

Novelty biases attention and gaze in a surprise trial

Gernot Horstmann^{1,2,3} · Arvid Herwig^{1,3}

Published online: 21 October 2015
© The Psychonomic Society, Inc. 2015

Abstract While the classical distinction between task-driven and stimulus-driven biasing of attention appears to be a dichotomy at first sight, there seems to be a third category that depends on the contrast or discrepancy between active representations and the upcoming stimulus, and may be termed novelty, surprise, or prediction failure. For previous demonstrations of the discrepancy-attention link, stimulus-driven components (saliency) may have played a decisive role. The present study was conducted to evaluate the discrepancy-attention link in a display where novel and familiar stimuli are equated for saliency. Eye tracking was used to determine fixations on novel and familiar stimuli as a proxy for attention. Results show a prioritization of attention by the novel color, and a de-prioritization of the familiar color, which is clearly present at the second fixation, and spans over the next couple of fixations. Saliency, on the other hand, did not prioritize items in the display. The results thus reinforce the notion that novelty captures and binds attention.

Keywords Surprise · Attention · Prioritization · Unexpectedness · Eye movements · Gaze · Fixation

✉ Gernot Horstmann
gernot.horstmann@uni-bielefeld.de

¹ Department of Psychology, Bielefeld University,
33501 Bielefeld, Germany

² Centre for Interdisciplinary Research, Bielefeld University,
Bielefeld, Germany

³ Cognitive Interaction Technology – Excellence Cluster, Bielefeld
University, Bielefeld, Germany

Introduction

An important distinction has been drawn between task-driven attention and stimulus-driven attention. According to this distinction, task-driven attention operates via attentional control settings (Folk, Remington, & Johnston, 1992) or target templates (Duncan & Humphreys, 1989) that are generated in the context of a task with the aim of optimally guiding attention to task-relevant items. For example, in a laboratory task where participants search for a green letter among red letters, the color green would be the most important property of the attentional set or target template, leading to a fast attentional orienting to the green item when the display appears (Yantis & Egeth, 1999).

On the other hand, task-irrelevant objects may also involuntarily capture attention even when mismatching the attentional set or target template. The general function of involuntary attention capture is to enable the organism to respond to possibly relevant events beyond the task. For example, a rapidly onsetting stimulus in an otherwise stationary display will quickly draw attention, even if participants have firm knowledge that the onsetting stimulus is irrelevant to the task (Jonides, 1981).

The dichotomy of task-driven versus stimulus-driven attention becomes problematic when considering involuntary attention to surprising, unexpected, or novel stimuli. For example, Horstmann (2002) tested spontaneous attentional orienting towards a task-irrelevant, unannounced novel color singleton. Participants performed a demanding letter search during the pre-critical trials, where all letters were presented in the same color on each trial (e.g., red). In the critical trial the target letter was rendered in a novel color (e.g., green), while the remaining non-target letters were presented in the familiar color. This changed search performance from very inefficient search to almost efficient search; evidently participants

quickly oriented attention to the novel color singleton. These results were replicated and extended (Horstmann, 2005; 2006; Horstmann & Becker, 2008; 2011; Horstmann & Herwig, 2015; Becker & Horstmann, 2011).

Clearly, since the target never appeared with the novel color or as a singleton before, the attentional response on the first presentation of the novel color singleton is unlikely to have been task-driven (Gibson & Jiang, 1998). Task-driven processes should rather have prioritized the familiar color in the display. At the same time, stimulus-driven attention also seems not to apply, as previous research has shown that expected, repeated, or familiar color singletons do not capture attention in the context of a visual search task (e.g., Eimer & Kiss, 2008; Horstmann & Ansorge, 2006; Yantis & Egeth, 1999), except when participants actively search for a singleton on another dimension (Bacon & Egeth, 1994) that changes from trial to trial (Theeuwes, deVries, & Godijn, 2003). Theorists of singleton-driven attention (for an overview, see Theeuwes, 2010) also predict effects of singletons to be early, in the time range of 60–150 ms (Kim & Cave, 1999) in covert attention tasks, and on the first eye movement in overt attention tasks (Theeuwes et al., 2003). In contrast, with novel singletons attention capture occurred later, after around 400 ms, which has been shown for covert shifts (Horstmann, 2006; Horstmann & Becker, 2008), as well as for overt shifts (Horstmann & Herwig, 2015) of attention.

Thus, there seems to be a third category besides task-driven and stimulus-driven attention that is related to novelty. According to the discrepancy-hypothesis (for a recent overview, see Horstmann, 2015), this third category of attentional responses depends on the discrepancy between an expectation and the upcoming stimulus. The expectation can be conceptualized as an active representation, and is usually based on the objects experienced before in the respective type of situation. The discrepancy between the expectation and the novel stimulus is thought to directly influence priority values, which in turn determine the next focus of covert and overt attention. On this view, stimuli that are novel and deviate from an expectation or prediction have the potential to bias attention.

Aim of the present study

In previous studies on the discrepancy hypothesis, the unexpected, surprising, or novel stimulus was always a perceptual singleton – an item that was perceptually unique within the display. From a systematic point of view, the novel color singleton thus had two characteristics: (a) it presented a novel feature and (b) it was perceptually salient. It is thus reasonable to ask whether a novel item biases attention when it is not at the same time salient.

A number of experiments have tested whether salience alone (i.e., salience without feature novelty) captures attention on a surprise presentation (i.e., the first and unannounced

occurrence). Generally it was found that it did not (Horstmann, 2005; Becker & Horstmann, 2011). For example, there was no reduction in the set size effect for a green-among-red singleton target when the surprise presentation was preceded by a random sequence of all-red and all-green non-singleton displays (Horstmann, 2005). The only exception to the general absence of attentional effects of saliency was when all items in a display were novel (Becker & Horstmann, 2011, Experiment 3): After being presented with stationary squares in non-singleton displays, a diamond among rotating squares captured attention on a surprise presentation.

The results discussed thus far suggest that salience alone is not sufficient to drive attention, except in connection with a global surprising display change. However, is feature novelty sufficient? The present study sought evidence for the assumption that feature novelty biases attention even in the absence of salience.

The task was a demanding search where participants searched for a closed ring (target) among rings with a small gap (distractors). In half of the trials, the eight-object display contained a target among seven distractors (target trials), and in half of the trials, eight distractors were presented (target-absent, or “blank” trials). Participants reported the presence versus absence of the target. In Experiment 1, targets and distractors were presented on gray disks, which were – depending on condition – either darker or lighter than the medium gray background. For each participant, all pre-critical trials were identical in terms of object color. The eight disk-ring compounds were presented equidistantly on the outline of an imaginary circle. In the pre-critical trials, all discs in a display were of the same gray (e.g., light gray). In the critical trial, the disks of one half of the array were presented in the alternative gray (e.g., dark gray). The alternative feature was not presented in the experiment before, nor was the display change announced or otherwise signaled to the participants. Importantly, in this arrangement, none of the two features is more perceptually salient than the other in the display. Thus, if the alternative feature is prioritized for attentional selection, this must be entirely due to its novelty or unexpectedness. Experiment 2 used the colors red and green on a black background (e.g., Horstmann, 2002) to test the robustness of the results from Experiment 1.

Eye tracking was used to track gaze position, which in turn served as a proxy for attention (Deubel & Schneider, 1996). Of most importance is performance in target-absent trials, because here eye movements are not influenced by the presence of the target. The critical trial was always a target-absent trial.

If feature novelty biases attention, we expect that early fixations prefer a novel feature over a familiar feature. Although the present experiments are not specifically designed to test a saliency-based account, this view predicts that gaze is attracted in particular to the color borders in the display where stimuli are salient because one of their neighbors has a

different color. Therefore, early fixations should be preferentially directed to the four stimuli in the display that are flanked by a different color (salient stimuli), rather than to the four stimuli in the display that are flanked by the same color (non-salient stimuli).

Experiment 1

Method

Participants

Thirty-two participants were tested in a 10-min session in the eye-tracking laboratory. They were paid €1.

Stimuli

The target was a 1.1° diameter ring with stroke strength of 0.17° . The distractors equaled the target with the only difference of a small radial gap of 0.08° height. Fifteen different gap positions, evenly distributed between 24° and 360° , were used. The rings had the same color as the background, which was medium gray (50 % of maximum luminance, 32 cd/ms^2). The rings were presented on light gray (70 % of maximum luminance, 63 cd/m^2) or dark gray (30 % of maximum luminance, 13 cd/m^2) disks with a diameter of 2.7° (see Fig. 1). Eight disk-ring compounds were presented on each trial arranged equidistantly on the circumference of an invisible circle with a radius of 8.4° .

Apparatus

Stimuli were presented on a 19-in display monitor (100-Hz refresh rate, resolution 1,024 x 768 pixels) at a distance of 71 cm. A video-based tower mounted eye tracker (EyeLink

1000, SR Research, Ottawa, Ontario, Canada) with a sampling rate of 1,000 Hz was used for recording eye movements. The participants' head was stabilized by a chin rest, and the right eye was monitored in all participants.

The experiment was controlled by the Experiment Builder 1.10.165, and data processing was done using Data Viewer 1.11.1 (both from SR Research). For the determination of fixations, standard software settings for cognitive research were used (a velocity threshold of $30^\circ/\text{sec}$ and an acceleration threshold of $8,000^\circ/\text{sec}^2$). Fixations were scored to the nearest stimulus. This choice has two advantages. First, given the single critical trial in a surprise experiment, data loss due to between-stimulus fixations is avoided. Second, as the main hypothesis is about the biasing of attention towards the novel versus familiar colored half of the stimuli, analyzing fixations that are more in the direction of novel versus familiar stimuli seems to be an optimal strategy.

Procedure

The experiment proper comprised one single block of 31 trials; 30 pre-critical trials where all gray disks were of the same gray (either all light gray or all dark gray) and one critical trial where half of the disks were light gray and half were dark gray. The 30 pre-critical trials included 12 practice trials (plus two filler trials between practice and the experimental trials) to stabilize performance, which were not analyzed. Half of the displays in each condition were target trials, and half were blank trials (with no target). The participants' task was to report the presence or absence of the target with a corresponding key press. The critical trial was always a target-absent trial.

Each trial began with a drift correction where participants fixated on a fixation marker at the screen's center and confirmed fixation with a key press with their left hand. Then the search display was presented until a response key press (right hand) was registered.

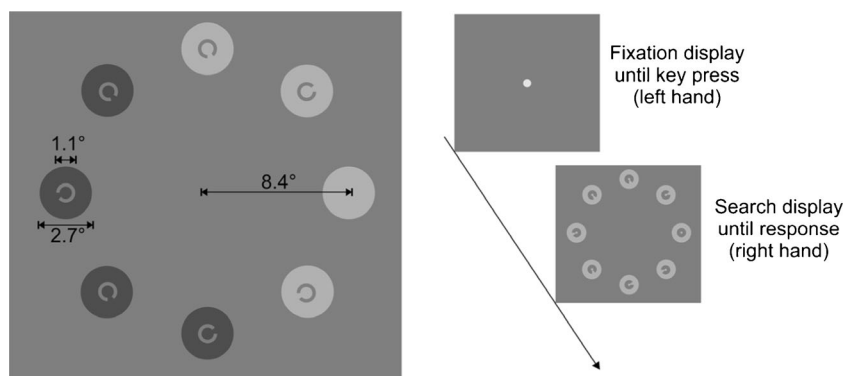


Fig. 1 Schematic representation of the target, a 1.1° diameter ring with stroke strength of 0.17° . The distractors equaled the target with the only difference of a small radial gap of 0.08° height. Fifteen different gap positions, evenly distributed between 24° and 360° , were used. The

rings had the same color as the background, which was medium gray (50 % of maximum luminance, 32 cd/ms^2). The rings were presented on light gray (70 % of maximum luminance, 63 cd/m^2) or dark gray (30 % of maximum luminance, 13 cd/m^2) disks with a diameter of 2.7°

Design

Participants were randomly assigned to two conditions that only differed in the pre-critical trials. For half of the participants, light gray was used in the pre-critical trials, such that in the critical trial, light gray was the familiar and dark gray was the novel color. This assignment was reversed for the other half of the participants. The spatial layout in the critical trial was controlled such that all eight possible rotations of the critical trial display were used equally often. In target trials, the position of the target was random. Distractors were chosen randomly without replacement, with independent sampling for each trial.

Results

General results

Mean valid reaction time (RT; excluding few RTs >6,000 ms to reduce the influence of outliers on average RT) was 1,996 ms in pre-critical target trials, 2,497 ms in pre-critical blank trials, and 3,649 ms in the critical trial (which was a blank trial). All t-tests reported hereafter are two-tailed. Cohen's *d* is reported as a measure of effect size, computed as the difference between the means, divided by the standard deviation of the difference. RT was shorter in the target than in the blank trial condition, $t(30) = 6.49, p < .001, d = 1.16$. Importantly, it was longer in the critical trial, as compared to the corresponding (blank) trials in the pre-critical block, $t(30) = 8.34, p < .001, d = 1.50$. Error rates in the pre-critical trials were 14.6 % in target and 7.3 % in blank trials, $t(31) = 1.93, p = .063, d = 0.34$. In the critical trial, one error (3 %) occurred, reducing the sample size to 31. Although the error rate seems to be lower in the critical trial than in the pre-critical blank trials, this difference did not prove to be significant, $t(31) = 1.01, p = .318, d = 0.18$. A t-test is used for consistency here; an alternative approach is to test whether the proportion from the critical trial (.03) is within the 95 % confidence interval (CI) around the proportion from the pre-critical trials (.02–.22), which is the case. The CI is computed according to a procedure described in Hays (1973, pp. 378–380).

There were fewer fixations in the pre-critical target trials as compared to pre-critical blank trials, (6.91 vs. 10.18), $t(30) = 8.70, p < .001, d = 1.56$. This is expected on the assumption of a self-terminating search in target trials and an exhaustive search in blank trials. Fixations were more frequent in the critical trial (13.67) than in the corresponding pre-critical (blank) trials, $t(30) = 5.38, p < .001, d = 0.97$.

Average fixation duration in the pre-critical trials was 179 ms in target and 195 ms in blank trials, $t(30) = 2.48, p = .018, d = 0.45$. It was longest in the critical trial (216 ms), which differed significantly from the pre-critical (blank) trial duration, $t(30) = 3.00, p = .005, d = 0.54$.

Average saccade amplitude was 5.60 in pre-critical target trials and 6.19 in pre-critical blank trials, $t(30) = 5.72, p < .001, d = 1.03$. Thus, there was a small tendency towards longer saccades in blank trials. Average saccade amplitude in the critical trial (5.89) was somewhat shorter than in the pre-critical (blank) trials, $t(30) = 2.39, p = .023, d = 0.43$.

Main results

The main results are reported in the following. As a first approximation, we looked at the times of two fixations from the critical trial: the first fixation to the novel side (novel side fixation latency) and the first fixation to the familiar colored side (familiar side fixation latency). The novel side fixation latency was shorter on average than the familiar side fixation latency (362 ms vs. 719 ms), $t(30) = 2.6, p = .014, d = 0.46$.

For comparison, two randomly chosen exclusive sets of four adjacent positions were dummy coded for each pre-critical trial, and the corresponding first two fixations to these two sets were determined in the same way as for the critical trial. These two sets were of course indistinguishable for the participant; it is thus expected that each set would on average get the first fixation on half of the trials, and a later fixation on the other half of the trials. In other words, the average reflects the unbiased or random value for the first fixation latency for each of the two sets. The average first fixation time on the subsets of stimuli in the pre-critical trials was 495 ms. This latency differed from the novel side fixation latency, $t(30) = 3.58, p = .001, d = 0.36$, and from familiar side fixation latency: $t(30) = 2.0, p = .053, d = 0.64$. In other words, the gaze shift to the novel feature was earlier, and the shift to the familiar feature was later than expected on the assumption of random priorities.

A second analysis was done on the proportion of fixations on either the novel or the familiar color as a function of *fixation sequence*. That is, we asked how many of the first fixations went to the novel color, how many of the second fixations, and so on. If the novel color is prioritized for selection, one expects that many of the early fixations are on the novel color. This was what we found. While the first fixation was on either of the two colors with equal probabilities, the second and the third fixations were significantly more probably directed to the novel color ($\chi^2 = 5.45, p = .019$, for both). For the ninth and the tenth fixations the pattern is reversed, with more probable fixations on the familiar color ($\chi^2 > 4.17, p < .041$).

Figure 2 depicts part of the results in overview. Note that the X-axis shows the mean latency of the respective fixation. Thus, it can be seen that the first fixation occurs after 197 ms on average, and the second fixation after 374 ms. These latencies give an additional indication about when the prioritization of the novel color begins.

It is conceivable that the first fixation went indiscriminately to either the novel or the familiar color because it was directed

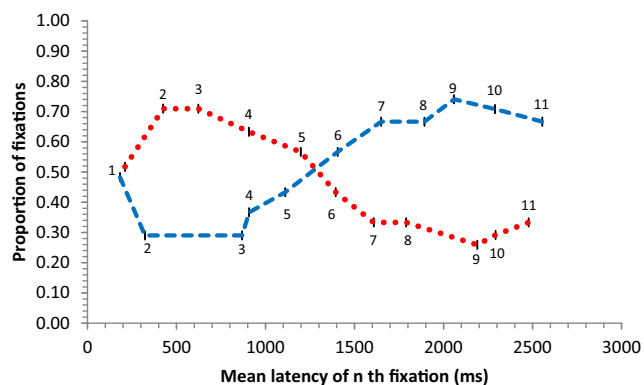


Fig. 2 Plot of the proportion of fixations per ordinal number of the new or old color. X-axis shows the mean latency of the 1st, 2nd, 3rd, etc. fixation. Y-axis shows the proportion of fixations that went either to the new (dotted line) or the old (dashed line) color. The numbers near the line indicate the ordinal number of the fixation

to the salient hotspots in the display, which are the two color borders between the familiar and the novel color (see Fig. 1). To test this possibility, we examined how often the first, second, third, etc. fixation was on one of the four stimuli that made up the salient color border (salient stimuli) versus on one of the four stimuli that are flanked only by same-colored stimuli (non-salient stimuli). The proportions for selecting the salient stimuli were .45, .42, .65, .67, .47, .47, .40, .57, .48, and .46 when analyzed as a function of fixation sequence for the first ten fixations. None of the proportions differed from chance, with only the fourth fixation¹ approaching significance $\chi^2(1; N = 30) = 3.33, p = .067$ (other $\chi^2 < 2.61$). Thus, the saliency of the color border does not play an important role for the first or one of the following fixations.

Discussion

Experiment 1 demonstrated the biasing of attention towards a novel color. The first fixation on a novel item was earlier than on a familiar item, and it was earlier than predicted on the assumption of random priorities. On closer inspection, the second and third fixations were directed more frequently to the novel than to the familiar color. The first fixation, however, was not biased by novelty.

The salient hot-spots in the display at the borders between the two colors did not attract attention. This result suggests that saliency did not bias attention within the present task and with the present stimuli.

Another surprise-related effect was an increase in the number of fixations in the critical trial. Although the exact causes are unknown, it is an indication that the routine search behavior from the pre-critical trials has been interrupted in the critical trial, possibly by shifting search from routine to a more controlled mode (Foerster & Schneider, 2015).

¹ Note that the 4th fixation had a mean latency of 919 ms.

To draw an interim conclusion, the most important aim of the experiment was successfully met: to demonstrate the attentional prioritization for a novel stimulus when the novel stimulus is not confounded with being a singleton. Before other results are discussed in detail, Experiment 2 is reported.

Experiment 2

Experiment 2 is a close replication of Experiment 1, with chromatic color (red and green) rather than achromatic color as familiar and novel features. Many of the preceding experiments on surprise capture and saliency capture have been conducted using chromatic color. It is thus of interest to analyze a replication of Experiment 1 with chromatic color.

Method

Participants

Forty participants were tested individually in 10-min sessions. They were paid €1.

Stimuli, apparatus, and procedure

The stimuli, apparatus, and procedure were with the following two exceptions the same as in Experiment 1. First, background was black (0.8 cd/m^2). Second, instead of the two gray tones, the colors red (RGB intensities: 20 %, 100 %, and 20 %; 31 cd/m^2) and green (20 %, 100 %, and 20 %; 72 cd/m^2) were used.

Design

Participants were randomly assigned to two conditions that differed only in the pre-critical trials. For one condition, red was used in the pre-critical trials, such that in the critical trial, red was the familiar and green was the novel color. For the other condition, this assignment was reversed. Accidentally, eight participants more were tested in the condition with green as the familiar color. As both conditions rendered practically identical results, this difference in sample sizes does not seem to affect the results pattern, and the conditions are pooled for all further analyses. All other details were identical to Experiment 1.

Results

General results

One RT in the critical trial $>10 \text{ s}$ was treated as an outlier and excluded from RT-analysis. The mean valid RT was 1,955 ms in pre-critical target trials, 2,490 ms in pre-critical blank trials,

and 3,797 ms in the critical trial (which was a blank trial). RT was shorter in the target than in the blank trial condition, $t(39) = 6.88, p < .001, d = 1.06$, and it was longer in the critical trial than in the corresponding (blank) trials in the pre-critical block, $t(37) = 15.99, p < .001, d = 1.75$. Error rates in the pre-critical trials were 7.8 % in target trials and 8.9 % in blank trials, $t(39) = 0.50, p = .617, d = 0.07$. One error occurred in the critical trial, reducing the sample size to 39 for analyses including the critical trial. The error rate in the critical trial (2.5 %) was lower than in the pre-critical trials, $t(39) = 2.06, p = .046, d = 0.32$. (The same implication applies when the CI for proportions is used, see Experiment 1.)

There were fewer fixations in the pre-critical target trials as compared to pre-critical blank trials, (7.34 vs. 10.45), $t(39) = 8.40, p < .001, d = 1.23$. Replicating Experiment 1, there were many more fixations in the critical trial (15.43) than in the corresponding (blank) pre-critical trials, $t(38) = 8.28, p < .001, d = 1.13$.

Average fixation duration in the pre-critical trials was 175 ms in target and 185 ms in blank trials, $t(39) = 2.92, p = .005, d = 0.27$. It was longest in the critical trial (201 ms), which differed significantly from the pre-critical fixation duration, $t(38) = 3.40, p = .002, d = 0.46$.

Average saccade amplitude was 5.92 in pre-critical target trials and 6.22 in pre-critical blank trials, $t(39) = 4.09, p < .001, d = 0.64$, similar to Experiment 1. Saccade amplitude in the critical trial (5.92) did not differ significantly from the pre-critical target blank trial amplitudes, $t(39) = 1.67, p = .103, d = 0.09$.

Main results

As before, we looked at the latencies of two fixation latencies from the critical trial, the novel side-fixation latency, and the familiar side-fixation latency. Novel side-fixation latency was much shorter on average than familiar side-fixation latency (397 ms vs. 896 ms), $t(38) = 3.08, p = .004, d = 0.49$.

The analysis of the pre-critical trials (see Experiment 1 for details) revealed an average first-fixation time of 513 ms, as a measure of random priorities. This latency differed from both of the latencies in the critical trial: familiar side-fixation latency: $t(38) = 2.85, p = .007, d = 0.45$; novel side-fixation latency: $t(38) = 2.65, p = .012, d = 0.42$. In other words, the gaze shift to the novel feature was earlier than expected on the assumption of random priorities, while the shift to the familiar feature was later than expected.

A second analysis was done on the proportion of fixations on either the novel or the familiar color as a function of fixation sequence. As in Experiment 1, the first fixation did not discriminate between the two colors. Similar to Experiment 1, fixations 2–6 significantly preferred the novel color, $\chi^2 > 5.77,$

$ps < .016$. In contrast, fixations 9–11 significantly preferred the familiar color ($\chi^2 > 5.11, ps < .023$).

As before, we tested the attraction of gaze by saliency. The proportions for selecting the salient stimuli for fixations 1–10 were .49, .54, .59, .46, .38, .50, .61, .58, .55, and .43. None of the proportions differed from chance (highest $\chi^2 = 2.07, p = .149$). As in Experiment 1, the saliency of the color border played no role for fixations.

Figure 3 shows the main results in overview. Importantly, the first saccade, which did not differentiate between the two colors, occurred at a latency of 211 ms on average, while the second saccade that preferred the novel color occurred at a latency of 375 ms.

Discussion

Experiment 2 perfectly replicated the main results of Experiment 1. Again, the novel color quickly attained priority over selection. In fact, the pattern of manifest prioritization of the novel over the familiar color spanned from the second to the sixth fixation, and was thus stronger than in Experiment 1. Error rate in the critical trial was significantly lower than in pre-critical blank trials, which is consistent with a trend in Experiment 1. Actually, in previous experiments on surprise capture there seems to be a general trend towards lower error rates in the critical trial. As RTs are also almost always elevated in the critical trial, this might reflect a speed-accuracy trade-off. An alternative explanation is that surprise, in itself, induces a more conservative response criterion, which would well make sense from an ecological point of view (i.e., changing from routine processing to controlled processing in view of an unexpected change in the environment). Next, the results of both experiments will be discussed together in the General discussion.

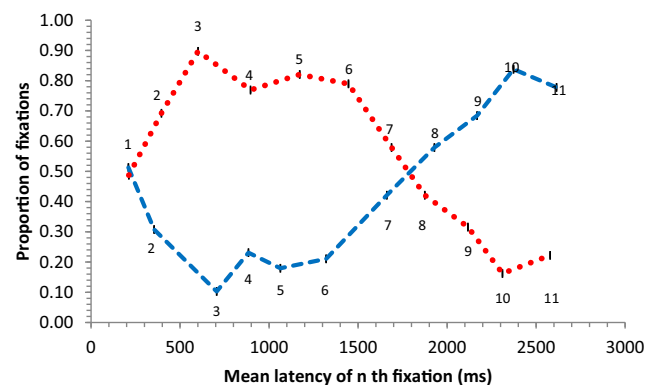


Fig. 3 Plot of the proportion of fixations per ordinal number on the new or old color. X-axis shows the mean latency of the 1st, 2nd, 3rd, etc. fixation. Y-axis shows the proportion of fixations that went either to the new (dotted line) or to the old (dashed line) color. The numbers near the line indicate the ordinal number of the fixation

General discussion

The aim of the present study was to test attentional prioritization of a novel color over a familiar color when the novel color was not at the same time a singleton. For this reason, the display in the critical trial consisted of two equally large sets of colored items such that none of the two sets was a singleton. None of the two colors in the critical trial could be expected to have a higher priority for attentional selection than the other when judged from a saliency perspective.

Results showed that when assessing the first fixation on the familiar color and the first fixation on the novel color for each participant, mean first fixation latency was lower for the novel than for the familiar color. Moreover, analyzing the probability of a fixation on the novel color as a function of fixation sequence shows a significant preference for the novel color between the second and the third fixations (Experiment 1) or to the sixth fixation (Experiment 2). Interestingly, the first fixation was indifferent with respect to feature novelty. Results thus indicate that the novel feature was prioritized for attentional selection over the familiar color, but only after the first fixation.

Effects were somewhat larger in Experiment 2, with chromatic color, than in Experiment 1, with achromatic color. From a discrepancy perspective, the novel chromatic color might be more discrepant than the novel achromatic gray scale. This is because a novel chromatic color (e.g., green) is very dissimilar to anything presented before in the experiment (e.g., red), whereas any achromatic color (e.g., light gray) is in a sense always somewhat similar to another achromatic color (e.g., dark gray). This would predict that Experiment 1 would have yielded stronger effects if black and white had been used instead of dark and light gray.

The critical trial in the present experiments was also the first trial where two colors were presented. One might argue that this resulted in the emergence of two objects (two semi-circles) in the critical trial through grouping by similarity, instead of a single object (a full circle) in the pre-critical trial. The emergence of two objects would be another expectancy discrepant change. The reasons for conducting the experiments in this way are explained below. Importantly, the discrepancy induced by two objects (instead of one) does not prioritize any of them over the other. Thus the emergence of two objects does not challenge the aim of this study to demonstrate that a novel feature is prioritized over a familiar feature.

The most important reason for presenting only one color in each pre-critical trial was not to alert participants to color differences. Gibson and Kelsey (1998) have argued that display-wide features might be incorporated in an attentional set even when not correlated with the target position. It was decided to present only one color in each pre-critical trial to avoid concerns about subtle attentional sets induced by

heterogeneous colors (Burnham, 2007). A secondary reason is that varying colors in the pre-critical trial presumably renders a novel color less discrepant (e.g., Schützwohl, 1998), such that keeping color constant maximizes the novelty manipulation.

Prioritization in the surprise trial occurs rather late: (1) the average novel color fixation latency was 362 ms in Experiment 1 and 397 ms in Experiment 2; (2) the first fixation in the trial was not systematically influenced by the novel feature. The present latencies are very similar to the first fixation latency on a novel-colored singleton in a recent eye-tracking experiment (Horstmann & Herwig, 2015). The present latencies also correspond closely to perceptual gains around 400 ms following the onset of the novel color stimulus (Horstmann, 2002; Horstmann, 2006; Horstmann & Becker, 2008). In addition to the average latency one might ask when the prioritization emerges. The lower and upper limit is given by the latency of the first fixation in the display, which was found to be unbiased, and the latency of the second fixation, which was found to be biased. On this analysis, prioritization emerged between 200 and 375 ms in the present experiments. This is a conservative measure because it ignores saccade dead time (e.g., Ludwig, Mildinhal, & Gilchrist, 2007), which is the time where newly acquired information cannot be used to change the course of an already planned saccade. Taking saccade dead time into account would reduce the estimates by 60–80 ms.

What implications does the late effect of novelty have? Theeuwes (e.g., Theeuwes, 2010) argues that involuntary capture effects can be observed only when early attentional shifts are examined, because top-down control dominates later shifts. This does not apply to the present results. It seems rather that the first fixations are dominated by task-driven variables, whereas novelty gains influence somewhat later.

We adopt a biased competition perspective here where visual objects compete for attentional selection (Bundesen, 1990; Desimone & Duncan, 1995; Schneider, Einhäuser, & Horstmann, 2013), with priority for attentional selection being influenced by multiple sources simultaneously. For the present purpose, these sources are in particular the task and novelty. A task-driven top-down set for target candidate stimuli would be expected to include patch shape, patch color, possible patch positions, rapid common onset, etc. All eight disc-ring compounds should receive rather high top-down priorities in the critical trial as they match shape, positions, manner of appearance, etc., with priorities for the four familiar colored stimuli being somewhat higher because they additionally match color. The four novel items in turn are prioritized for their novelty.

Items are selected in the order of decreasing priority, with inhibition being applied to already visited items to allow search to progress (e.g., Klein & MacInnes, 1999). This suggests that for the first fixation, priorities are very similar for the familiar and the novel items, while for the next two

(Experiment 1) or next five (Experiment 2) fixations, priorities were higher for the novel items.

The most straightforward way to explain this pattern is to assume that the time-courses of task-driven and novelty-driven prioritization are different. In particular, it is possible that task-driven priority is computed faster than novelty-driven priority. Feature-based attention has been assumed to modulate the gain of feature detectors in early visual processing (e.g., Wolfe, 1994; Zelinsky & Bisley, 2015). For example, if task-relevant objects are red, the gain for red detectors is increased in early vision feature maps. Importantly, this happens even before the stimulus appears, giving task-relevant stimuli optimal processing conditions right from the start of the display. This would give the familiar color, which had been task relevant during the pre-critical trials, a head start. Stimuli that are not task relevant (e.g., because their color was not presented before) would thus be at a disadvantage, with novelty overcoming this disadvantage only after a short delay.

A second possible explanation assumes that novelty has the highest priority right from the start but that the first eye movement was not based on priorities but was determined strategically. It seems not implausible that participants use tacit strategies and heuristics to program their first eye-movement before the display appears. After all, the task was designed such that the target does not allow attentional guidance. Thus, priorities were the same for all stimuli within a pre-critical trial display, or at least very similar when assuming random fluctuations through noise. Participants may have solved the decision problem where to start the search by choosing the starting point habitually, for example by always starting at the top, or always starting at the position where they left the previous display. To summarize, assuming that the first eye-movement was not determined by priorities but by strategy would explain why novelty did not affect the first eye-movement.

A saliency perspective (e.g., Theeuwes, 2010, Itti, & Koch, 2000) predicts frequent first fixations on the salient hot-spots at the color borders in the display. According to both psychological and neuro-computational models, saliency is computed early in visual processing and is the direct input to the attention guiding priority map (see Zelinsky & Bisley, 2015, for a recent overview on priority maps). According to saliency models, the color borders should be the most salient regions in the display. Therefore, one might expect attention to be biased towards the stimuli forming that border. In contrast to these predictions there was no evidence for such an effect of saliency, which seems to contradict strong versions of these models.

The absence of evidence for a prioritization of saliency resonates with skeptical theorists who view saliency to be as task dependent as feature-based attention (Ansorge, Horstmann & Scharlau, 2010; Folk et al., 1992). According to this view, saliency can be used to guide attention, but this would depend on explicit or implicit task demands. For

example strong evidence for saliency capture (both covert attention and gaze) in psychological experiments is found when participants search for a salient singleton on another dimension, but not if they search for a particular feature in a different dimension (Bacon & Egeth, 1994). If these or similar task demands are missing as in the present experiments, saliency has little immediate effect.

Awh, Belopolsky, and Theeuwes (2012) criticized the dichotomy between top-down and bottom-up processing, emphasizing that variables such as priming (Maljkovic & Nakayama, 1994) or valence (Della Libera & Chelazzi, 2006) do not fit in either category, yet contribute to the control of attention. We agree with these authors on the general argument. Actually we view novelty as an even better example of why the dichotomy fails, because priming by features of previous targets or habitual attentional responses to stimuli of learned importance are arguably variants of task-driven priorities that had been relevant before. Novelty, on the other hand, is conceptually very distinct from task influences and from bottom-up effects. Novelty is different from task influences as it is not based on a match to task-related attentional control settings, but rather on a mismatch to what has been part of the context so far. Novelty is also not purely bottom-up in that it is not in the stimulus (such as rapid onset, e.g. Jonides, 1981) but is dependent on variable internal conditions, in this case expectations.

To summarize, gaze is biased towards the novel color even if the novel color is not at the same time a singleton in the display. The present results also show that it is the second fixation that is strongly biased towards the novel feature, while the first fixation seems unaffected. This indicates that the prioritization begins not before 200 ms post-stimulus onset, but before 375 ms post-stimulus onset, with 350–400 ms as the average first fixation latency of the novel color. In conclusion, the present study provides support to the discrepancy-prioritization hypothesis that novel or expectancy discrepant stimuli are prioritized for attentional selection.

Acknowledgments This research was conducted in the context of the ZiF-research group “Competition and priority control in mind and brain: New perspectives from task-driven vision” that took place at the Centre for Interdisciplinary Research in Bielefeld. The work was supported by the Cluster of Excellence Cognitive Interaction Technology ‘CITEC’ (EXC 277) at Bielefeld University, which is funded by the German Research Foundation (DFG), and by DFG-Grant HO 3248/2-1 to Gernot Horstmann.

References

- Ansorge, U., Horstmann, G., & Scharlau, I. (2010). Top-down contingent attentional capture during feed-forward visual processing. *Acta Psychologica, 135*, 123–126.

- 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.
- Bacon, W. F., & Egeth, H. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, *55*, 485–496.
- Becker, S. I., & Horstmann, G. (2011). Novelty and saliency in attentional capture by unannounced motion singletons. *Acta Psychologica*, *36*, 290–299.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523.
- Burnham, B. R. (2007). Displaywide visual features associated with a search display's appearance can mediate attentional capture. *Psychonomic Bulletin & Review*, *14*(3), 392–422.
- Della Libera, C., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary rewards. *Psychological Science*, *17*(3), 222–227.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222.
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*, 1827–1837.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.
- Eimer, M., & Kiss, M. (2008). Involuntary attentional capture is determined by task set: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *20*(8), 1423–1433.
- Foerster, R. M., & Schneider, W. X. (2015). Expectation-violations in sensorimotor sequences: Shifting from LTM-based attentional selection to visual search. *Annals of the New York Academy of Sciences*, *1339*, 45–59.
- 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*, 1030–1044.
- Gibson, B. S., & Jiang, Y. (1998). Surprise! An unexpected color singleton does not capture attention in visual search. *Psychological Science*, *9*, 176–182.
- Gibson, B. S., & Kelsey, E. M. (1998). Stimulus-driven attentional capture is contingent on attentional set for display-wide visual features. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 699–706.
- Hays, W. L. (1973). *Statistics for the social sciences*. New York: Holt, Rinehart and Winston.
- Horstmann, G. (2002). Evidence for attentional capture by a surprising color singleton in visual search. *Psychological Science*, *13*, 499–505.
- Horstmann, G. (2005). Attentional capture by an unannounced color singleton depends on expectation discrepancy. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1039–1060.
- Horstmann, G. (2006). Time course of intended and unintended orienting of attention. *Psychological Research Psychologische Forschung*, *70*, 13–25.
- Horstmann, G. (2015). The surprise attention link – A review. *Annals of the New York Academy of Sciences*. doi:10.1111/nyas.12679
- Horstmann, G., & Ansorge, U. (2006). Attentional capture by rare singletons. *Visual Cognition*, *14*, 295–325.
- Horstmann, G., & Becker, S. (2008). Effects of stimulus onset asynchrony and display duration on implicit and explicit measures of attentional capture by a surprising singleton. *Visual Cognition*, *2*(3), 290–306.
- Horstmann, G., & Becker, S. I. (2011). Evidence for goal-independent attentional capture from validity effects with unexpected novel color cues - a response to Burnham (2007). *Psychonomic Bulletin & Review*, *18*, 512–517.
- Horstmann, G. & Herwig, A. (2015). Surprise attracts the eye and binds the gaze. *Psychonomic Bulletin & Review*, in press.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10), 1489–1506.
- Jonides, J. (1981). Voluntary versus automatic control over the mind's eye movement. In J. Long & A. D. Baddeley (Eds.), *Attention & Performance IX* (pp. 187–203). Hillsdale, NJ: Erlbaum.
- Kim, M. S., & Cave, K. R. (1999). Top-down and bottom-up attentional control: On the nature of interference from a salient distractor. *Perception & Psychophysics*, *61*(6), 1009–1023.
- Klein, R. M., & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science*, *10*(4), 346–352.
- Ludwig, C. J., Mildinhal, J. W., & Gilchrist, I. D. (2007). A population coding account for systematic variation in saccadic dead time. *Journal of Neurophysiology*, *97*(1), 795–805.
- Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition*, *22*(6), 657–672.
- Schneider, W. X., Einhäuser, W., & Horstmann, G. (2013). Attentional selection in visual perception, memory and action: A quest for cross-domain integration. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *368*(1628), 20130053.
- Schützwohl, A. (1998). Surprise and schema strength. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1182–1199.
- Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*(2), 77–99.
- Theeuwes, J., deVries, G.-J., & Godijn, R. (2003). Attentional and oculomotor capture with static singletons. *Perception & Psychophysics*, *65*, 735–756.
- Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238.
- Yantis, S., & Egeth, H. (1999). On the distinction between visual salience and stimulus-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(3), 661–676.
- Zelinsky, G. J., & Bisley, J. W. (2015). The what, where, and why of priority maps and their interactions with visual working memory. *Annals of the New York Academy of Sciences*, *1339*(1), 154–164.