion specialized magnocellular pathways (Merigan & Maunsell, 1993). In addition, object selective neurons that respond whether the object is defined by motion, luminance or texture size have been found in temporal cortex (Sary et al., 1993). In identification tasks redundancy gain violating the race model can be modeled by hypothetical neural mechanisms that are analogous to the saliency map summation model, but here the saliency map represents objects rather than locations of salient features in the visual field. Accordingly, this object space probably resides in LOC and/or IT.

In the widely recognized saliency map models, pop-out of odd targets occur where bottom up driven activity in corresponding loci at the saliency map reaches a specific threshold (Cave & Wolfe, 1990; Koch & Ullman, 1985; Müller, Heller & Ziegler, 1995; Wolfe, 1994; 2001). FA rates in identification are higher and RTs are longer than in the detection task, probably due to more top down influences in higher processing stages. Accordingly, Nothdurft (2006) found longer RTs when the task was to identify the feature that made a target element pop out,

than just detecting an odd element. This does not mean that detection tasks are free from top-down influences. Krummenacher, Müller and Heller (2001) found that, in detection tasks, when feature weights need to be shifted under variable feature conditions, race model violations are less reliably obtained than under constant feature conditions. Their explanation was that limited attentional weights are better balanced between features in redundant conditions, permitting their features to interact.

In pre-grouping detection tasks it is known that RTs are faster when target defining features repeats rather than changes between trials (Found & Müller, 1996). This is in line with Wolfe's (2001) Guided search model, where both pre-attentive guidance from saliency driven bottom-up, and top-down observer driven guidance that directs attention toward target features, is used to direct attention in visual search. When the target defining feature is known both processes act in concert to speed up detection. Müller Heller and Ziegler (1995) proposed a dimension-weighting account, in which master map units compute the weighted sum of dimension-specific saliency signals. Greater weight is assigned to only one feature at a time on feature specific maps and persists to the next trial. If the target dimension is known in advance, signals from that dimension are amplified. A time consuming weight shift occurs when the feature changes between trials. Analyses of ERPs obtained during such weight shifts demonstrate the involvement of relatively early perceptual processes (Gramann et al, 2007). Although conditions for coactivation were optimized by running blocks with constant feature conditions, no race model violations were found here in the pre-grouping task for color-motion in Experiment 2, or in the previous study (Poom, 2009). Whether similar

costs, caused by weight shifts between features, occur in post-grouping stages as in pre-grouping stages is an interesting question for future research.

The pattern of results obtained by comparing RT redundancy gain in detection and identification may be a general effect of task demand. Kastner, Busse and Treue (2006) measured RTs for identification of a particular color, direction of motion, or both combined in redundant conditions. They found redundancy gain that violated the race model boundary and that the effect occurred although the target color and motion occurred at separate spatial locations, as opposed to an odd target detection task (Poom, 2009). On the other hand, detection of coherent motion is facilitated by texture and disparity signals but not by color (Poom & Börjesson, 2006). This may imply that detection tasks generally, not only those involving visual search, are mediated by separate color and motion processes, and that identification generally, not only shape or object identification, are mediated by processes where color and motion produce coactivation.

In summary, comparing the results for identification and detection suggests that color and motion signals are kept separate until identification processes take place and signals converge at the shape map and therefore produce more coactivation than other feature conjunctions that has already coactivated at lower stages of processing. Thus, in visual search the decision stage relies on signals on separate color and motion specific saliency maps. These separate activities therefore lead to statistical facilitation for color-motion conjunctions in visual search tasks as described by the race model (Koene & Zhaoping, 2007; Poom, 2009). In identification tasks, involving temporal cortex where motion and color signals are finally integrated (Self & Zeki, 2004), the decision stage rely on coactivation of

common color-motion sensitive shape selective neurons producing greater amount of race model violations.

Acknowledgements

This research was supported by the Swedish Research Council and the Magnus Bergwall foundation.

References

- Cave, K. R., & Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology*, **22**, 225-271.
- Chambers, C.D., & Brown, M. (2003). Timing accuracy under Microsoft Windows revealed through external chronometry. *Behavior Research Methods, Instruments, & Computers*, **35**, 96-108.
- Colonus, H., & Diederich, A. (2006). The race model inequality: interpreting a geometric measure of the amount of violation. *Psychological Review*, **113**, 148-54.
- Diederich, A., & Colonus, H. (1987). Intersensory facilitation in the motor component? A reaction time analysis. *Psychological Research*, **49**, 23-29.
- Field, D. J., Hayes, A. & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision Research*, **33**, 173-193.
- Found, A. P., & Müller, H. J., (1996). Searching for feature targets on more than one dimension: investigating the dimension weighting account. *Perception & Psychophysics*, **58**, 88-101.
- Gegenfurtner, K. R., Kiper, D. C. & Fenstemaker, S. B. (1996). Processing of color, form, and motion in macaque area V2. *Visual Neuroscience*, **13**, 161-172.
- Gramann, K., Toellner, T., Krummenacher, J., Eimer, M., & Müller, H. J. (2007). Brain electrical correlates of dimensional weighting: An ERP study. *Psychophysiology*, **44**, 277-292.
- Gulyas, B., Heywood, C. A., Popplewell, D. A., Roland, P. E., & Cowey, A. (1994). Visual form discrimination from color or motion cues: Functional-

- anatomy by positron emission tomography. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 9965-9969.
- Hughes, H. C., & Townsend, J. T. (1998). Varieties of binocular interaction in human vision. *Psychological Science*, 9, 53-60.
- Kastner, S., Busse, L., & Treue, S. (2006). Feature-based attentional integration of color and visual motion, *Journal of Vision*, 6, 269-284.
- Koch C, & Ullman S, (1985) Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4, 219-227.
- Koene, A. R., & Zhaoping, L. (2007). Feature-specific interactions in saliency from combined feature contrasts: Evidence for a bottom-up saliency map in V1. *Journal of Vision*, 7, 1-14.
- Krummenacher, J., Müller, H. J., & Heller, D. (2001) Visual search for dimensionally redundant pop-out targets: Evidence for parallel-coactive processing of dimensions. *Perception & Psychophysics*, 63, 901-917.
- Merigan, W. H., & Maunsell J. H. R. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience*, 16, 369-402.
- Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cognitive Psychology*, 14, 247-279.
- Mordkoff, J. T, Miller, J., & Roch, A. C. (1996). Absence of coactivation in the motor component: evidence from psychophysiological measures of target detection. *Journal of Experimental Psychology: Human Perception & Performance*, 22, 25-41.
- Mordkoff, J. T, & Yantis, S. (1993). Dividing attention between color and shape: Evidence of coactivation. *Perception & Psychophysics*, 53, 357-366.
- Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature

- targets within and across feature dimensions. *Perception & Psychophysics*, 57, 1-17.
- Nothdurft, H. C. (1993). The role of features in preattentive vision: Comparison of orientation, motion and color cues. *Vision Research*, 33, 1937-1958.
- Nothdurft, H. C. (2006). Saliency and target selection in visual search. *Visual Cognition*, 14, 514-542
- Poom, L. (2001). Visual inter-attribute contour completion. *Perception*, 30, 855-865.
- Poom, L. (2002). Visual binding of luminance, motion and disparity edges. *Vision Research*, 42, 2577-2591.
- Poom, L. (2009). Integration of color, motion, orientation and spatial frequency in visual search. *Perception*, 38, 708-718.
- Poom, L., & Börjesson, E. (2006). Color, polarity, disparity, and texture contributions to the detection of global motion. *Perception*, 34, 1193-1203.
- Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, 24, 574-590.
- Sary, G., Vogels, R. & Orban, G. A. (1993). Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science*, 260, 995-997.
- Segalowitz, S. J., & Graves, R.E. (1990). Suitability of the IBM XT, AT, and PS/2 keyboard, mouse, and game port as response devices in reaction time paradigms. *Behavior Research Methods, Instruments, & Computers*, 22, 283-289.
- Self, M. W., & Zeki, S. (2004). The integration of color and motion by the human visual brain. *Cerebral Cortex*, 15, 1270-1279.
- Todd, J. W. (1912). Reaction to multiple stimuli. *Archives of Psychology*, 25, 1-65.

- Ulrich, R, Miller, J., & Schrötter, H. (2007). Testing the race model inequality: An algorithm and computer programs. *Behavior Research Methods*, 39, 291-302.
- von Mühlelen, A., & Müller, H. J. (2001). Visual search for motion-form conjunctions: is form discriminated within the motion system? *Journal of Experimental Psychology: Human Perception and Performance*, 27, 707-718.
- Wolfe, J. M. (1994) Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, 1, 202-238.
- Wolfe, J. M. (2001) Guided search 4.0: A guided search model that does not require memory for rejected distractors. *Journal of Vision*, 10, 349.
- Zehetleitner, M., Krummenacher, J., & Müller, H. J. (2009). The detection of feature singletons defined in two dimensions is based on saliency summation, rather than on serial exhaustive or interactive race architectures. *Attention, Perception & Psychophysics*, 71, 1739-1759.