

Evidence for goal-independent attentional capture from validity effects with unexpected novel color cues—A response to Burnham (2007)

Gernot Horstmann · Stefanie I. Becker

© Psychonomic Society, Inc. 2011

Abstract Can a stimulus capture attention independent of the observer's goals and intentions? In a recent review, (Burnham, *Psychonomic Bulletin & Review*, 14, 392–422, 2007) argued that there is no convincing evidence that attention capture is ever completely independent of the goals and intentions of the observer. By contrast, surprise capture studies have shown that a color singleton can capture attention on its unannounced first occurrence, if it is new and unexpected, and hence is not part of the intentional set. However, the evidence from surprise capture studies has been criticized on methodological grounds. Here, we tested surprise capture in a new paradigm that avoids previous methodological complications. The results refute the prior criticisms and reinstate surprise capture as prime evidence for goal-independent capture.

Keywords Attention capture · Attention · Surprise

One of the most widespread views in attention research is that some visual properties can capture attention independently of the goals and intentions of the observers. Yet, in a recent review, Burnham (2007) argued that there is still no convincing evidence for goal-independent attention (see also Becker & Horstmann, 2010).

The majority of studies reviewed by Burnham tested goal-independent attention by asking observers to search for a specific target in a visual search array (e.g., a

particular letter in a letter array) while ignoring an irrelevant salient distractor (e.g., a red item among all-green items). To establish goal independence, the distractor was presented at chance level at the location of the target, so that search could not benefit from voluntarily attending to it. Goal-independent attention capture would be inferred if the distractor captured attention despite its task irrelevance.

As Burnham observes, this procedure is vulnerable to side effects of goal-dependent processes. In particular, salient distractors typically signal the beginning of the target display, and thus may be incorporated in goal-dependent attentional control settings as a trigger to initiate search (e.g., Folk, Remington, & Johnston, 1992). Thus, the conclusion that shifts of attention to irrelevant distractors reflect involuntary capture (e.g., Theeuwes, 1994) may be unfounded.

From a broader perspective, the principal weakness of this approach is that salient distractors are presented repeatedly during an experiment. With repeated presentations, it is virtually impossible to ensure that distractor features do not become part of goal-dependent attentional settings. Gibson and Jiang (1998) were the first to note that this problem can be circumvented by testing the distractor at its unannounced first presentation. To date, a number of studies using this surprise capture paradigm have shown that an unexpected distractor can capture attention in a goal-independent manner (Horstmann, 2002, 2005, 2006; Horstmann & Becker, 2008).

Burnham acknowledged that unexpected distractors could reveal truly goal-independent attention, but he raised a methodological concern. For example, in Horstmann (2002, Exp. 3), the task was a demanding visual search for a target letter among 11 or 3 perceptually similar nontarget letters, which resulted in a set-size effect: RTs were longer with 11 than with 3 nontargets. This set-size effect indicates

G. Horstmann (✉)
Department of Psychology, Bielefeld University,
PO-Box 100 131, Bielefeld 33 501, Germany
e-mail: gernot.horstmann@uni-bielefeld.de

S. I. Becker
School of Psychology, University of Queensland,
St Lucia, Queensland, Australia

that the target could be found only by selecting letters in a serial manner in these precritical trials. On the critical trial, a salient color distractor was presented unexpectedly at the target's position. This eliminated the set-size effect, which is the hallmark of attentional capture.

Burnham questioned the latter conclusion. The methodological flaw, according to Burnham, was that the set-size effect was inferred on the basis of the mean performance in the precritical trials. These included the early trials for which there was very little practice, yielding large set-size effects, as well as the late trials, where the set-size effect had possibly already been eliminated by practice. Thus, according to Burnham, it is possible that the reduction in the set-size effect on critical trials was due not to goal-independent attention capture, but to practice during the precritical trials.

Aim of the present experiments

The present experiments aim to address Burnham's critique and to reinstantiate surprise capture as prime evidence for goal-independent attention. To that end, we used a new paradigm in which surprise capture was not inferred from a set-size effect reduction in the critical trial relative to the precritical trials. In fact, set size was kept constant. To assess attention capture at the unannounced first presentation of a salient color, the effects of a valid color cue at the position of the target and of an invalid color cue away from the target were compared. If attention was immediately directed to the cue, the time to find the target should be short with a valid cue but relatively long with an invalid cue, because after directing attention to the invalid cue, a serial search for the target must still be conducted. Thus, attention capture would here be inferred from the sudden emergence of a difference in critical-trial performance, rendering Burnham's practice explanation inapplicable. Furthermore, for the assessment of precritical trial performance, two baselines were used: (1) mean RT, averaged over all precritical trials, and (2) the RT of the single trial immediately preceding the critical trial (cf. Burnham, 2007).

Experiment 1

Participants searched for the target letter H or U in a circular array of 12 letters (e.g., A B C D E F I L P S T). In the precritical trials (1–48), all letters were of the same color (gray), rendering search difficult. In the critical trial (49), one of the letters was unexpectedly presented in a novel color (green). In the postcritical trials (50–72), the novel color was presented under conditions of expected-

ness. Depending on the condition, the new color was presented at the position of the target (valid) or four positions away from the target (invalid).

Method

Participants A total of 23 women and 8 men, with a mean age of 23.7 years ($SD = 3.4$), took part in this experiment. Their compensation was €2.

Apparatus A microcomputer (80486/100 MHz CPU), a keyboard, and a 15-in. CRT monitor were used for stimulus presentation and response registration. The response keys were the adjacent ← and ↓ keys in the lower row of the keyboard.

Stimuli All stimuli were presented against a black background with a viewing distance of 60 cm. The fixation cross subtended 0.2° of visual angle. The 12 letters ($1.0^\circ \times 0.7^\circ$) consisted of vertical and horizontal lines only ("block letters") and were presented equidistantly along the outline of an imaginary circle (diameter: 6.4°).

The targets were the letters H and U. The 11 nontarget letters were A, B, C, D, E, F, I, L, P, S, and T. The letter colors were light gray or green, depending on the condition. These colors were similar in luminance, although not formally matched.

Procedure A fixation cross was presented at the screen's center throughout a trial. Letters were presented for a maximum of 3,000 ms; a response within this period terminated presentation and started the next trial. The intertrial interval was 1,600 ms.

On each trial, one target was presented among the 11 nontargets. The target identity and letter positions were randomized. The participants pressed the ← or the ↓ key, depending on target letter identity (H or U); the two targets appeared equally often. Response speed was emphasized, but participants were instructed to be accurate as well; false responses were immediately followed by a 100-ms tone.

After 12 familiarization trials, there was a single experimental block of 72 trials. Trials 1–48 constituted the precritical trials, Trial 49 the critical trial, and Trials 50–72 the postcritical trials. The transitions between trials were all the same.

In the precritical trials, all letters were light gray. In the critical and postcritical trials, either the target letter (valid cue condition) or a distractor letter four positions away from the target (invalid cue condition) was colored green. Validity was varied between participants: If the cue was valid (or invalid) in the critical trial, it was also valid (or invalid) in the postcritical trials. Participants were randomly

assigned to the valid ($N = 15$) or the invalid ($N = 16$) condition.

Results

Correct RTs between 200 and 3,000 ms were analyzed; excluding long RTs (e.g., above 2,000 ms) did not change the results pattern meaningfully. In the invalid cue condition, six RTs from critical trials were lost due to errors (four trials) or to failures to respond (two trials). Figure 1 displays the mean correct RTs.

Precritical trials versus critical trial A first ANOVA compared RTs for the first presentation of the new color in the critical trial and for the precritical trials. The 2 (trial: precritical vs. critical) \times 2 (validity: valid vs. invalid cue) ANOVA revealed a main effect of validity, $F(1, 23) = 13.55$, $p = .001$, which was modified by the predicted Trial \times Validity interaction, $F(1, 23) = 10.27$, $p = .004$. The main effect of block was not significant, $F(1, 23) = 2.87$, $p = .104$. The interaction was due to a validity effect in the critical trial, $t(23) = 3.24$, $p = .008$, whereas validly and invalidly cued groups did not differ in their performance in the precritical trials, $t < 1$. The effect size of the interaction was large, Cohen's $d = 1.21$. A corresponding analysis of errors (see Table 1) revealed no indication of a speed-accuracy trade-off [for trial, $F(1, 29) = 2.27$, $p = .143$; for validity, $F(1, 29) = 4.65$, $p = .040$; for Trial \times Validity, $F(1, 29) = 4.60$, $p = .040$].

These results did not depend on the aggregation of RTs over the first 48 trials: When only Trials 48 and 49 were compared (24 valid responses), the Trial \times Validity interaction was significant as well, $F(1, 22) = 9.70$, $p =$

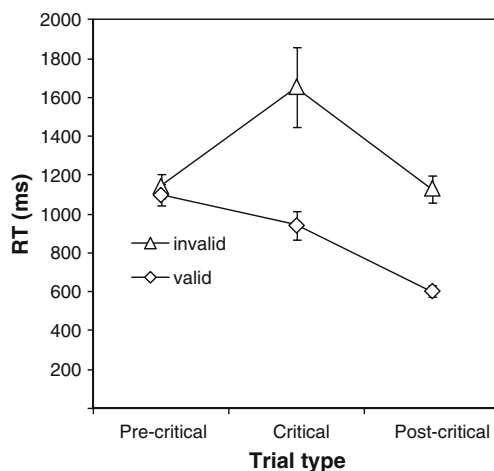


Fig. 1 Mean RTs in the precritical trials, the critical trial, and the postcritical trials for the valid and invalid cue conditions in Experiment 1

Table 1 Error percentages

	Experiment 1		Experiment 2			
	Valid	Invalid	Repetition		Change	
			Valid	Invalid	Valid	Invalid
Trial Type (1–48)	3.8	3.5	4.9	6.8	5.3	7.1
Last precritical (48)	6.7	0.0	0.0	0.0	13.3	0.0
Critical (49)	0.0	25.0	6.7	0.0	6.7	26.3
Postcritical (50–72)	5.6	1.6	6.4	7.4	3.9	5.5

.005 [trial, $F(1, 22) = 3.25$, $p = .085$; validity, $F(1, 22) = 2.61$, $p = .121$]. Critical-trial RTs were shorter with a valid cue than with an invalid cue (949 vs. 1,652 ms), $t(22) = 3.19$, $p = .008$, whereas Trial 48 RTs did not differ (1,138 vs. 945 ms), $t(22) < 1$.

Critical versus postcritical trials To assess whether cuing effects are modulated by the observers' expectations, we computed a corresponding ANOVA over the RTs of the critical trial versus the postcritical trials. The results showed main effects of trial, $F(1, 23) = 27.40$, $p < .001$, and validity, $F(1, 23) = 28.83$, $p < .001$, but no interaction, $F(1, 29) = 1.22$, $p = .280$. Thus, no differences were found between the size of the cuing effect for the unexpected and the expected cues. A corresponding analysis of errors revealed a pattern similar to the one found before [trial, $F(1, 29) = 2.33$, $p = .138$; validity, $F(1, 29) = 3.30$, $p = .080$; Trial \times Validity, $F(1, 29) = 6.12$, $p = .019$].

Discussion

Experiment 1 showed significant validity effects indicative of an attentional shift to the new color. Because the new color was tested on its unannounced first presentation, side effects of goal-dependent processing or voluntary strategies can be excluded as explanations. In addition, the results cannot be attributed to practice effects (Burnham, 2007), rendering them prime evidence for goal-independent attention.

The cuing effect was as large in the critical trial as in the postcritical trials. However, RTs were about 400 ms longer in the critical than in the corresponding postcritical trials. Such an RT inflation is regularly observed in the surprise capture paradigm and can be explained by nonspatial interference induced by the surprising change itself, which is additive to the spatially specific attentional effects (cf. Horstmann, 2005). This additive nonspatial interference

also explains why the benefits of valid cues appear relatively small, whereas the costs induced by invalid cues are relatively large when compared to costs in the precritical trials.

Experiment 2

Experiment 2 extended on the results of **Experiment 1** by testing surprise capture by an unexpected color (green) when two colors (red and gray) had been present in the precritical trials. The purpose of this change was twofold. First, previous demonstrations of surprise capture had used only a single homogeneous color in all precritical trials. Thus, it was of interest to see whether strict homogeneity is necessary for observing involuntary attention capture by the unannounced color cue. Second, Goolsby, Grabowecky, and Suzuki (2005) found that detection of a color singleton was faster when all items in the previous, target-absent trial had been presented in the distractor color than when all items had been presented in the target color or a neutral color (distractor preview effect). To test whether surprise capture depends on a preview of same-color distractors, we included conditions in which the distractor color on the critical trial was preceded by two trials with the same versus a different color.

Methods

Participants A total of 41 women and 22 men, with a mean age of 23.03 years ($SD = 3.31$), took part in this experiment.

Apparatus, stimuli, and procedure These were the same as before, with the following exceptions. On half of the 48 precritical trials, all stimuli were presented in gray, and on the other half, all stimuli were presented in red. In order to test the distractor preview hypothesis, the order of presentation was randomized for the first 44 trials and then controlled in Trials 45–48, where two fixed sequences (AABB and BBAA) were used. On the critical trial, a green singleton was presented among gray nonsingleton letters. Participants were randomly assigned to the valid ($N = 30$) or the invalid ($N = 33$) condition.

Results

Correct RTs in the range between 200–3,000 ms were analyzed (Fig. 2). Seven critical-trial RTs were lost due to errors committed on that trial: two in the valid and five in the invalid cue condition.

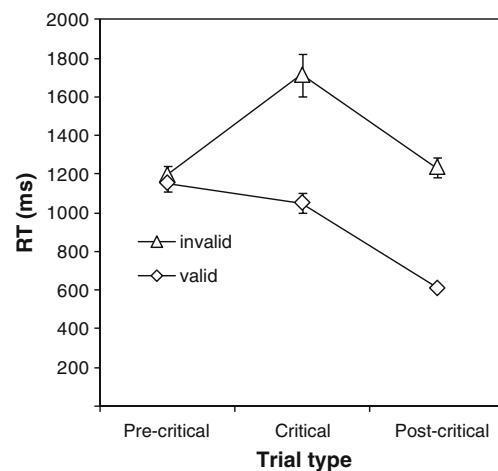


Fig. 2 Mean RTs in the precritical trials, the critical trial, and the postcritical trials for the valid and invalid cue conditions in **Experiment 2**

Precritical trials versus critical trial The 2 (trial: precritical vs. critical) \times 2 (validity: valid vs. invalid) \times 2 (sequence: repeated vs. nonrepeated distractor color) ANOVA revealed main effects of validity, $F(1, 52) = 18.45, p < .001$, and trial, $F(1, 52) = 15.23, p < .001$, as well as the predicted Trial \times Validity interaction, $F(1, 52) = 33.60, p < .001$ (all other $F_s < 1.01$). The interaction (Cohen's $d = 1.55$) was due to large differences between the valid and invalid groups in the critical trial, $t(54) = 5.41, p < .001$, but not in the precritical trials, $t(54) < 1$. An error analysis (Table 1) revealed no significant effects, although the main effect and interactions involving sequence approached significance ($F_s = 2.69$ to $2.91, p_s = .093$ to $.106$), which is apparently due to an error peak in the critical trial of the invalid nonrepeated-color condition (26%).

The results were similar when only Trial 48 was used as a baseline. The ANOVA revealed the predicted Trial \times Validity interaction, $F(1, 50) = 33.76, p < .001$, as well as main effects of trial, $F(1, 50) = 16.63, p < .001$, and validity, $F(1, 50) = 9.98, p = .003$. The Trial \times Sequence interaction was significant, $F(1, 50) = 5.69, p = .021$ (all other $F_s < 1$). In the critical trial, responses were faster with valid than with invalid cues (1,029 vs. 1,713 ms), $t(52) = 5.57, p < .001$, but not in Trial 48 (1,132 vs. 1,039 ms), $t(52) < 1$. The significant Trial \times Sequence interaction revealed more nonspatial interference in the critical trial when the distractor color was repeated from the previous trial than when it changed (441 vs. 115 ms).

Critical versus postcritical trials The ANOVA comparing RTs across the critical and postcritical trials revealed main effects of block, $F(1, 52) = 67.98, p < .001$, and validity, $F(1, 52) = 67.85, p < .001$, but the other effects were not significant, $F_s < 1$. A corresponding analysis of the errors revealed two (almost) significant interactions [Trial \times Sequence, $F(1, 59) = 4.31, p = .042$; Trial \times Sequence \times

Validity, $F(1, 59) = 3.03, p = .087$; all other $F_s \leq 2.60, p_s \geq .112$], reflecting mainly the high incidence of errors in the invalid nonrepeated-color condition.

Discussion

Experiment 2 extended the results from **Experiment 1** by showing that surprise capture also occurs when the precritical trials contain two different colors, regardless of whether the distractors on the critical trial match the color of the items presented in the preceding trial. This indicates that surprise capture does not require highly repetitive or strictly uniform precritical presentations. There was evidence that repeating the distractor color leads to longer RTs in the critical trial. Thus, although a distractor preview effect was not found, it seems that with a more familiar context (repeated distractors), general interference is higher.

General discussion

Evidence for goal-independent attention from the surprise capture paradigm is important because of the persistent problem of excluding nonobvious top-down strategies with repeated, and hence expected, cues. Burnham (2007), however, rejected previous results from the surprise capture paradigm that are relevant to the issue of goal-independent attention, because of possible methodological problems caused by differential training effects. The present experiments, however, are not subject to the same criticism: First, attention capture was inferred from a difference in the critical trial between groups with a valid versus an invalid cue, which could hardly be explained by practice effects. Second, an extra check of the last precritical trial ensured that the pattern observed in the critical trial was not already present in the immediately preceding trials.

The present results also put in question Burnham's practice explanation of the surprise capture in previous studies, and indicate that the prior results using the set-size method were probably due not to practice, but to capture. This interpretation merges with other evidence that the practice explanation is invalid. First, the practice explanation seems to be inconsistent with the fact that performance gains are observed only with a novel, not with a familiar, cue color (e.g., Horstmann, 2005, Exp. 5). Second, surprise capture has also been demonstrated in accuracy tasks in which practice effects were taken into account statistically (e.g., Horstmann, 2002, Exps. 1 and 2; Horstmann, 2006; Horstmann & Becker, 2008).

Thus, the present results confirm previous tests of surprise capture using a new paradigm and analysis, which are not subject to the criticism that has been raised against the previous tests. **Experiment 2**, moreover, demonstrated that the new and unexpected color (green) captured attention regardless of whether the distractor color (red) matched the colors of all items in the preceding, precritical trial, and despite the fact that observers had been exposed to two different colors in the precritical trials. This indicates that surprise capture does not depend on presenting same-color distractors in the precritical trials. A second implication of **Experiment 2** is that exposure to multiple colors in the precritical trials does not lead to a generic top-down setting to ignore color variation.

How can surprise capture be explained? According to the saliency capture hypothesis, salient stimuli—that is, those that deviate in a visual feature from their surroundings—capture attention in a bottom-up fashion (see, e.g., Itti & Koch, 2000; Theeuwes, 1994), and the present results may also seem consistent with saliency capture. However, there is evidence that saliency alone is not sufficient for capture by unexpected stimuli. Other experiments have shown that one critical precondition for surprise capture is that the feature itself be novel: When the salient feature in the critical trial (e.g., a red singleton among all green items) was rendered familiar—for example, because all-red and all-green displays had been randomly mixed in the precritical trials—the set-size effect in the critical trial was not reduced, reflecting inefficient search (Horstmann, 2005, Exp. 5). This result pattern is consistent with an expectancy discrepancy hypothesis, stating that stimuli capture attention in a goal-independent manner if they violate feature expectations.

To conclude, by ruling out alternative explanations for surprise capture and applying a new method to measure it, surprise capture has been reinstated as currently the best evidence for attentional capture in the absence of a relevant attentional set.

References

- Becker, S. I., & Horstmann, G. (2010). Novelty and saliency in attentional capture by unannounced motion singletons. *Acta Psychologica*. doi:10.1016/j.actpsy.2010.12.002
- Burnham, B. R. (2007). Displaywide visual features associated with a search display's appearance can mediate attentional capture. *Psychonomic Bulletin & Review*, *14*, 392–422.
- 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.

- Goolsby, B. A., Grabowecky, M., & Suzuki, S. (2005). Adaptive modulation of color salience contingent upon global form coding and task relevance. *Vision Research*, *45*, 901–930.
- 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 allocation of attention. *Psychological Research*, *70*, 13–25.
- Horstmann, G., & Becker, S. I. (2008). Effects of stimulus onset asynchrony and display duration on implicit and explicit measures of attentional capture by a surprising singleton. *Visual Cognition*, *16*, 290–306.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506.
- Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 799–806.