Salient stimuli capture attention and action

Dirk Kerzel & Josef Schönhammer

🖉 Springer

Springer

Salient stimuli capture attention and action

Dirk Kerzel · Josef Schönhammer

Abstract Reaction times in a visual search task increase when an irrelevant but salient stimulus is presented. Recently, the hypothesis that the increase in reaction times was due to attentional capture by the salient distractor has been disputed. We devised a task in which a search display was shown *after* observers had initiated a reaching movement toward a touch screen. In a display of vertical bars, observers had to touch the oblique target while ignoring a salient color singleton. Because the hand was moving when the display appeared, reach trajectories revealed the current selection for action. We observed that salient but irrelevant stimuli changed the reach trajectory at the same time as the target was selected, about 270 ms after movement onset. The change in direction was corrected after another 160 ms. In a second experiment, we compared manual selection of color and orientation targets and observed that selection occurred earlier for color than for orientation targets. Salient stimuli support faster selection than do less salient stimuli. Under the assumption that attentional selection for action and perception are based on a common mechanism, our results suggest that attention is indeed captured by salient stimuli.

Keywords Reaching · Attentional capture · Saliency · Action

Introduction

Selection of visual information is achieved by attentional prioritization of visual stimuli and may be controlled in different ways. Attentional selection is said to be top-down when it reflects expectations and goals of the observer, and it is said to be bottom-up if it reflects the saliency of the stimuli (reviewed by Theeuwes, 2010). We will briefly present the contingent capture and the additional singleton paradigms, which are believed to provide evidence for top-down and bottom-up control, respectively. However, it should be noted that the two paradigms confound bottom-up and top-down factors to some degree and also involve mechanisms, such as intertrial priming, that defy the theoretical dichotomy (Awh, Belopolsky, & Theeuwes, 2012).

The contingent attentional capture paradigm consists of a target display that is preceded by a cue display (e.g., Folk, Remington, & Johnston, 1992). In separate blocks of trials, observers search the target display for a colored item among white distractors (i.e., a color target), or they look for a single white item (i.e., an onset target). Color targets and onset targets are shown in separate blocks, resulting in an attentional set for color or onset, respectively. Over blocks of trials, the cue display contains either a color cue or an onset cue. It was observed that reaction times (RTs) in a speeded discrimination task are shorter when the cue is shown at the same location as the target, but only when the cue characteristics match the target characteristics (i.e., onset cue/onset target or color cue/color target). In sum, only cues that match the current attentional set capture attention, which is strong evidence for top-down control. However, subsequent studies showed that the repetition of target features contributes to the pattern of results (e.g., Ansorge & Horstmann, 2007), which does not support the notion of top-down control, because trial history is not easily classified as bottom-up or top-down (Awh et al., 2012).

In the additional singleton paradigm developed by Theeuwes (1991), observers search for a shape singleton in a circular array of items and perform a discrimination task. In one variant of the paradigm, the target and distractors change roles randomly from trial to trial (Theeuwes, 1991). That is, a

D. Kerzel (✉) · J. Schönhammer
Faculté de Psychologie et des Sciences de l'Éducation, Université de Genève, 40 Boulevard du Pont d'Arve, 1205 Genève, Switzerland
e-mail: dirk.kerzel@unige.ch

diamond target among circle distractors may be followed by a circle target among diamond distractors, and vice versa. On half of the trials, a salient color singleton is shown at a nontarget location, which increases RTs. Presumably, attention is attracted to the salient color distractor before moving to the less salient shape singleton. However, the conclusion that attention was involuntarily attracted by the salient element and the notion of bottom-up control were subsequently challenged.

Bacon and Egeth (1994) introduced a slight change to the additional singleton paradigm by adding another shape to the array, whereby the target lost its status as shape singleton. After adding the shape, attentional capture was abolished. Bacon and Egeth concluded that attentional capture occurred only when observers were in singleton detection mode, which may induce selection of singletons along the wrong dimension (i.e., color instead of shape). In contrast, when a particular feature was looked for, no capture occurred. However, Theeuwes (1992) found that attentional capture persisted even when the target feature was fixed in a block of trials, instead of changing randomly from trial to trial. In this case, observers did not have to detect the singleton shape but could look for a particular feature in a block of trials (i.e., a feature search).

Furthermore, Folk and Remington (1998) accounted for increases in RTs with an additional singleton by nonspatial filtering costs. According to Treisman, Kahneman, and Burkell (1983), "any object on which attention could, but should not, be focused must be excluded at a cost, in time or in accuracy" (p. 530). Filtering costs arise even for objects that are highly dissimilar from the target and are, therefore, not attended (Kahneman, Treisman, & Burkell, 1983; Treisman et al., 1983). Folk and Remington showed that nonmatching cues in the contingent capture paradigm increased RTs, as compared with a condition with neutral cues. At the same time, RTs with nonmatching cues were not different at cued and uncued locations, suggesting that attention had not been captured. These results confirm the idea that it takes time to exclude nonmatching cues even if they do not attract attention.

Importantly, nonspatial filtering costs predict that the distance between target and distractor should not affect RTs, whereas if attention was captured by the salient distractor, the cost should be larger for distractors that are further away from the target, because attention has to travel a longer distance. Becker (2007) concluded that increases in RTs caused by distractors were mostly consistent with nonspatial filtering costs. Further evidence against attentional capture by salient distractors comes from studies investigating the role of practice. In these studies, two groups of participants were shown different displays in the training phase but the same displays in the subsequent test phase. After training with displays inducing feature search mode, a

salient distractor in the test phase did not affect performance, whereas it did after a training phase inducing singleton detection mode (Leber & Egeth, 2006; Zehetleitner, Goschy, & Müller, 2012). Even more surprisingly, the resistance to distraction of participants who performed in feature search mode disappeared when a distractor was presented in the test phase, which they had not encountered during training (Zehetleitner et al., 2012), pointing to the important influence of training and prior exposure in the additional singleton paradigm.

Furthermore, electrophysiological measures of attentional deployment, such as the N2pc, have been used to investigate the control of attention. The N2pc is a negativity contralateral to the attended stimulus that occurs 200–300 ms after stimulus onset (Eimer, 1996; Luck & Hillyard, 1994). Importantly, the N2pc is a spatial measure of the focusing of visuospatial attention because the negativity shows which hemifield was attended. In contrast, differences in RT could be due to shifts of spatial attention or to other processes, such as spatial filtering. Hickey, McDonald, and Theeuwes (2006) reported an N2pc in response to an irrelevant color singleton, which supports the notion of attentional capture by salient stimuli. In their experiment, target and distractor shapes changed roles randomly from trial to trial. In subsequent research, the target shape was fixed, allowing for feature search, which abolished the N2pc to the irrelevant element (Burra & Kerzel, 2013; Schubö, Schröger, Meinecke, & Müller, 2007; Töllner, Müller, & Zehetleitner, 2011; Wykowska & Schubö, 2010, 2011). The absence of an N2pc to irrelevant distractors was confirmed in studies using the contingent capture paradigm, where no N2pc occurred to nonmatching cues (Ansorge, Kiss, Worschech, & Eimer, 2011; Lien, Ruthruff, Goodin, & Remington, 2008).

In the present study, we reexamined attentional selection in a feature search using a novel approach that is based on action execution (Song & Nakayama, 2009). The main purpose of attention outside the laboratory may be to select targets for goal-directed action (Allport, 1987; Neumann, 1987). In the General Discussion section, we will argue that perception, as investigated by the manual discrimination tasks in attentional and contingent capture paradigms, shares a common attentional mechanism with manual reaching action. We therefore looked for changes in goal-directed action attributable to the attentional selection of salient stimuli. In our task, the search display was presented after participants had initiated a reaching movement toward the screen, forcing them to do the search with the hand in flight (see also Chapman et al., 2010a, 2010b). We recorded reaching trajectories and measured deviations toward the irrelevant distractor. If the distractor was selected for goal-directed action, the hand should move toward its location. Importantly, our measure allows for a continuous and spatial

measure of attentional selection, quite similar to the N2pc. In related research, it was observed that reaching movements deviated toward salient elements that were presented at potential target locations before the target appeared (Wood et al., 2011), suggesting that salient elements during our visual search task may have a similar effect.

We presented a matrix of white vertical bars containing an orientation singleton (see Fig. 1a). Participants were asked to touch the orientation singleton. The orientation singleton was always tilted to the left, allowing for feature search. A red vertical bar was the irrelevant but salient color singleton that was presented on 50% of the trials. Because feature search was possible (i.e., fixed target), we consider our experiments a conservative test of the attentional capture hypothesis. For instance, the electrophysiological marker of attentional deployment, the N2pc, was absent for irrelevant distractors when feature search was possible, and the increase in RTs was much smaller (Burra & Kerzel, 2013; Lamy & Yashar, 2008; Pinto, Olivers, & Theeuwes, 2005).
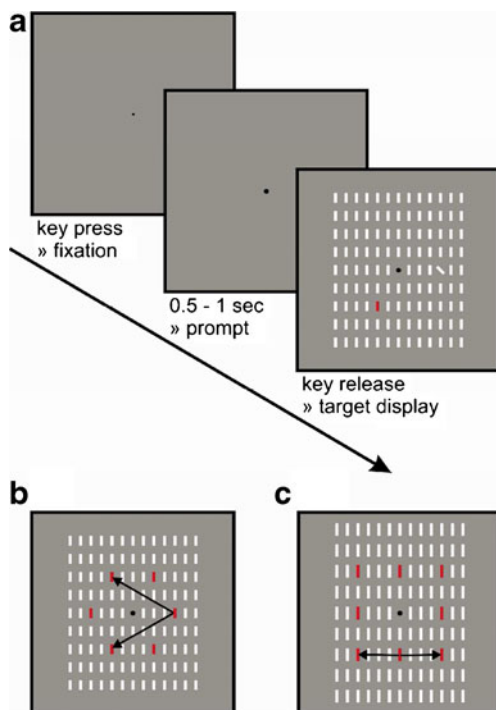

Fig. 1 a Experimental procedure. All stimuli are drawn to scale. In panels b and c, the possible target/distractor positions are represented by red bars. **b** Stimulus layout in Experiments 1, 3, and 4. The distractors were at ±120° of rotation from the target. For a target at the 3 o'clock position, the distractor was either at the 7 or at the 11 o'clock position. **c** Stimulus layout in Experiment 2. Distractors were presented only when the targets were above, below, left, and right of the fixation mark. The distractors were presented at ±45° of rotation. For a target below fixation, the distractor was in either the lower left or the lower right corner

Experiment 1

Method

Participants

Sixteen psychology students at the University of Geneva participated for class credit. They reported normal or corrected-to-normal vision. The study was approved by the ethics committee of the Faculty of Psychology and Educational Sciences, and informed consent was given before the experiment started.

Apparatus and stimuli

The 3-D coordinates of manual movements were recorded at a sample frequency of 150 Hz by means of a marker positioned on the nail of the right index finger (CMS20S, Zebris Medical GmbH, Isny im Allgäu, Germany). The stimuli were presented on a 21-in. CRT monitor (85 Hz, 1,280 × 1,024 pixels) equipped with a touch interface (IntelliTouch, Elo Touchsystems, Menlo Park, CA) at a distance of ~65 cm from the participant. On the screen, 1 cm corresponded to ~0.88° of visual angle. Reaching responses were initiated by pressing the "arrow down" key on a computer keyboard. The central fixation mark was 39 cm in front of and 27 cm above the index finger on the start key. The center-to-center distance between stimuli was 2.6 cm vertically and 1.5 cm horizontally. There were 13 columns and 9 rows. The central element was replaced by a black fixation disk 0.6 cm in diameter. The bars had a width of 0.4 cm and a height of 1.4 cm. The target and distractor elements were placed on a circle with a radius of 6 cm. There were six possible target/distractor positions, which are shown in Fig. 1b. Adjacent positions were separated by 60° of rotation, starting at the 3 o'clock position. Target and distractor were separated by 120° of rotation, which resulted in a distance of 10.4 cm. The context elements and target were white (110 cd/m^2), and the colored bar was red (CIE 1931: $x = 0.612$, $y = 0.338$, $l = 20.5$ cd/m^2), on a gray background (54 cd/m^2). The combined color and luminance difference made the color singleton very salient. The orientation singleton was a bar tilted to the left by 45° of rotation.

Procedure

Participants started a trial by pressing the "arrow down" key, which triggered the appearance of the fixation mark (see Fig. 1a). After a random interval (uniform distribution) between 0.5 and 1 sec, a change in the size of the fixation point (from 0.3° to 0.6°) and a beep prompted participants to lift the finger. When participants lifted the finger, the search display was shown, and observers reached the tilted element.

The release of the key was considered the RT, and the interval between key release and contact with the touch screen as movement time (MT). RTs had to be between 100 and 500 ms, and MTs had to be less than 600 ms. The short allowable MT forced observers to start moving toward the screen right after key release. Doing the visual search task without moving toward the screen inevitably resulted in time-out errors. The main purpose of the strict MT limit was to avoid a "wait-and-search" strategy, so it was set to be unrealistically difficult: Feedback about slow movements was given on 11% of the trials.

Furthermore, the ultrasound microphones did not capture the ultrasound pulse from the markers if the finger was turned away from the microphone or when the movement was very rapid, as in jerky movements. Observers were therefore instructed to move smoothly and continuously toward the target position after key release. The visual search task was to be performed while the movement was on-going. Visual error feedback was given at the end of the trial.

There were six target/distractor positions (cf. Fig. 1b), and all positions were equally likely. On half of the trials, no distractor was shown. On the other half, the distractor was shown at 120° clockwise or counterclockwise from the target. Observers worked through four blocks of 120 trials that were separated by a short break. Before data collection, participants received at least 30 practice trials.

Results

The data set of 1 participant was removed because 32% of the trials contained missing samples. After inspection of the distribution of MTs, a limit of 700 ms was chosen to remove MT outliers, which amounted to 2% of the trials. Error rates are presented in Table 1.

Each trajectory was resampled to yield 200 samples that were normalized with respect to depth (i.e., the axis from the participant to the screen). After spatial normalization, the time of each sample was recovered by interpolation. We then averaged trajectories for the six target positions without distractor and for the six target positions with distractor, separately for each of the two distractor positions (counterclockwise and clockwise). Thus, there were 6 unique conditions without distractor and 12 unique conditions with distractor.

The analysis of reach trajectories is illustrated in Fig. 2. For each of the six conditions without distractor, we calculated the distance between the trajectory to a target at position P1 and the trajectory to a target at position P2 that was rotated 120° clockwise or counterclockwise. The distance was calculated in 3-D for each of the 200 samples. Because the trajectories were normalized with respect to depth, the samples were approximately equidistant in depth. Therefore, the distance in depth between samples can be graphically represented by the grid lines in Fig. 2.

For each of the 12 conditions with distractor, we calculated how much the trajectory toward the target at P1 deviated toward the distractor at position P2. To this end, we referenced the trajectory with target at position P1 and distractor at position P2 to trajectories with targets at P1 and P2. The resulting distance indicated not simply the difference to the trajectory with a target at P1 in the absence of distractors, but also the deviation from the trajectory to target at P1 in the direction of the trajectory to a target at P2.

More precisely, the following calculations were carried out (see Fig. 2). A sample i on the trajectory to the target T1 is referred to as $T1_i$ and the corresponding point at about the same depth when moving to a target at T2 as $T2_i$. The corresponding point on the trajectory to a target at T1 with a distractor at D2 is referred to as $T1D2_i$. We first determined a straight line between $T1_i$ and $T2_i$, which we refer to as g. Next, we dropped a perpendicular of $T1D2_i$ on g and calculated the distance of this point to $T1_i$. If the point was between $T1_i$ and $T2_i$, the sign of the distance was positive. If the point was opposite of $T1_i$, the sign was negative. If the distractor had no effect, the distance between $T1_i$ and $T1D2_i$ would be zero. If the participant went to the distractor at P2 instead of the target at P1, there would be a distance between $T1_i$ and $T1D2_i$ that was equal to the distance between $T1_i$ and $T2_i$.

We averaged the distances between T1 and T2 and between T1 and T1D2 across all target positions. The distances for the target at position P1 with distractor at position P2 are

Table 1 Mean error rates and percentage of retained trials (error types are not exclusive)

	Error Type						Retained
	Choice	Early RT	Late RT	Slow MT	Recording	Fixation	Trials
Exp. 1	1%	3%	1%	2%	6%	–	90%
Exp. 2	0%	3%	1%	1%	7%	–	90%
Exp. 2 fix.	1%	2%	2%	2%	5%	14%	79%
Exp. 3	0%	4%	0%	1%	6%	–	90%

Note. RT = reaction time, MT = movement time, recording = missing samples in the traces, fix. = with eye fixation
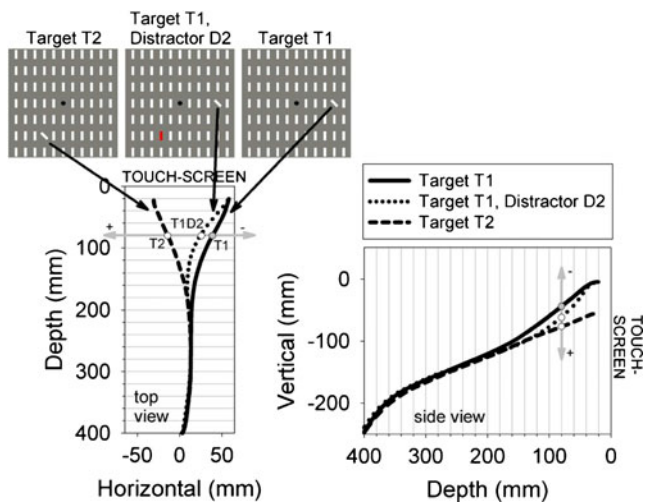
Fig. 2 Top view (left panel) and side view (right panel) of mean trajectories of a single participant. The finger was at about 39 cm in front and 27 cm below the center of the touch screen at the start of a trial. The trajectories are shown for the targets at positions P1 (3 o'clock) with and without distractor and for the target P2 (7 o'clock). Trajectories were resampled such that 200 samples resulted that were equally spaced along the depth axis. The distance between the trajectory to target T1 and T2, as well as the distance between target T1 with and without distractor, was calculated. The latter distance was determined in the direction of the trajectory to target T2. Because the trajectories were normalized with respect to depth, the calculated distances correspond graphically to the distance along the gray grid lines

shown in Fig. 3. As a reference, we also plot the distance between trajectories to targets at P1 and P2. Observers follow a default trajectory for the first ~270 ms, which results in very small differences between the trajectories during the

first phase of the movement. At ~270 ms, the finger starts to move toward a specific stimulus, which increases the distance between trajectories to targets at P1 and P2. At the same time, there is an increase in the deviation toward the distractor.

To quantify the deviation toward the distractor, we compared the mean deviation in the peak interval with the mean deviation in a baseline interval at the beginning of the movement. The baseline interval was defined from the 20th to the 60th sample, which corresponds to the initial 10%–30% of the trajectory. Note that there were 200 equally spaced samples from the starting point of the hand on the keyboard to the endpoint of the movement on the screen. To determine the peak interval, we analyzed the deviation toward the distractor after averaging across participants. First, we determined the peak deviation and calculated the mean deviation in the baseline interval. Next, we used the value at 50% between the baseline deviation and the peak deviation to delimit the peak interval at the rising and falling flank. Then the limits of the peak interval were applied to individual data. Thus, the range of samples (out of 200) in the baseline and peak intervals was the same across participants, but the time of the two intervals varied between participants because individual velocities were different. The average time intervals and average deviations in the baseline and peak intervals are shown in Table 2.

Average trajectories are shown in Fig. 3. The baseline interval was between 72 and 170 ms after response onset, and the peak interval was from 273 to 445 ms, comprising 67 samples (cf. Table 2). The mean deviation was larger during
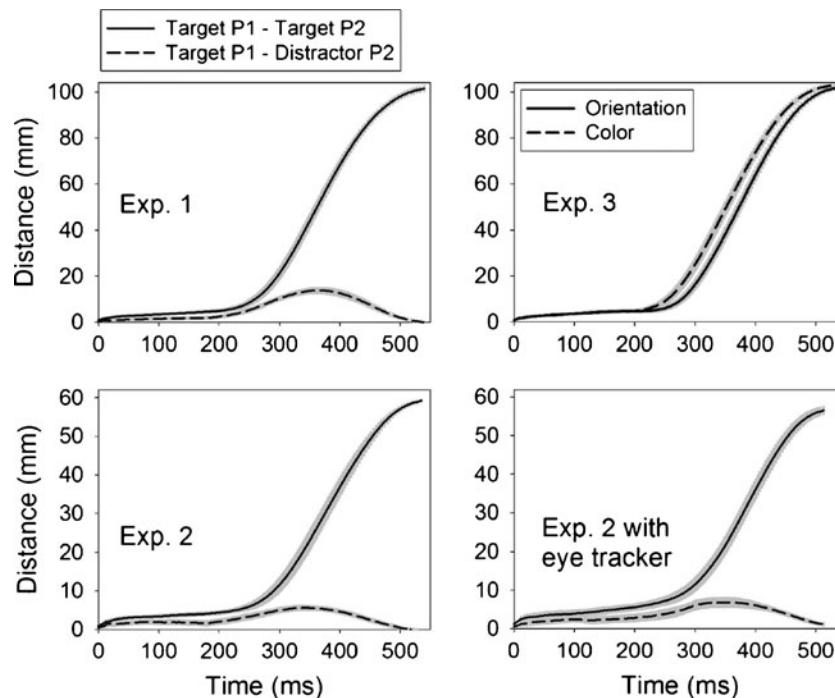

Fig. 3 Results of Experiments 1–4. The gray bars show the between-subjects standard error of the mean

Table 2 Mean interval times (measured from movement onset) and mean deviation toward the distractor for the baseline and peak intervals

	Interval (ms)		Deviation (mm)	
	Baseline	Peak	Baseline	Peak
Exp. 1	72–170	273–445	1.6	11.5
Exp. 2	81–180	261–421	1.8	4.9
Exp. 2 fix.	88–183	272–434	2.4	6.1

Note. fix. = with eye fixation

the peak interval than during the baseline interval (11.5 vs. 1.6 mm), $t(14) = 10.9$, $p < .001$.

RTs did not differ between trials with and without distractor (247 ms in both conditions), $p = .595$, but MTs were slightly longer with than without distractor (568 vs. 557 ms), $t(14) = 7.1$, $p < .001$. In this and the following experiments, choice errors were rare (1%), suggesting that attentional capture did not affect the landing position of the finger on the screen.

Discussion

We observed that goal-directed reaching movements deviated toward the position of a salient distractor in a visual search task. About 270 ms after movement onset, participants left the default trajectory and moved toward the orientation singleton. At the same time, there was a deviation toward the color singleton. The deviation toward the distractor was corrected after 160 ms when the hand returned to the trajectory without a distractor. Because the target remained the same throughout the experiment, participants were able to search for a particular feature and did not have to rely on singleton detection. In contrast to previous studies using the N2pc as a spatial marker of visual selection, we observed evidence for attentional capture by task-irrelevant but salient visual elements in a feature search.

Experiment 2

While we wish to interpret the spatial deviation as evidence for attentional selection of the irrelevant object, an alternative explanation is possible. Because target and distractor locations were on opposite sides of fixation, simply continuing on the default trajectory would also result in a deviation toward the distractor. Figure 4a, b shows that without measuring the default trajectory toward the center of the screen, it is not possible to attribute the deviation unambiguously to attraction by the irrelevant distractor. It may just as well be a delay in the selection of the target that could be attributed to spatial filtering.
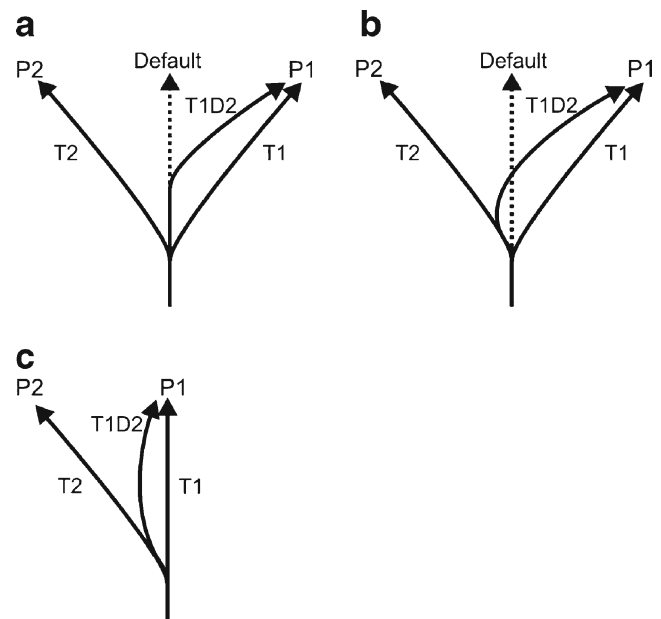

Fig. 4 Schematic drawing of reach trajectories. On target-only trials, the hand moves to targets at positions P1 or P2. The respective trajectories are referred to as T1 and T2. When a distractor is present at position P2 and the target is at P1, the trajectory is referred to as T1D2. **a** The distractor delays the decision to move to P1, and participants continue on the default trajectory before moving toward P1. **b** The distractor attracts the reach before participants correct the movement toward the target at P1. Using the analysis described in the text, the situations illustrated in panels a and b will show that the trajectory T1D2 was between trajectories T1 and T2, resulting in a deviation toward the distractor. **c** The target T1 is shown on the default trajectory. Only if participants move toward the distractor at P2 will the analysis indicate a deviation toward the distractor. Staying on the default trajectory does not result in a deviation toward the distractor

To decide between attraction and delay, we changed the spatial layout of target and distractor positions. Now the targets were presented on the cardinal axes and the distractors in the corners of a virtual square (cf. Fig. 1c). Because moving to the target required staying on the default trajectory, only attraction toward the distractor would result in deviations. For instance, let us consider the position above fixation as P1 and the position in the upper left as P2 (cf. Fig. 4c). To move to P1, participants continue on the default horizontal trajectory to the center of the screen. If a distractor at P2 attracted the reach, this would result in a horizontal deviation that would fall between target-only reaches toward P1 and P2. If selection of the target at P1 was delayed because of a distractor at P2, the trajectories would remain horizontally aligned with a reach to P1.

Method

Fourteen new students participated, but the data set of 1 was discarded due to excessive errors (38%). The methods were as in Experiment 1, with the following exceptions. The spatial layout of the target positions was changed to a square,

with possible target positions at the midpoints and corners of the square outline (see Fig. 1c). Thus, there were eight, not six, possible target positions. Distractors were shown at 45° clockwise or counterclockwise with respect to the target. The vertical and horizontal distance between adjacent target positions was 6.2 cm. To keep the number of repetitions in a reasonable range, no distractors were shown when the target appeared in the corner positions. These conditions could be safely omitted, since only the positions on the cardinal axes decide between the two hypotheses. Overall, there were 480 trials administered in four blocks. During the experiment, feedback about long MTs was given on 10% of the trials.

In a group of 10 additional participants, we monitored eye fixation by means of an EyeLink 2 eyetracker (SR Research Ltd, Kanata, Ontario, Canada). If a saccade larger than 1° occurred between stimulus onset and 100 ms before contact with the screen, a fixation error was signaled to the participant. The data set of 1 participant was removed due to excessive errors (43%).

Results

The deviation toward the distractor was larger during the peak than during the baseline interval (4.9 vs. 1.8 mm), $t(12) = 4.92$, $p < .001$. The baseline interval was from 81 to 180 ms, and the peak interval from 261 to 421 ms, comprising 72 samples. RTs were not different with and without distractor (255 vs. 257 ms), $p = .296$, but MTs were 2 ms shorter with than without distractor (559 vs. 561 ms), $t(12) = 2.35$, $p = .037$.

The results were unchanged with fixation control. The deviation was larger during the peak than during the baseline interval (6.1 vs. 2.4 mm), $t(8) = 4.77$, $p = .001$. The baseline and peak intervals were from 88 to 183 and 272 to 434 ms, respectively. The peak interval comprised 71 samples. RTs were not different with and without distractor (267 vs. 266 ms), $p = .365$, but MTs were 3 ms shorter with than without distractor (536 vs. 539 ms), $t(8) = 2.44$, $p = .041$.

Discussion

We changed the spatial layout of distractor and target positions to rule out an alternative account in terms of selection delay. As compared with Experiment 1, the results were unchanged, suggesting that reach trajectories were attracted by the salient distractor. Furthermore, it did not matter whether observers freely viewed the stimuli or maintained fixation during the reach. It has been observed that there is a close temporal (Adam, Buetti, & Kerzel, 2012; Neggers & Bekkering, 2000; Prablanc, Echalier, Komilis, & Jeannerod, 1979) and spatial (Sailer, Eggert, & Straube, 2002; Song & McPeek, 2009) coupling between eye and hand movements. If the eye was captured by salient stimuli, as has been shown in previous research (van Zoest & Donk, 2008; van Zoest,

Donk, & Theeuwes, 2004), the hand may have followed the eye. If this was true, changes in movements of the hand would result from oculomotor capture. However, we observed similar results with and without eye fixation, and we therefore think that it is unlikely that eye movements contributed substantially to the deviations in reach trajectories.

Experiment 3

Bottom-up models of attention assume that attention is drawn to the most salient element first, regardless of whether this element is the target or not (Itti & Koch, 2001; Theeuwes, 2010). Consistent with this idea, previous studies have reported that RTs and an electrophysiological marker of attention, the N2pc, have shorter latencies with more salient stimuli (Töllner, Zehetleitner, Gramann, & Müller, 2011). To confirm the previously reported latency differences in our paradigm, we compared reaches to color and orientation targets. The same spatial layout as in Experiment 1 was used, but no distractors were presented.

Method

Fourteen new students participated, but the data set of 1 was discarded due to excessive errors (39%). The methods were as in Experiment 1, with the following exceptions. The color and orientation singletons were shown with equal probability. No distractors were presented. Participants were instructed to touch the element that was different from the context. During the experiment, feedback about long MTs was given on 12% of the trials.

Results

We calculated the distance between trajectories to targets that were separated by 120° clockwise and counterclockwise, as in Experiment 1. We used a jackknife method to compare onset times (Miller, Patterson, & Ulrich, 1998). The onset time was defined as the midpoint between baseline and peak distance. Average onset times were calculated after removing each participant once, resulting in 13 values. The t-values were adjusted to take into account the reduced variability. The onset of the deviation occurred earlier with color than with orientation targets (357 vs. 379 ms), $t_{adj}(13) = 2.8$, $p = .015$. RTs were not different between color and orientation targets (243 ms for both), $p = .932$, but MTs were 3 ms shorter with color than with orientation targets (563 vs. 560 ms), $t(13) = 3$, $p = .01$.

Discussion

The reach trajectories showed that the decision to move toward color singletons was taken 22 ms earlier than the

decision to move to orientation singletons. In models of visual search (e.g., Wolfe, 2007), salience is defined as the summed discrepancy of one display element with the surrounding with respect to feature dimensions such as color or orientation. Because the color singleton stood out from the context in two dimensions (color and luminance), whereas the orientation singleton differed only in one, we assumed that the color singleton was more salient than the orientation singleton, which is confirmed by the shorter latencies for color singletons.

Theeuwes (1991) already showed that attentional capture occurred only for distractors that are more salient than the target (but see Zehetleitner, Koch, Goschy, & Müller, 2013). He proposed that feature differences are computed in parallel across the visual field and that larger feature differences are available earlier. Subsequently, attention is directed to the feature differences in the order of availability. The present results are consistent with this idea, because the combined color and luminance differences constitute a powerful distractor in Experiments 1 and 2, and the present experiment shows that the combined color and luminance differences were available earlier than orientation differences.

The shorter latencies may at least partially account for the involuntary selection of salient color singletons and their influence on reach trajectories. However, it should be noted that contrary to Theeuwes's (1991) initial claim, attentional capture may also be observed with distractors that are slightly less salient than the target, suggesting a probabilistic nature of the selection process (Zehetleitner et al., 2013). Furthermore, note that the presumed latency differences concern feature differences and not the features themselves. Purushothaman, Patel, Bedell, and Ogmen (1998) found that brighter stimuli have a shorter latency. In our study, the color target had a lower luminance than the orientation target. Therefore, luminance-related differences in latency would predict slower selection of the color singleton, but we found the opposite.

General discussion

We recorded hand movements while participants searched for an orientation singleton. On half of the trials, a salient but irrelevant color singleton was presented, and we observed that reaches deviated toward the distractor. The deviations were not accounted for by a delay in the selection of the target, as predicted by filtering costs, or by eye movements to the salient distractor. Therefore, we conclude that attentional selection of a salient distractor for action biased the trajectories of reaching movements toward the distractor. Furthermore, we found that selection of salient color targets occurs more rapidly than selection of less salient orientation targets. Overall, our results show that action is captured by salient stimuli in an online visual search task.

There are a number of reasons to believe that attentional selection for visually guided reaching movements is tightly coupled to attentional selection for perception. First, it has been demonstrated that perception is enhanced at the goal of a reaching movement briefly before the hand movement is initiated (Coutté, Faure, & Olivier, 2012; Deubel, Schneider, & Paprotta, 1998), which is also true for several targets in multiple target reaches (reviewed in Baldauf & Deubel, 2010). Thus, attention moves to the reach target(s) before movement initiation.

Second, a number of studies have shown that various movement parameters of reaches mirror the pattern of results observed with keypress responses. Essentially, reaching tasks require target localization, whereas keypresses indicate target detection or identification. Thus, the question is whether the same attentional mechanisms are involved with both task requirements. In a relevant study, Welsh (2011) ran a modified version of the contingent attentional capture paradigm (cf. the introduction) with reaching movements. He observed that reaches deviated toward the cue only if cue and target properties matched, which mirrored the typical pattern of RT results. His interpretation of results was framed in the action-centered theory of attention. According to the theory, the distribution of attention is shaped by the planned action (i.e., attention is allocated to the target of the action), and the attentional selection of an object automatically activates responses that would allow for interactions with the object (Tipper, Lortie, & Baylis, 1992; Welsh, Elliott, & Weeks, 1999). If more than one object is attended, multiple action plans are generated (Cisek & Kalaska, 2005), and the actual movement may represent a mixture between these action plans if there was not sufficient time to inhibit the movement plans to the wrong locations. With little time to inhibit conflicting movement plans, the final response represents a mixture between a response to the target and a response to the distractor, which is what was observed in the present and in previous studies (Buetti & Kerzel, 2009; Scherbaum, Dshemuchadse, Fischer, & Goschke, 2010; Welsh, 2011; Welsh & Elliott, 2004).

A number of other studies also support the idea that target localization in reaching and target detection or identification with keypresses reflects the same selection mechanism. Zehetleitner, Hegenloh, and Müller (2011) observed that the pattern of RTs and MTs in a reaching task mirrored RTs in a detection task. In the reaching task, participants had to touch the odd element in a search display, while a buttonpress was required in the detection task. The temporal parameters in both tasks were equally influenced by variables known to influence search times, such as the feature contrast between target and distractors, repetition of the same target dimension, or cuing of the upcoming target dimension. Similarly, Song and Nakayama (2008) observed that spatial parameters of a reaching movement were influenced by the

same perceptual competition mechanisms as buttonpresses. When participants were asked to touch the odd-colored item in a search display, the trajectories of the reaching movements often deviated toward the distractors when the visual search task was difficult because there were only a few distractors (Song & Nakayama, 2006, 2008). Manual discrimination responses are slower in this condition (Bravo & Nakayama, 1992), which is accounted for by poor perceptual grouping with few distractors.

Finally, Hunt, von Mühlenen, and Kingstone (2007) argued that capture of saccades and manual joystick responses by onset distractors has a common origin. Consistent with previous studies (Wu & Remington, 2003), they observed that without strong speed pressure, the appearance of sudden onsets in a search display slowed down RTs of manual responses, while the RTs of eye movements were unaffected. In contrast, manual responses did not deviate toward the sudden onset, while eye movements did. However, when observers were forced to respond rapidly, the joystick movements showed the same pattern of deviation toward the onset as eye movements, suggesting that the underlying mechanisms were the same but masked by the different latencies of manual and saccadic responses. In our paradigm, the visual search had to be performed online, which produced short-latency deviations of the reach trajectories toward the distractor. Possibly, forcing participants to move even more rapidly to the touch screen would have brought about endpoint errors. That is, observers would have touched the color distractor instead of the orientation target.

On the basis of the assumption that attentional selection for goal-directed action and perception share a common mechanism, our results speak to a recent debate on the causes of attentional capture with perceptual discrimination tasks. It has been claimed that increases in RT with salient distractors are accounted for not by attentional capture, but by nonspatial filtering costs. An important argument against attentional capture was the absence of the N2pc to irrelevant but salient distractors with feature searches (Ansorge et al., 2011; Burra & Kerzel, 2013; Kiss, Jolicoeur, Dell'acqua, & Eimer, 2008; Lien et al., 2008; Schubö et al., 2007; Wykowska & Schubö, 2010, 2011). Similar to our reaching trajectories, the N2pc is a continuous and spatial measure of the distribution of attention. Contrary to findings in the N2pc literature, the deviations of reaches toward salient singletons in our experiments suggest that attention is indeed attracted by salient stimuli.

Furthermore, we would like to point out that our paradigm is similar to those in previous studies in which saccadic eye movements were investigated. In the studies by van Zoest et al., observers had to make a saccade to an orientation target, while color (van Zoest & Donk, 2008) or orientation (van Zoest & Donk, 2006) distractors were shown. Saccadic RTs were longer with distractors than on no-distractor trials. Additionally, saccades often landed on the distractor instead

of the target. The tendency to erroneously select the distractor increased with short RTs. Also, saccades often go to sudden onsets that appear at the same time as the target, a phenomenon known as oculomotor capture (Theeuwes, Kramer, Hahn, & Irwin, 1998). Similar to the hand trajectories in our studies, some saccade trajectories veer toward the onset distractor before turning toward the correct target stimulus (Godijn & Theeuwes, 2002).

In sum, our results show that reaches deviate toward salient color singletons in an online visual search task, suggesting that action is captured by salient stimuli. Our results have implications for a recent debate on attentional capture in the additional singleton paradigm. Whereas electrophysiological results suggest that attention was not captured, our results indicate that attentional selection of the salient color singleton occurred, even in a feature search that supposedly prevents attentional capture. Similar to the N2pc, our measure of attentional deployment is spatial and continuous.

References

Adam, J. J., Buetti, S., & Kerzel, D. (2012). Coordinated flexibility: how initial gaze position modulates eye-hand coordination and reaching. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 891–901. doi:10.1037/a0027592

Allport, A. (1987). Selection for action: Some behavioural and neurophysiological considerations of attention and action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action*. Hillsdale, NJ: Erlbaum.

Ansorge, U., & Horstmann, G. (2007). Preemptive control of attentional capture by colour: Evidence from trial-by-trial analyses and orderings of onsets of capture effects in reaction time distributions. *Quarterly Journal of Experimental Psychology, 60*(7), 952–975. doi:10.1080/17470210600822795

Ansorge, U., Kiss, M., Worschech, F., & Eimer, M. (2011). The initial stage of visual selection is controlled by top-down task set: new ERP evidence. *Attention, Perception & Psychophysics, 73*(1), 113–122. doi:10.3758/s13414-010-0008-3

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Sciences, 16*(8), 437–443. doi:10.1016/j.tics.2012.06.010

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics, 55*(5), 485–496.

Baldauf, D., & Deubel, H. (2010). Attentional landscapes in reaching and grasping. *Vision Research, 50*(11), 999–1013. doi:10.1016/j.visres.2010.02.008

Becker, S. I. (2007). Irrelevant singletons in pop-out search: attentional capture or filtering costs? *Journal of Experimental Psychology: Human Perception and Performance, 33*(4), 764–787. doi:10.1037/0096-1523.33.4.764

Bravo, M. J., & Nakayama, K. (1992). The role of attention in different visual-search tasks. *Perception & Psychophysics, 51*(5), 465–472.

Buetti, S., & Kerzel, D. (2009). Conflicts during response selection affect response programming: reactions toward the source of

stimulation. *Journal of Experimental Psychology: Human Perception and Performance, 35*(3), 816–834. doi:10.1037/a0011092

Burra, N., & Kerzel, D. (2013). Attentional capture during visual search is attenuated by target predictability: Evidence from the N2pc, Pd, and topographic segmentation. *Psychophysiology, 50*(5), 422–430. doi:10.1111/psyp.12019

Chapman, C. S., Gallivan, J. P., Wood, D. K., Milne, J. L., Culham, J. C., & Goodale, M. A. (2010a). Reaching for the unknown: multiple target encoding and real-time decision-making in a rapid reach task. *Cognition, 116*(2), 168–176. doi:10.1016/j.cognition.2010.04.008

Chapman, C. S., Gallivan, J. P., Wood, D. K., Milne, J. L., Culham, J. C., & Goodale, M. A. (2010b). Short-term motor plasticity revealed in a visuomotor decision-making task. *Behavioural Brain Research, 214*(1), 130–134. doi:10.1016/j.bbr.2010.05.012

Cisek, P., & Kalaska, J. F. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron, 45*(5), 801–814. doi:10.1016/j.neuron.2005.01.027

Coutté, A., Faure, S., & Olivier, G. (2012). Influence de la préparation d'une séquence de pointages de cibles sur la recherche visuelle. *L'année psychologique/Topics in Cognitive Psychology, 12*(4), 561–573. doi:10.4074/S0003503312004022

Deubel, H., Schneider, W. X., & Paprotta, I. (1998). Selective dorsal and ventral processing: Evidence for a common attentional mechanism in reaching and perception. *Visual Cognition, 5*(1/2), 81–107.

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology, 99*(3), 225–234.

Folk, C. L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 847–858. doi:10.1037/0096-1523.24.3.847

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 1030–1044.

Godijn, R., & Theeuwes, J. (2002). Programming of endogenous and exogenous saccades: evidence for a competitive integration model. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1039–1054. doi:10.1037/0096-1523.28.5.1039

Hickey, C., McDonald, J. J., & Theeuwes, J. (2006). Electrophysiological evidence of the capture of visual attention. *Journal of Cognitive Neuroscience, 18*(4), 604–613. doi:10.1162/jocn.2006.18.4.604

Hunt, A. R., von Mühlenen, A., & Kingstone, A. (2007). The time course of attentional and oculomotor capture reveals a common cause. *Journal of Experimental Psychology: Human Perception and Performance, 33*(2), 271–284. doi:10.1037/0096-1523.33.2.271

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194–203. doi:10.1038/35058500

Kahneman, D., Treisman, A., & Burkell, J. (1983). The cost of visual filtering. *Journal of Experimental Psychology: Human Perception and Performance, 9*(4), 510–522.

Kiss, M., Jolicoeur, P., Dell'acqua, R., & Eimer, M. (2008). Attentional capture by visual singletons is mediated by top-down task set: new evidence from the N2pc component. *Psychophysiology, 45*(6), 1013–1024. doi:10.1111/j.1469-8986.2008.00700.x

Lamy, D., & Yashar, A. (2008). Intertrial target-feature changes do not lead to more distraction by singletons: target uncertainty does. *Vision Research, 48*(10), 1274–1279. doi:10.1016/j.visres.2008.02.021

Leber, A. B., & Egeth, H. E. (2006). Attention on autopilot: Past experience and attentional set. *Visual Cognition, 14*(4–8), 565–583. doi:10.1080/13506280500193438

Lien, M.-C., Ruthruff, E., Goodin, Z., & Remington, R. W. (2008). Contingent Attentional Capture by Top-Down Control Settings: Converging Evidence From Event-Related Potentials. *Journal of Experimental Psychology: Human Perception and Performance, 34*(3), 509–530. doi:10.1037/0096-1523.34.3.509

Luck, S. J., & Hillyard, S. A. (1994). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance, 20*(5), 1000–1014.

Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology, 35*(1), 99–115.

Neggers, S. F., & Bekkering, H. (2000). Ocular gaze is anchored to the target of an ongoing pointing movement. *Journal of Neurophysiology, 83*(2), 639–651.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on Perception and Action* (pp. 361–394). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Pinto, Y., Olivers, C. N. L., & Theeuwes, J. (2005). Target uncertainty does not lead to more distraction by singletons: Intertrial priming does. *Perception & Psychophysics, 67*(8), 1354–1361. doi:10.3758/BF03193640

Prablanc, C., Echalier, J. F., Komilis, E., & Jeannerod, M. (1979). Optimal response of eye and hand motor systems in pointing at a visual target. I. Spatio-temporal characteristics of eye and hand movements and their relationships when varying the amount of visual information. *Biological Cybernetics, 35,* 113–124. doi:10.1007/BF00337436

Purushothaman, G., Patel, S. S., Bedell, H. E., & Ogmen, H. (1998). Moving ahead through differential visual latency. *Nature, 396*(6710), 424. doi:10.1038/24766

Sailer, U., Eggert, T., & Straube, A. (2002). Implications of distracter effects for the organization of eye movements, hand movements, and perception. *Progress in Brain Research, 140,* 341–348. doi:10.1016/S0079-6123(02)40061-1

Scherbaum, S., Dshemuchadse, M., Fischer, R., & Goschke, T. (2010). How decisions evolve: the temporal dynamics of action selection. *Cognition, 115*(3), 407–416. doi:10.1016/j.cognition.2010.02.004

Schubö, A., Schröger, E., Meinecke, C., & Müller, H. J. (2007). Attentional resources and pop-out detection in search displays. *Neuroreport, 18*(15), 1589–1593. doi:10.1097/WNR.0b013e3282efa08e

Song, J.-H., & McPeek, R. M. (2009). Eye-hand coordination during target selection in a pop-out visual search. *Journal of Neurophysiology, 102*(5), 2681–2692. doi:10.1152/jn.91352.2008

Song, J.-H., & Nakayama, K. (2006). Role of focal attention on latencies and trajectories of visually guided manual pointing. *Journal of Vision, 6*(9). doi: 10.1167/6.9.11

Song, J.-H., & Nakayama, K. (2008). Target selection in visual search as revealed by movement trajectories. *Vision Research, 48*(7), 853–861. doi:10.1016/j.visres.2007.12.015

Song, J.-H., & Nakayama, K. (2009). Hidden cognitive states revealed in choice reaching tasks. *Trends in Cognitive Sciences, 13*(8), 360–366. doi:10.1016/j.tics.2009.04.009

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics, 50*(2), 184–193.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics, 51*(6), 599–606.

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica, 135*(2), 77–99. doi:10.1016/j.actpsy.2010.02.006

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go where we want them to go: Capture of the eyes by new objects. *Psychological Science, 9*(5), 379–385.

Tipper, S. P., Lortie, C., & Baylis, G. C. (1992). Selective reaching: evidence for action-centered attention. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 891–905.

Töllner, T., Müller, H. J., & Zehetleitner, M. (2011a). Top-down dimensional weight set determines the capture of visual attention: Evidence from the PCN component. *Cerebral Cortex*. doi:10.1093/cercor/bhr231

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

Treisman, A., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the cost of filtering. *Perception & Psychophysics, 33*(6), 527–532.

van Zoest, W., & Donk, M. (2006). Saccadic target selection as a function of time. *Spatial Vision, 19*(1), 61–76.

van Zoest, W., & Donk, M. (2008). Goal-driven modulation as a function of time in saccadic target selection. *The Quarterly Journal of Experimental Psychology: Section A, 61*(10), 1553–1572. doi:10.1080/17470210701595555

van Zoest, W., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 746–759. doi:10.1037/0096-1523.30.4.746

Welsh, T. N. (2011). The relationship between attentional capture and deviations in movement trajectories in a selective reaching task. *Acta Psychologica, 137*(3), 300–308. doi:10.1016/j.actpsy.2011.03.011

Welsh, T. N., & Elliott, D. (2004). Movement trajectories in the presence of a distracting stimulus: evidence for a response activation model of selective reaching. *The Quarterly Journal of Experimental Psychology: Section A, 57*(6), 1031–1057. doi:10.1080/02724980343000666

Welsh, T. N., Elliott, D., & Weeks, D. J. (1999). Hand deviations toward distractors. Evidence for response competition. *Experimental Brain Research, 127*(2), 207–212. doi:10.1007/s002210050790

Wolfe, J. M. (2007). Guided Search 4.0: Current Progress with a model of visual search. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York: Oxford.

Wood, D. K., Gallivan, J. P., Chapman, C. S., Milne, J. L., Culham, J. C., & Goodale, M. A. (2011). Visual salience dominates early visuomotor competition in reaching behavior. *Journal of Vision, 11*(10). doi: 10.1167/11.10.16

Wu, S. C., & Remington, R. W. (2003). Characteristics of covert and overt visual orienting: Evidence from attentional and oculomotor capture. *Journal of Experimental Psychology: Human Perception and Performance, 29*(5), 1050–1067. doi:10.1037/0096-1523.29.5.1050

Wykowska, A., & Schubö, A. (2010). On the temporal relation of top-down and bottom-up mechanisms during guidance of attention. *Journal of Cognitive Neuroscience, 22*(4), 640–654. doi:10.1162/jocn.2009.21222

Wykowska, A., & Schubö, A. (2011). Irrelevant singletons in visual search do not capture attention but can produce nonspatial filtering costs. *Journal of Cognitive Neuroscience, 23*(3), 645–660. doi:10.1162/jocn.2009.21390

Zehetleitner, M., Goschy, H., & Müller, H. J. (2012). Top-down control of attention: it's gradual, practice-dependent, and hierarchically organized. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 941–957. doi:10.1037/a0027629

Zehetleitner, M., Hegenloh, M., & Müller, H. J. (2011). Visually guided pointing movements are driven by the salience map. *Journal of Vision, 11*(1). doi: 10.1167/11.1.24

Zehetleitner, M., Koch, A. I., Goschy, H., & Müller, H. J. (2013). Salience-based selection: Attentional capture by distractors less salient than the target. *PLoS One, 8*(1), e52595. doi:10.1371/journal.pone.0052595