Journal of Experimental Psychology:
Human Perception and Performance
2005, Vol. 31, No. 5, 1039–1060

0096-1523/05/$12.00 DOI: 10.1037/0096-1523.31.5.1039

Attentional Capture by an Unannounced Color Singleton Depends on Expectation Discrepancy

Gernot Horstmann
Bielefeld University

Eight experiments examined the conditions under which a color singleton that is presented for the 1st time without prior announcement captures attention. The main hypothesis is that an unannounced singleton captures attention to the extent that it deviates from expectations. This hypothesis was tested within a visual-search paradigm in which set-size effects were used to infer attentional capture. The results showed that attentional capture by an unannounced color singleton was due to a mismatch with expectations concerning the color of the object and not due to its being a singleton. Thus, the results imply that theories of attention have to consider expectation discrepancy as a determinant of attention shifts.

Keywords: attention capture, involuntary orienting, visual search, expectation discrepancy, surprise

Expected and unexpected events differ fundamentally in their nature. Expected events can, in principle, be attended to and further processed according to their relevance for current goals or intentions as they occur. The relevance of unexpected events, however, is unknown before their appearance; yet it has to be determined to safeguard adaptive behavior. This presupposes that focused attention is directed to the unexpected event. The present research is concerned with the conditions of attentional capture by an unexpected color singleton (Gibson & Jiang, 1998; Horstmann, 2002, 2004, in press; Wilcocks, 1928). That is, it examines the conditions under which an unexpected salient event quickly summons attention.

The main hypothesis pursued is that the unannounced—and hence unexpected—presentation of a new color captures attention to the extent that it is discrepant from expectations (e.g., Horstmann, 2002; Wilcocks, 1928). In the reported experiments, this hypothesis is tested and refined via a visual-search paradigm (e.g., Jonides, 1981; Theeuwes, 1992; Treisman & Gelade, 1980; Yantis & Egeth, 1999): Participants searched for target letters (e.g., *U* or *H*) within an array of distractor letters and indicated the target letter's identity with a key press. Reaction time (RT) was the dependent variable. The number of distractor letters (set size) was varied to test search efficiency: If RT was longer with 12 letters than with 4 letters, search was assumed to have proceeded nonefficiently, whereas RTs that were approximately the same with 12 and 4 letters revealed efficient search, indicating attentional capture to the target's location. To test whether the expectancy-discrepant color captured attention, I presented it at the position of the target only; that is, it was a color singleton (e.g., Nothdurft, 2000; Theeuwes, 1992, 1994; Yantis & Egeth, 1999).

Contingent Capture

Color singletons can be used to guide visual search quite effectively (cf. Krummenacher, Müller, & Heller, 2002; Nothdurft, 2000; Wolfe, 1994; Yantis & Egeth, 1999): When observers know that the singleton's position predicts the target's position, search performance (as indicated by RT) is independent of set size, revealing efficient search. Moreover, the phenomenal experience is one of immediate awareness of the singleton, as if attention had been attracted or captured by the color singleton. However, this does not necessarily mean that attentional capture is completely independent of intentions. On the contrary, if the singleton's position is known to be nonpredictive of the target's position, singletons can often be completely ignored (e.g., Yantis & Egeth, 1999; but see Theeuwes, 1992, 1994; Theeuwes & Burger, 1998; Theeuwes & Godijn, 2001). This pattern of results indicates that attentional capture critically depends on an attentional set: It occurs in the presence of the intention to attend to the singleton but not in the absence of such an intention and is thus contingent on the attentional set (or attentional control settings; cf. Folk, Remington, & Johnston, 1992, 1993). According to the contingent-capture hypothesis, the process underlying attentional capture is a hybrid of controlled and automatic processes: The settings of an attentional control system are determined "offline" (i.e., prior to the presentation of the stimulus) depending on the respective intention; thereafter, stimuli that match the attentional set capture attention without the need of intervening intentional processes.

Several variants of attentional control settings have been described. First, attentional control can be set to color, brightness, size, or motion (e.g., Yantis & Egeth, 1999) if the target is defined by one of these features. Initially, it was suggested that attentional

---

This research was partially supported by Deutsche Forschungsgemeinschaft Grant HO 3248/1 and by University Grant 27102030. I am indebted to Wilhelm Banmann, Christine Broermann, Sabine Dlugosch, Kathrin Holtmann, and Lydia Wittenbreder for conducting the experiments; Lily-Maria Silny for her assistance in manuscript preparation; and Ulrich Ansorge, Stefanie Becker, Bradley Gibson, Wulf-Uwe Meyer, Ingrid Scharlau, Daniel Simons, and Jan Theeuwes for their helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Gernot Horstmann, Department of Psychology, Bielefeld University, PO Box 100 131, Bielefeld D-33 501, Germany. E-mail: gernot.horstmann@uni-bielefeld.de

control settings discriminate between static features, such as color or brightness, and dynamic features, such as motion or flicker, but not within the respective categories (Folk et al., 1992). Recent evidence, however, suggests that the probability of capture depends on physical similarity between stimulus and attentional set even within dimensions (Ansorge & Heumann, 2003; Folk & Remington, 1998). Second, apart from this set for target features there may be a "set for displaywide features": Gibson and Kelsey (1998) argued that attention should often be set to features that discriminate the search display from other displays within a given trial, and they provided evidence supporting this hypothesis. A third set type was described by Bacon and Egeth (1994). If a singleton discriminates between targets and distractors, performers may adopt a singleton-search strategy, which entails searching for any singleton present in the search display. In summary, attentional control may be set to the target's features, to displaywide features, or to any singleton present in the display.

Surprise Capture

That attentional capture by color singletons can be contingent on the intentions of the observer does not imply that this is always the case. Indeed, there is evidence that a color singleton can capture attention in the complete absence of any intention to attend to the singleton—that is, when the singleton is presented for the first time without prior announcement and following a number of trials without a singleton (e.g., Horstmann, 2002, 2004, in press). Clearly, the presentation of a singleton under such circumstances precludes any intentions toward the singleton (cf. Gibson & Jiang, 1998). Additionally, in Horstmann's (2002) experiments, color was not a discriminative cue related to any aspect of the task to prevent a set for color as a displaywide feature (Gibson & Kelsey, 1998): All stimuli preceding the critical trial (including the fixation cross and the messages that informed the participants about the beginning and the end of the practice trials) were presented in the color of the stimuli in the homogeneous trials.

Two strategies have been used to test attentional capture by an unannounced singleton. In the first paradigm, participants performed a search task on very briefly presented letters, with accuracy as the dependent variable (Gibson & Jiang, 1998). For example, in Horstmann (2002, Experiment 1), the positions of 12 letters were preceded by color cues of the same color with a stimulus onset asynchrony (SOA) of 500 ms, until, in the critical trial, the cue at the position of the target letter had a new color. In this critical trial, accuracy was significantly improved, indicating that the color cue captured attention to its location, thereby facilitating the identification of the target letter. These results were replicated and extended (Horstmann, 2004, in press); they also indicated that the time lag of surprise capture (relative to contingent capture) was about 300–400 ms (Horstmann, in press) and that both benefits from predictive singletons and costs from nonpredictive singletons accrued (Horstmann, 2004).

The second approach was instantiated in Horstmann (2002, Experiment 3), where the presentation time of 4 or 12 colored letters was not restricted and RT was the dependent variable. In the precritical trials, in which all letters were colored equally, RT was strongly affected by the number of presented letters (set size), indicating that finding the target required serial search; in contrast, there was only a weak set-size effect in the critical trial, in which the target letter appeared in a new color. This pattern of results would be expected on the hypothesis that the new color captured spatial attention.

These results are consistent with a hypothesized surprise–attention interface (Darwin, 1872; Horstmann, 2002; Meyer, 1988; Meyer, Niepel, Rudolph, & Schützwohl, 1991; Meyer, Reisenzein, & Schützwohl, 1997; Niepel, Rudolph, Schützwohl, & Meyer, 1994; Prinz, 1983, 1990; Schützwohl, 1998; Selz, 1922; Wilcocks, 1928). In short, events that deviate from expectations attract attention (besides other surprise-related changes, e.g., an interruption of current information processing and action). That is, the unannounced color singleton captures attention because it is discrepant from expectancies concerning the color of the stimuli in the critical trial, with surprise capture being a direct consequence of the color being expectancy discrepant. Such a shift of attention to the surprise-eliciting stimulus is thought to be adaptive in that it provides access to limited processing and response-production capabilities.

Differences Between Contingent and Surprise Capture

Contingent and surprise capture differ in their conditions. Whereas contingent capture depends on a stimulus's *match with intentions*, surprise capture depends on a stimulus's *mismatch with expectations*. However, they should not be seen as being in conflict. Rather, they can be conceived of as complementary mechanisms, one concerned with expected events for which intentions for action can be formed in advance, and the other dealing with unexpected events whose relation to intentions and goals is initially unknown.

Contingent capture and surprise capture also differ in their temporal characteristics. Several experiments presented the search display only briefly (less than 100 ms) and assessed accuracy (Horstmann, 2002, 2004, in press). With this procedure, it is possible to compare the relative speed of contingent capture and surprise capture. For example, Horstmann (in press, Experiment 1) varied the SOA between the singleton and the search display and found that with a 200-ms SOA, performance was not improved by the unannounced first-time presentation of a singleton at the target's position after 48 homogeneous trials. In contrast, with 400-ms or 600-ms SOAs, performance was improved. Experiment 2 of the same study, in which a more extended range of SOAs was realized, replicated this result with no significant benefits with 0-ms, 100-ms, or 200-ms SOAs but with significant benefits with 400-ms, 500-ms, and 600-ms SOAs. These results are consistent with previous studies that did not find performance to have improved with an SOA of 0 ms (Gibson & Jiang, 1998, Experiment 1; Horstmann, 2002, Experiment 2) but did find improvement with an SOA of 500 ms (Horstmann, 2002, Experiment 1). Horstmann (2004) obtained a similar time course for the performance costs that incurred with a nonpredictive singleton (i.e., a singleton that appears at a distance from the target's position at the position of a distractor). In contrast, when observers know that the singleton is presented in each trial at the same position as the target of the search, performance is scarcely influenced by SOA; that is, performance is nearly as good with simultaneous presentation (0-ms SOA) as it is with positive SOAs (Horstmann, 2002, in press), even when the singleton is presented unpredictably in only in 4% of the trials (Horstmann & Ansorge, in press).

Another discriminating temporal aspect is the time course of surprise capture and contingent capture at repeated presentations. Surprise capture can be elicited only once (or a very limited number of times) with the same type of event, because surprise is elicited by expectancy-discrepant events, and events are anticipated (expected) with repeated presentations (e.g., Horstmann & Ansorge, in press; Meyer et al., 1997; Schützwohl, 1998). Note that expectation is not equated with low probability in the present context, because low-probability events are often expectation congruent (Gibson & Peterson, 2001; Horstmann, in press; Horstmann & Ansorge, in press). For example, there may be only a 10% chance that a traffic light will change from green to red when the driver approaches it—yet this is certainly an expected event. In contrast, expectation is conceptually closely related to the concepts of anticipation, hypothesis, and belief. *Expectation discrepant* thus means contrary to anticipations, hypotheses, or beliefs (cf. Meyer & Niepel, 1994). Whereas surprise capture can be elicited by the same type of event only a limited number of times, contingent capture can be elicited as often as the task requires the participant to attend to a feature.

The differences in temporal characteristics suggest different underlying mechanisms for contingent capture and surprise capture. In particular, contingent capture is assumed to result from a *match* between the attentional control setting and characteristics of the stimulus. In contrast, surprise capture is assumed to result from a *mismatch* between aspects of a schematic expectation (e.g., that the presented stimuli are red) and corresponding aspects of the discrepant stimulus (e.g., that one stimulus is green). The expectation is termed *schematic* for three reasons. First, the concept of expectation (as used here) focuses on its content rather than on judgments of probability. Second, the idea of a schema has been used by Neisser (e.g., 1976, 1979) and others to emphasize the role of anticipation and expectation in selective attention and conscious perception, which is also assumed to be important to explain surprise capture. In particular, it is assumed that expectations are constant concomitants of cognitive activity and that they are generated continuously, without intention, without interference with other cognitive processes, and without being necessarily conscious. That is, expectations are assumed to result from processes that are automatic in the traditional notion (e.g., Posner & Snyder, 1975).

Third, schema theories (e.g., G. Mandler, 1985; J. M. Mandler, 1984; Rumelhart, 1984; Rumelhart & Ortony, 1977; Rumelhart, Smolensky, McClelland, & Hinton, 1986; Schützwohl, 1998) emphasize the importance of knowledge in perception and cognition. In particular, preceding and concurrent input suggest knowledge structures as initial candidates for activated schemas, and once a reasonable fit has been established between the schema's components and the input, the schema is instantiated as a reasonable hypothesis that accounts for the input. Highly conflicting information on whether a given schema accounts for a given input constitutes a schema discrepancy. The input to this process stems from various sources, including basic features, objects, relations among objects, concepts, and so forth. Only schema discrepancies that stem from preattentive processes (i.e., simple features, e.g., color or form) can produce surprise capture in a narrower sense, whereas schema discrepancies that stem from attentive processes (e.g., recognized objects) only cause spatial attention to be located at the surprising stimulus longer or more frequently.

Last, contingent and surprise capture differ in function. Contingent capture allows for the quick selection of task-relevant features and thereby allows one to ignore features that are known to be task irrelevant or distracting. In contrast, the function of surprise capture is to make expectancy-discrepant—and thus highly informative—stimuli available for further processing.

Inattentional Blindness and Other Failures of Unexpected Events to Capture Attention

A number of results from the literature might seem to contradict the surprise-capture hypothesis. In the first experiment in which an unexpected singleton was presented in a search task, Gibson and Jiang (1998) did not find evidence for attentional capture for a surprise singleton. The authors presented eight letters briefly, with the participants' task being to search for *H* and *U*. All letters had the same color in the precritical trials. In the critical trial and the postcritical trials of Experiment 1, the target had a color different from that of the remaining letters. Gibson and Jiang found performance in the first unannounced presentation of the singleton target in the critical trial to be no better than in the precritical trials but significantly worse than in the postcritical trials. In Experiment 2, the precritical trials corresponded to those of Experiment 1; in the critical and in the postcritical trials, however, a distractor letter was a singleton, whereas the target had the same color as the remaining letters. If the singleton distractor captured attention, interference with the identification of the target letter would be expected, given the very limited presentation time of the letters. In contrast to this hypothesis, performance turned out not to be significantly reduced in the critical trial. Horstmann (2002, Experiment 2; see also Horstmann, 2004, in press) replicated the absence of attentional capture for a singleton target under conditions similar to those used by Gibson and Jiang (1998). However, using a positive SOA between the onset of the singleton and the onset of the letters, Horstmann (2002) found clear evidence of attention capture in the critical trial, a result that was subsequently replicated and extended by a systematic variation of SOA. The failure to find evidence for surprise capture with zero or small positive SOAs, which are normally sufficient to measure attention to expected singletons, is probably due to the fact that the latency of surprise capture lags behind attention to an expected singleton by roughly 300 ms (see also Horstmann, 2004, in press).

A further source of apparent evidence against the surprise-capture hypothesis—inattentional blindness (IB)—is more difficult to review, not least because of the large number of experiments that have been conducted, which in some cases were also exploratory in nature (Mack & Rock, 1998). However, it is important to note at first that IB and surprise capture are not in conflict as phenomena. In contrast, experiments on surprise often reveal rates of IB similar to those found in experiments directly concerned with IB. For example, Meyer et al. (1991; see also Horstmann, 2004, in press) reported that about 20% of the participants did not notice the unexpected stimulus change in the critical trial, which is similar to the IB rates reported by Mack and Rock (1998). That is, in both types of experiments, the majority of participants were not blind to the stimulus change, whereas a minority were.

Second, in some of the most striking demonstrations of IB, a crucial factor appears to be the similarity of the unexpected stimulus to the to-be-attended-to stimuli, on the one hand, and the to-be-ignored stimuli, on the other hand. In particular, Most et al. (2001) showed that IB was linearly related to the unexpected event's similarity to the to-be-ignored stimuli and inversely related to the similarity to the to-be-attended-to stimuli. For example, if black items are to be attended to and white items are to be ignored, a dark gray item is noticed with a higher probability than a light gray item. This is consistent with a contingent capture account of involuntary orienting, which states that attentional capture by an item that is not searched for should be a function of its similarity to the currently used attentional set (e.g., Folk et al., 1992).

Third, in some cases, the discrepancy of the stimulus from the current set of schemas pertains to a semantic aspect: The notorious gorilla among white- and black-dressed basketball players in Simons and Chabris's (1999) experiment, which was not noticed by many observers who attended to the white-dressed players, was completely expectancy congruent with regard to its color (black), way of movement (walking), and gross appearance (primate with upright posture); it was expectancy incongruent, though, for its identity. As stated above, it is quite probable that an expectancy discrepancy can be preattentively detected only for rather simple features (e.g., Treisman & Gelade, 1980) but not for conjunctions of features or conceptual attributes implied by the stimuli.

Fourth, many of the experiments discussed in the book on IB by Mack and Rock (1998) used display durations that were probably too short for surprise capture. In particular, Mack and Rock's standard procedure entailed that the stimuli were displayed for 200 ms and pattern masked afterward. As indicated before, surprise capture appears to lag about 300 ms behind contingent capture, and a display duration sufficient for intentional attention shifts may not be sufficient for surprise capture.

Last, some of the displays might have induced rather nonspecific expectations with regard to the features that happened to characterize the target in the critical trial. For example, in Experiment 3 of Most et al.'s (2001) study, many participants who attended to white circles and ignored black circles did not notice a red plus sign that was visible for 30 s. Although I do not want to suggest this to be the sole factor responsible for the impressively high rate of IB, the black and white circles had already instantiated two colors, and prior research has shown that variability in the critical feature during the precritical trials tends to reduce the surprise response (Schützwohl, 1998).

Another Account of Surprise Capture: Singleton Capture

Some models of visual selection propose that salient stimuli can capture attention in their own right. For example, Guided Search 2.0 (Wolfe, 1994) proposes that intention-driven and data-driven processes independently contribute to the selection of stimuli, implying that salient stimuli can capture attention even in the absence of top-down activation. Similar positions have been adopted by other authors (e.g., Itti & Koch, 2000; Kim & Cave, 1999; Koch & Ullman, 1985; Theeuwes, 1992, 1994; Theeuwes & Godijn, 2001). For example, Theeuwes and Godijn (2001) assumed that singletons capture attention in an involuntary and unavoidable fashion. However, if observers know that the singleton is irrelevant to the task (i.e., nonpredictive of the target), they are able to quickly reorient attention, and this reorientation of attention is used to explain the ease with which singletons are ignored under appropriate conditions (e.g., Yantis & Egeth, 1999).

Evidence concerning this position is heavily disputed. That nominally irrelevant singletons modify performance in a way that could be predicted by a *singleton-capture* hypothesis turned out to be true in a number of studies (e.g., Bacon & Egeth, 1994; Kim & Cave, 1999; Theeuwes, 1992, 1994; Theeuwes & Burger, 1998; Todd & Kramer, 1994). What is debatable, however, is whether such results reflect attentional capture and, if so, whether it is truly noncontingent on intentions. For example, Kim and Cave (1999) presented two singletons in each trial. Although only the form singleton was task relevant and the color singleton was task irrelevant, the color singleton interfered with task performance. This result, however, may indicate not purely bottom-up-driven capture but rather imperfect intentional control over the deployment of attention, because the task required the participants to search for a singleton (cf. Bacon & Egeth, 1994). Furthermore, Folk and Remington (1998) noted that the small costs (in term of RT or accuracy) that incur with irrelevant singletons need not necessarily indicate visuospatial shifts of attention but can also be explained with nonspatial filtering costs that precede shifts of attention. For example, with two singletons, one defining the target (e.g., form) and the other being task irrelevant (e.g., color), the target-defining singleton has to be discriminated from the task-irrelevant singleton, producing small amounts of performance costs.

The singleton-capture view is relevant in the current context because attentional capture by the unannounced first presentation of a singleton might indicate singleton capture rather than surprise capture. For example, Theeuwes (e.g., Theeuwes & Godijn, 2001; see also Kim & Cave, 1999) suggested that singletons generally capture attention in a purely bottom-up fashion but that observers may reorient attention very quickly if it is desirable to ignore the singleton. Thus, it might be argued that the surprise singleton captures attention precisely because the two conditions for singleton capture are met: (a) The singleton is salient, producing a large amount of bottom-up activation, and (b) intentions to ignore the singleton are absent. A first indication that attentional capture by a surprise singleton may not be explained by a singleton-capture account is that the attentional shift in the surprise trial is relatively slow: As pointed out before, attentional capture appears to be somewhat slower for unexpected than for expected singletons, and the singleton-capture hypothesis cannot explain this difference (Horstmann, 2004). Rather, proponents of the singleton-capture hypothesis have proposed that singleton capture occurs within 50–150 ms after the singleton's presentation (e.g., Theeuwes & Godijn, 2001; see also Kim & Cave, 1999). Notwithstanding this point, the present experiments address the singleton-capture hypothesis directly.

Objectives of the Present Research

According to the present hypothesis, an unexpected singleton captures attention to the degree that it is discrepant from expectations. As in prior experiments, the contents of expectations are manipulated via the precritical trials in the present experiments. For example, if all items presented in the precritical trials are green, it is assumed that the expectation concerning the critical trial is that the items in that trial will also be green. If one or more red items are presented in the critical trial, these items will also be expectancy discrepant and draw attention to themselves. In contrast, if some of the items in the precritical trials are green and some are red, a red item in the critical trial will be expectation congruent.

Second, as already proposed by Selz (1922), the detection of an expectancy discrepancy is assumed to be an automatic process that does not need an intention or conscious testing of the expectation. There is plenty of evidence from previous experimental work on surprise that this assumption holds, as participants were not required to test or monitor expectations but were distracted by a stimulus change (e.g., Horstmann, 2001; Meyer et al., 1991; Niepel et al., 1994; Schützwohl, 1998).

Third, whether a discrepancy can be detected preattentively (which is a precondition for attentional capture) depends on whether the discrepant stimulus aspect is preattentively available. Of course, the detection of an expectancy discrepancy often depends on attending to the stimulus first, in particular for stimuli that are expectancy discrepant with respect to their specific combination of basic perceptual features (e.g., Treisman & Gelade, 1980). This is because the combination of basic features as well as their precise spatial relations are available only through attention. In contrast, basic perceptual features, such as color, motion, or orientation, are available prior to attention. Thus, an expectancy discrepancy of those features can also be—at least in principle—detected preattentively.[1]

Fourth, it is assumed that, following the discrepancy detection, attention is involuntarily directed to the discrepant stimulus's features and thereby to the discrepant stimulus's location. An involuntary shift of attention is evident from the RT costs revealed in the critical trials of many experiments in which the physical appearance of a task-irrelevant experimental stimulus was different from that in the precritical trials (e.g., Meyer et al., 1991; Schützwohl, 1998). Of course, these experiments did not demonstrate spatial shifts but showed selective attention in the sense of a selection of the discrepant stimulus for further processing at the cost of a selection of the target stimulus of the RT task.

I have previously described the most straightforward examples of an expectancy-discrepant event (the first presentation of a new color after consistent presentations of a different color) and of an expectancy-congruent event (the presentation of a color in the critical trial that has been already presented in the precritical trials). As already reported, attentional capture by the first presentation of a new color after consistent presentations of another color has been tested and repeatedly found in several experiments (Horstmann, 2002, 2004, in press). The aim of the present experiments is to specify more precisely the conditions of surprise capture. To this end, the present experiments vary the relation between the expectancy and the unannounced singleton in the critical trial. More precisely, the critical trial is essentially the same in all experiments, that is, a color singleton in an otherwise color-homogeneous display. To clarify the conditions of surprise capture, I varied the stimuli displayed in the precritical trials among experiments to induce different expectations and intentions. The variations included presentations in the precritical trials of differently colored nonpredictive singletons (Experiment 1), nonpredictive form singletons (Experiment 3), color-heterogeneous displays with two colors alternating in adjacent positions (Experiment 4), two differently colored homogeneous displays (Experiment 5), a single color change with homogeneous displays (Experiment 6), and form-heterogeneous displays with two forms alternating in adjacent positions (Experiment 7).

All experiments used a modified version of the paradigm introduced by Gibson and Jiang (1998, Experiment 1), in which instructions inform about the precritical trials but make no reference to the display types in the critical and the postcritical (i.e., color-heterogeneous) trials. The basic assumption underlying the present experiments is that the precritical trials induce implicit expectations about what the following displays will look like. More precisely, a schematic representation of the spatial and temporal aspects of each trial is formed on the basis of the instructions and the first trials, which represents the typical layouts and their sequence in each trial as well as their relation to the task. Thus, different presentation conditions are assumed to induce different expectations, and the presentation conditions in the critical trial may or may not match the expectations. If the singleton in the critical trial is discrepant from expectations, surprise capture should occur—that is, the position of the singleton should be attended to.

I assessed attentional capture by observing the combined effects of a set-size variation and the presentation of the singleton at the target's position. Search demands were high with homogeneous displays or heterogeneous displays with nonpredictive singletons in the precritical trials, with RT depending on set size. In contrast, if a feature at the target's position captures attention, RT should not depend on set size. That is, set size and trial type should interact when the singleton at the target's position captures attention. If, however, the singleton fails to capture attention, no Set Size × Trial Type interaction should occur.

For the interpretation of the critical trial RTs, it must be considered that RTs could be longer in the critical trial than in the postcritical trials because of the surprise-induced RT delay (cf. Gibson & Jiang, 1998; Horstmann, 2002; Meyer et al., 1991), caused by a distraction through the surprising event. The RT delay probably reflects decision-level processes concerning the surprising event (e.g., a determination of the meaning and the task relevance of the surprising event) as well as automatic processes, such as the change of the current sets of schemas (e.g., Schützwohl, 1998). The duration of these surprise-related processes should thus depend on the type of the surprising change but should be independent of set size. Further, because an attentional shift and processes concerning the surprising event occur at different processing stages, their effects must be additive (Sternberg, 1969). As a concrete example, if one assumes that the attention shift to the surprise singleton has a latency of 300 ms (independent of set size) and the decision-level processes have a duration of 500 ms (also independent of set size), the RT in the critical trial will be 800 ms, independent of set size (ignoring the duration of response-related

---

[1] Note that because the spatial position of a feature is preattentively not precisely available (Treisman & Gelade, 1980), the discrepancy cannot pertain to the "wrong" position of the feature but only to its presence or absence in the entire scene. That is, if a red object is presented consistently at a specific spatial position until its position is changed in the critical trial, this will normally not be detected preattentively but only after the prior or the actual position of the object has been attended to.

processes). If the mean RT with serial search (e.g., in the precritical trials) is 500 ms with a small set size and 1,200 ms with a large set size, RTs in the critical trial will be longer than with serial search with a small set size but shorter with a large set size. This pattern of results was observed in Horstmann (2002, Experiment 3) with the set sizes of 4 versus 12 stimuli. Thus, the critical result indicative of attentional capture in the present experiments is a Set Size × Trial Type interaction but not a general decrease of RTs with all set sizes. A general decrease of RTs with all set sizes is only to be expected if the set sizes tested are sufficiently large (see, e.g., the present Experiment 8).

The questions addressed in the experiments are as follows. Experiments 1, 3, and 7 examine the specificity of expectations within and between the static feature dimensions of color and orientation. More precisely, I test whether the presentation of a nonpredictive color versus form singleton in the precritical trials eliminates attentional capture by a color singleton defined by a new color. Experiments 4–6 further test whether the expectation discrepancy instigating attentional capture concerns color or other aspects of the display in the critical trial, such as heterogeneity. In addition, Experiments 4–6 test whether the alternative hypothesis that singletons always capture attention (e.g., Theeuwes, 1992, 1994; Theeuwes & Godijn, 2001) can account for attentional capture by the unannounced first presentation of a singleton. Finally, Experiment 8 seeks to replicate the original attentional capture effect after color-homogeneous precritical trials with large set sizes of 25 versus 49 stimuli.

Experiment 1A: Homogeneous Versus Heterogeneous Displays in the Precritical Trials

Experiment 1A seeks to replicate and extend Horstmann's (2002) Experiment 3, in which the first unannounced presentation of a color singleton at the target's position resulted in a strong reduction of the set-size effect on RTs. The replication part of the experiment was instantiated in the *homogeneous display* condition, which differs from Horstmann's Experiment 3 only in minor changes in the stimulus material. As an extension of that experiment, the *heterogeneous display* condition was created. In this condition, a color singleton was presented in each trial; in the precritical trials, however, the singleton did not predict the position of the target and should therefore be ignored (e.g., Yantis & Egeth, 1999). In the critical and the postcritical trials, a different color was used for the singleton; however, it was always the target.

I included the heterogeneous display condition to examine the specificity of expectations concerning color. I reasoned that if expectations are highly specific within the static feature dimension of color, the new color in the critical trial should be expectancy discrepant and thus capture attention. In contrast, if expectations are broad and fuzzy and do not discriminate strongly within the color dimension, the new color should be consistent with expectation and thus fail to attract attention. Schützwohl's (1998; see also Schützwohl, 1993) research suggests that variability within a dimension in the precritical trials reduces the unexpectedness of a new value on this dimension in the critical trial. For example, Schützwohl (1993, Experiment 6) presented words in the precritical trials either in seven different typefaces or in a constant typeface. A new typeface presented in the critical trial was considered as unexpected by fewer participants and was rated as less surprising in the variable-typeface condition than in the constant-typeface condition. Thus, there appears to be a degree of generalization over the specific stimuli presented, which suggests that the new color in the critical trial of the present experiment will probably not be highly expectancy discrepant. Moreover, because the task is to ignore the color singleton, there is no additional incentive to discriminate neatly between the presented colors.

Method

Participants. Fifty students (13 men and 35 women) participated in the study for a small monetary exchange. Their mean age was 25.2 years ($SD = 5.9$).

Apparatus. A microcomputer, equipped with an Intel 80486/100MHz central processing unit, a keyboard, and a 15-in (38.10-cm) cathode ray tube monitor, was used for stimulus presentation and response registration. Response keys were two adjacent keys in the lower row of the keyboard (the *arrow left* and *arrow down* keys). ERTS (BeriSoft, Frankfurt, Germany) was used for event scheduling and RT measurement.

Design. Four conditions resulted from orthogonally crossing the two between-subjects variables, set size (4 vs. 12 letters) and display type, in the precritical trials (homogeneous vs. heterogeneous).

Stimuli. Figure 1 illustrates the stimulus conditions. Stimuli were presented against a black background. Viewing distance was approximately 60 cm. The fixation cross subtended 0.2° of visual angle. The letters (all taken from the same sans serif font type) were uppercase and subtended about 0.7° in height and 0.6°–0.9° in width. The target letters were *L* and *R*. For the set-size 4 condition, the 3 distractor letters were *P*, *T*, and *W*. For the set-size 12 condition, the 11 distractor letters were *P*, *B*, *K*, *N*, *S*, *H*, *T*, *I*, *E*, *J*, and *W*. The colors of the letters were white, green, or red, depending on condition. Colors were not matched for luminance; rather, it was assumed that the subjectively large difference in hue would be the most important difference among the colors (e.g., Folk et al., 1992). Distance between letters (about 2°), rather than retinal eccentricity, was held constant across the two set-size conditions. For the set-size 4 condition, the letters were presented on the invisible circumference of a small (1.3° radius) circle; for the set-size 12 condition, the letters were presented on the circumference of a large (3.8° radius) circle. The reason for letting retinal eccentricity rather than the distance between the letters (density) vary between the two set sizes was that density is assumed to be an important determinant of salience (cf. Yantis & Egeth, 1999; see also Nothdurft, 2000): Other things being equal, the higher the density of stimuli is, the higher is the salience of a singleton. In contrast, eccentricity per se affects RTs in visual search only slightly, given that targets and nontargets are presented at equal eccentricities (Wolfe, O'Neill, & Bennett, 1998), as secured by the use of circular displays in the present experiment.

Procedure. Each trial began with the presentation of the fixation cross in the center of the screen for 1 s (see Figure 1). The fixation cross was replaced by the letter array, which was presented until a response was registered but for a maximum of 4 s. The imaginary circle on which the letters were presented had its center at fixation. In the set-size 4 condition, letters were presented on the 3, 6, 9, and 12 o'clock positions of the circle. In the set-size 12 condition, letters were presented on the 12 full-hour positions of the circle.

In each trial, 1 of the 2 target letters (*L* or *R*) plus 3 (in the set-size 4 condition) or 11 (in the set-size 12 condition) distractor letters were presented. Which of the 2 target letters appeared, its position, and the position of the distractor letters were determined at random, with a new random sequence computed for each subject. The 2 target letters (*L* and *R*) were equiprobable. The participants were instructed to press the left response key if the search array contained an *L* and the right response key if it contained an *R*. Response speed was strongly emphasized, but participants were instructed to be accurate as well; false responses were immediately followed by error feedback, which consisted of a short tone.

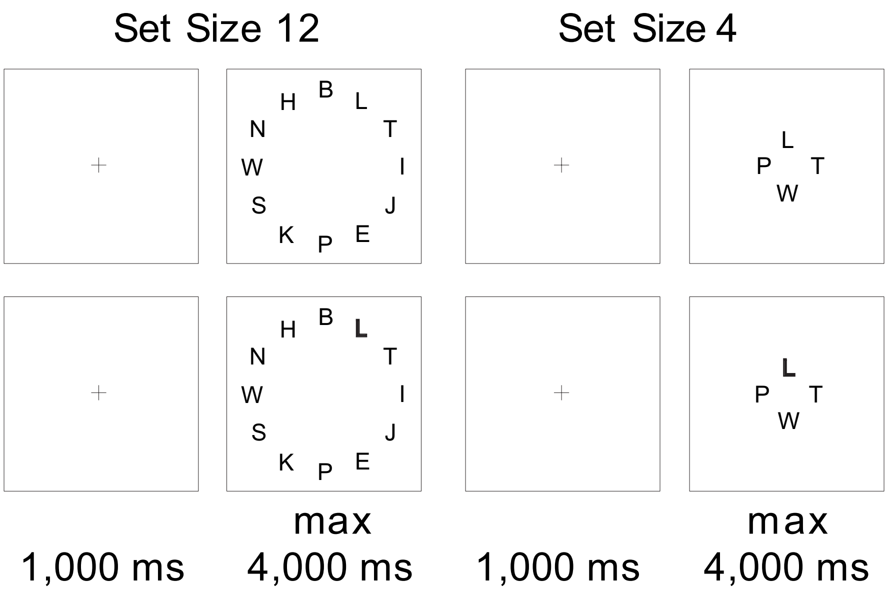

Figure 1. The trial structure in Experiment 1. Each trial consisted of two frames: a fixation cross presented for 1,000 ms, and the search display, which was presented until a response was made. Two display sizes—12 letters and 4 letters—are depicted. There were two types of displays. In homogeneous displays (upper row), all letters had the same color. In contrast, in heterogeneous displays (lower row), one letter had a different color (symbolized by a stronger line). max = maximum.

After 12 trials of familiarization with the task, there was one single experimental block that comprised 96 trials. The first 48 trials constituted the *precritical trials*. Trial 49 was the *critical trial*, and the following 47 trials constituted the *postcritical trials*. The transition between the last precritical trial and the critical trial did not differ from the transition between the other trials.

In the homogeneous-display condition, the distractor letters were all green throughout the experiment. In this condition, the target letter was likewise green in the precritical trials, and it was white in the critical and in the postcritical trials. In contrast, in the heterogeneous-display condition, one randomly determined letter was white in the precritical trials. In the critical and postcritical trials, the target letter appeared as a red singleton among green distractor letters.

Results

Errors and RTs longer than 3,000 ms were excluded from the RT analysis, which resulted in a loss of 4.6% (3.6% due to errors; see Table 1) of the experimental trials, two of them critical trials; this reduced the number of participants to 48. (The design assumed equal numbers of participants in each cell, but, because of an error, there were 13 participants in the set-size 4 homogeneous-display condition and 11 participants in the set-size 12 heterogeneous-display condition.) Figure 2 shows the mean RTs for the homogeneous-display condition (Figure 2, left panel) and the heterogeneous-display condition (Figure 2, right panel). Error rates are depicted in Table 1. In the precritical trials, the homogeneous and the heterogeneous display groups both showed large set-size effects (71 ms/letter and 91 ms/letter, respectively). In contrast, in the postcritical trials, the set-size effect was very small (9 ms/letter and 10 ms/letter, respectively). In the critical trial, however, the results differed markedly: Whereas the set-size effect was only 11 ms/letter in the homogeneous display condition, it remained high (86 ms/letter) in the heterogeneous display condition.

RT analysis paralleled that of Horstmann (2002, Experiment 3). First, an overall analysis of variance (ANOVA) was conducted. This 2 (set size: 4 vs. 12) × 2 (precritical trials: homogeneous vs. heterogeneous displays) × 3 (trial type: precritical vs. critical vs. postcritical trial) ANOVA revealed significant main effects of trial type, $F(2, 88) = 43.1$, $p < .01$, and set size, $F(1, 44) = 33.1$, $p < .01$, as well as a significant Trial Type × Set Size interaction, $F(2, 88) = 12.3$, $p < .01$, and two marginally significant interaction terms, Precritical Trials × Set Size, $F(1, 44) = 4.0$, $p = .05$, and Trial Type × Precritical Trials × Set Size, $F(2, 88) = 3.6$, $p = .05$ (here and henceforth, I computed the interaction terms involving

Table 1
Error rates (in percentages) for Experiments 1A and 1B

Condition	Precritical	Critical	Postcritical
Experiment 1A			
Homogenous display condition			
Set size 4	3.4	0.0	3.1
Set size 12	5.0	0.0	3.4
Heterogeneous display condition			
Set size 4	3.4	7.7	1.6
Set size 12	6.8	0.0	2.5
Experiment 1B			
Set size 4	4.1	0.0	4.7
Set size 12	6.0	0.0	2.9

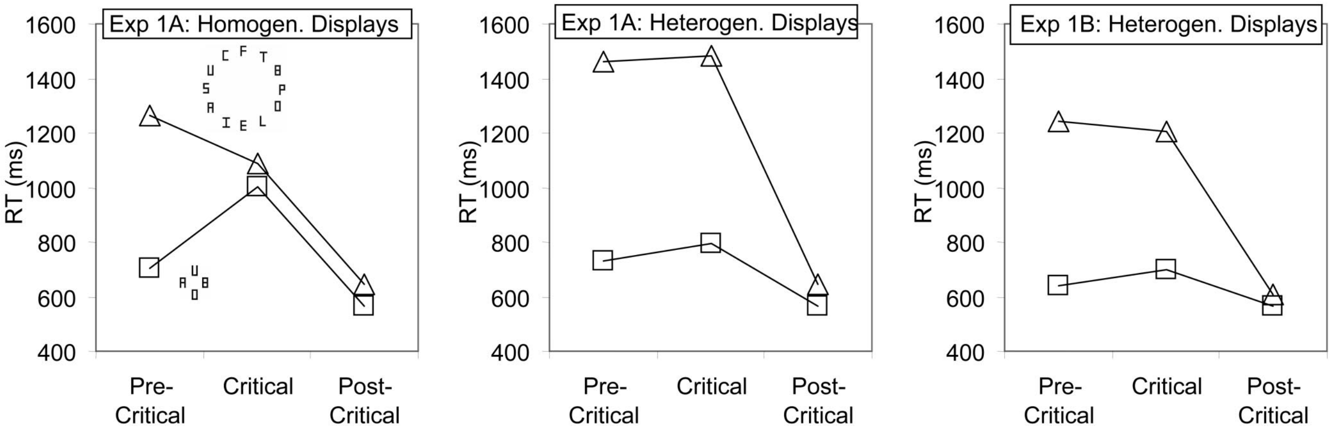

Figure 2. Results from Experiment 1. Mean reaction times (RTs) are presented for the three types of trials (precritical, critical, and postcritical), separately for each set size (set-size 12: triangles; set-size 4: squares). The left part of the figure shows the results of the homogeneous-display conditions, whereas the right part of the figure shows the results of the heterogeneous-display condition. Exp = Experiment; Homogen. = homogeneous; Heterogen. = heterogeneous.

trial type using Huynh–Feldt corrected degrees of freedom to adjust for violations of sphericity assumptions; the uncorrected degrees of freedom are reported to maintain readability). A corresponding error analysis (of error proportions) revealed no significant main effects or interactions (*F*s < 1.8, *p*s $> .10$), indicating that the interpretation of the RT results is not complicated by a speed–accuracy trade-off.

Whether attentional capture occurred in the homogeneous and in the heterogeneous display conditions was tested by two separate Set Size (4 vs. 12) × Trial Type (precritical vs. critical trial) ANOVAs. Search was predicted to be serial in the precritical trials in both conditions, implicating a large set-size effect in these trials. In contrast, surprise capture by the singleton in the critical trial should eliminate the set-size effects in the homogeneous-display condition but not—or not as strongly—in the heterogeneous-display condition. For this reason, I predicted a Set Size × Trial Type interaction for the homogeneous display condition (which would indicate a reduced set-size effect in the critical trial) but not for the heterogeneous display condition (which would indicate similar set-size effects in the precritical and the critical trial). The results were consistent with these predictions. The ANOVA revealed a significant main effect of set size in both conditions—homogeneous displays, $F(1, 23) = 6.2$, $p < .05$; heterogeneous displays, $F(1, 21) = 41.6$, $p < .01$—that was modified by a significant Trial Type × Set Size interaction in the homogeneous display condition, $F(1, 23) = 8.3$, $p < .01$, but not in the heterogeneous display condition ($F < 1$; the trial-type main effects were not significant, *F*s < 1). Follow-up *t* tests revealed that in the homogeneous display condition, set size had an effect on RTs in precritical trials, $t(23) = 6.7$, $p < .01$, but not in the critical trial, $t(21) < 1$, whereas in the heterogeneous display condition, set size had an effect on RTs both in the precritical trials, $t(21) = 16.6$, $p < .01$, and in the critical trial, $t(21) = 3.2$, $p < .01$.

I conducted the analogous analyses by comparing performance in the critical trial with performance in the postcritical trials. According to the present hypotheses, the RT in the critical trial of the homogeneous-display condition reflects primarily the additive components of (a) time for surprise capture to occur, (b) time consumed by the decision-level processes concerned with the surprising event, and (c) the processes that translate the stimulus letter into the response. In contrast, in the postcritical trials, RT should depend primarily on the additive components of (a) the latency of contingent capture and (b) stimulus–response (S-R) translation processes. Neither of the mentioned processes in the critical and the postcritical trials should be affected by set size. However, because the latency of contingent capture was shorter than the latency of surprise capture and because the processing of the surprising event was absent from the postcritical trials, the RTs in the critical trials should be longer than those in the postcritical trials. These predictions were supported by the ANOVA for the homogeneous display condition, which revealed a significant main effect of trial type only, $F(1, 23) = 25.8$, $p < .01$, whereas neither effect involving set size was significant (*F*s < 1).

For the heterogeneous-display condition, RTs in the critical and the postcritical trials should be different with respect to the efficiency of the target search, being nonefficient in the critical trial but efficient in the postcritical trials. Consistent with these predictions, there were significant main effects of block, $F(1, 21) = 23.7$, $p < .01$, and set size, $F(1, 21) = 12.3$, $p < .01$, as well as a significant Block × Set Size interaction, $F(1, 21) = 7.7$, $p < .05$. This pattern of results reveals that the set-size effect was reduced only after the critical trial.

When would observers in the heterogeneous-display condition change their search strategy in response to the new stimulus conditions? To answer this question, I examined the surprise trial and the immediately following trials. There was a rapid decrease in the set-size effect, beginning with the trial immediately following the surprise trial (i.e., search rates of 86, 29, 24, and 17 ms/item in the critical and the three following trials, Trials 49, 50, 51, and 52, respectively), which indicates that the majority of participants noticed the change in the surprise trial or in the following trial.

Discussion

The findings for the homogeneous-display condition closely replicate the results of Horstmann (2002, Experiment 3), which also found that the set-size effect present in the precritical trials was strongly reduced in the critical trial, which indicates that the surprise singleton captured attention. Also, as in that prior experiment, the set-size effect was similar to that in the postcritical trials, which indicates that attention capture was similarly efficient on its first presentation and in the postcritical trials, in which the singleton could be used intentionally to guide attention.

RTs were elevated in the critical trial as compared with the postcritical trials, reflecting distraction by the surprising event. This effect is usually obtained when a surprising event is presented during a reaction task (e.g., Horstmann & Schützwohl, 1998; Meyer et al., 1991, 1997; Niepel et al., 1994; Schützwohl, 1998). Because the decision-level processes reflected by the distraction and the attentional shift toward the surprising event logically occur at different processing stages, their effects should be additive (Sternberg, 1969). This assumption is consistent with the additivity of the RT increase with set size in the critical and the postcritical trials.

In contrast, the heterogeneous-display condition shows that not all types of display changes involving singletons capture attention: When observers were presented with a nonpredictive singleton in each precritical trial, no immediate attentional response occurred to an unannounced color change. This pattern of results contrasts sharply with the homogeneous display condition. In the present account, this result was due to expectations built up during the precritical trials that were matched—or at least did not mismatch—in the critical trial. That is, because of the heterogeneous precritical trials, participants expected a color singleton in each trial. An expectation of a color singleton could further be analyzed as consisting of two confounded expectations: the expectation of a certain color (or range of colors), and the expectation of a singleton (or of heterogeneity of the display). One aim of the following experiments is to clarify the contribution of these confounded expectation components to the mismatch in the critical trial. However, the following Experiment 1B was conducted to first replicate the heterogeneous display condition of Experiment 1 while eliminating a possible alternative account.

Experiment 1B: Heterogeneous Displays With Digital-Clock Stimuli

Experiment 1B seeks to replicate the absence of attentional capture in the heterogeneous display condition, with minor differences from Experiment 1A in stimuli and procedure. The most important change concerned the stimuli. In particular, digital-clock-like stimuli were used to eliminate any differences in basic features between targets and distractors (Gibson & Jiang, 1998; Horstmann, 2002, 2003). This was to ensure that the participants would not enter a "feature-detection mode" (Bacon & Egeth, 1994) on the basis of subtle differences between targets and distractors on a basic feature dimension. Second, the number of precritical trials was increased to test whether absence of attentional capture in the heterogeneous display condition was due to its insufficient exposure to the precritical stimuli. Third, the nonpredictive singleton in the precritical trials was a red letter among green letters, whereas the predictive singleton in the critical and the postcritical trials was a white letter. This change tests the possible contention that the red letter in the critical trial of Experiment 1A did not capture attention because it was less salient than the to-be-ignored white letter in the precritical trials.

Method

Participants. Participants were 10 men and 15 women who participated for a small monetary exchange. Their mean age was 21.8 years ($SD = 2.6$). None of them had participated in the previous experiment. (Naive participants were used in all experiments reported in this study.)

Design. Participants were randomly assigned to one of the two set-size conditions (4 vs. 12 letters).

Apparatus, stimuli, and procedure. Hardware and software were of the same type as in Experiment 1A. Stimuli were presented against a black background. The letters were constructed from vertical and horizontal line segments only. Each letter subtended 0.8° in height and 0.7° in width. The target letters were *H* and *U*. For the set-size 4 condition, the distractor letters were *A*, *B*, and *D*. For the set-size 12 condition, the distractor letters were *A*, *B*, *C*, *D*, *E*, *F*, *I*, *L*, *P*, *S*, and *T*. The stimuli were presented on the imagined circumference of a circle with radii of 0.9° and 2.3° in the set-size 4 and the set-size 12 conditions, respectively. The remaining stimuli were the same as in Experiment 1A.

In the 24 practice trials and the following 72 precritical trials, the distractor letters were green and the singleton letter was red. In the critical trial and the following 23 postcritical trials, the target letter was always white. For comparison purposes, note that the present experiment is a clone of Horstmann's (2002) Experiment 3, the only difference being in the precritical trials. That is, it realizes stimulus conditions in the critical trial that have proven to cause attentional capture after homogeneous precritical trials.

Results and Discussion

Errors and RTs longer than 3,000 ms were excluded from the analysis, resulting in a loss of 5.3% of the experimental trials, including one in the critical trial; this reduced the number of participants contributing to the RT analysis to 24 (12 in each group).

Figure 2 (right panel) shows the mean RTs for the two groups in the three trial types, and Table 1 shows the corresponding error rates. RTs were first analyzed by means of a 2 (set size: 4 vs. 12) × 3 (trial type: precritical vs. critical vs. postcritical trials) ANOVA. Both main effects were significant—set size, $F(1, 22) = 49.8$, $p < .01$; type of trial, $F(2, 44) = 39.8$, $p < .01$—as was the Set Size × Trial Type interaction, $F(2, 44) = 20.3$, $p < .01$.

A corresponding error analysis revealed a significant main effect for trial type only, $F(2, 44) = 15.4$, $p < .01$. The set-size main effect ($F < 1$) and the Trial Type × Set Size interaction, $F(2, 44) = 2.0$, $p > .10$, were not significant. Fewer errors occurred in the critical trial (0.0%) than in precritical trials (5.0%) and in postcritical trials (3.8%), indicating that observers traded speed for accuracy in the surprise trial. This trade-off, however, does not complicate the interpretation of the RTs, because the main effect for type of trial is not essential in the present analysis.

The main prediction concerned the interaction effect between set size (4 vs. 12) and trial type (precritical vs. critical trial). A corresponding ANOVA of the RTs revealed a set-size effect only, $F(1, 22) = 62.0$, $p < .01$; the other main effect and the interaction term were not significant (Fs < 1). There was thus no indication of attentional capture in the critical trial.

The critical and the first four postcritical trials were examined as in Experiment 1A. Again, there was a rapid decrease in the set-size effect, beginning in the trial immediately following the critical trial. Similar to Experiment 1A, this decline was strongest immediately after the critical trial, indicating that most of the participants recognized that the singleton was always the target letter and changed their attentional control settings to the effect that they rapidly shifted their attention to the singleton letter (calculated search rates were 63, 20, 11, and 6 ms/item in the critical and the three following trials, Trials 73, 74, 75, and 76, respectively).

One can draw the preliminary conclusion that two experiments, with minor differences concerning stimulus shape, number of precritical trials, and eccentricity, revealed nearly identical results; with nonpredictive singletons in the precritical trials, an unannounced color change of the singleton did not capture attention.

Experiment 2: Colored Squares

In some of the following Experiments 2–7, it was necessary to dissociate the singleton feature from the letter stimulus. For this reason, the singleton feature pertained to a color patch that was presented as a background for the letters. This procedure had already been used by Horstmann (2002, 2004, in press) to allow for a temporal disconnection of the color singleton and the letters relevant for the search task. However, because this procedure had not been used before in the RT paradigm, I conducted Experiment 2 to first establish the surprise capture effect with these stimuli before, in Experiments 3–7, testing further hypotheses.

Method

Participants. Participants were 6 men and 26 women with a mean age of 24.4 years (*SD* = 5.8). Compensation was €1.

Stimuli, procedure, and design. Black letters (0.7° × 0.8°), all composed of vertical and horizontal line segments (as in Experiment 1B), were presented against colored squares (1.2° × 1.2°). The background color was black. In the set-size 12 condition, the compound letter/square stimuli were presented in the 12 full-hour positions on the circumference of an imaginary 3.3°-radius circle, whereas in the set-size 4 condition, the stimuli were presented on the 3, 6, 9, and 12 o'clock positions of a 1.5°-radius circle. The compound letter/square stimuli were presented until a response was made, but a trial was aborted and counted as an error if there was no response within 4 s. Error feedback was given if the wrong reaction key was pressed.

The precritical trials comprised 48 experimental and 12 practice trials. The squares in the precritical trials all had the same color. In each trial, each position had an equal probability of containing the target. The critical trial and the following 23 postcritical trials differed from the precritical trials only in that one square (the color singleton) was presented in a different color. The color-singleton square always contained the target.

Participants were randomly assigned to one of four conditions that resulted from orthogonally combining the variables set size (4 vs. 12 letters) and standard color–singleton color (red–green vs. green–red).

Results

Errors and RTs longer than 3,000 ms were excluded from the RT analysis, resulting in a loss of 4.1% of the experimental trials. One of the errors occurred in the critical trial, reducing the number of participants in the RT analyses to 20, 10 per set size. Figure 3 (panel a) shows the main results. There was a large set-size effect in the precritical trials (87 ms/letter) but only a small set-size effect in the critical trial (4 ms/letter) and in the postcritical trials (15 ms/letter).

An overall ANOVA[2] with the variables' set size (4 vs. 12) and trial type (precritical vs. critical vs. postcritical) revealed significant effects of set size, $F(1, 18) = 5.6$, $p < .05$, and of trial type, $F(2, 36) = 23.9$, $p < .01$, and a significant Set Size × Trial Type interaction, $F(2, 36) = 13.6$, $p < .01$. A corresponding error analysis revealed no significant effects whatsoever (Fs < 1). Mean error rates appear in Table 2.

The main test determined whether the effect of set size was different for the precritical trials versus the critical trial. A 2 (set size: 4 vs. 12) × 2 (trial type: precritical vs. critical trial) ANOVA revealed the predicted Set Size × Trial Type interaction, $F(1, 18) = 16.6$, $p < .01$, and a significant set-size effect, $F(1, 18) = 6.1$, $p < .05$. The main effect of trial type was not significant, $F(1, 18) = 2.4$, $p > .10$. The significant interaction indicates that RTs changed differently in the two set-size conditions: They were 497 ms longer in the set-size 4 condition, $t(9) = 3.4$, $p < .01$, but 225 ms shorter in the set-size 12 condition, $t(9) = 2.3$, $p = .05$ (two-tailed).

I further tested whether the set-size effect was different in the critical and the postcritical trials. A 2 (set size: 4 vs. 12) × 2 (trial type: critical vs. postcritical trial) ANOVA revealed a main effect of trial type only, $F(1, 18) = 32.1$, $p < .01$ (all other Fs < 1), indicating that the set-size effects did not differ but that RTs in the critical trial were much longer than in the postcritical trials.

Discussion

The results replicate the results pattern of Experiment 1A, indicating that it is largely irrelevant whether the singleton color is a feature of the target or of a different object that occupies the same spatial position as the target. Having established that a color singleton captures attention with this experimental set-up, I return in Experiment 3 to the question of why the color singleton did not capture attention in the heterogeneous display conditions of Experiments 1A and 1B.

Experiment 3: Changed Singleton Dimension

Experiment 1 showed that when the stimuli in the precritical trials required the participants to treat color singletons as irrelevant, a new color singleton in the critical trial did not capture attention. This result was predicted on the assumption that the expectation concerning color is relatively broad within the dimension of color after the presentation of two rather different colors in the precritical trials of the heterogeneous display conditions. It is, however, unclear from this result whether it was indeed the color variation or rather the presentation of a singleton per se in the precritical trials that hindered attentional capture. To disentangle color variation and singleton presence, I included in each precritical trial of the present experiment an orientation singleton that was nonpredictive of the target's position. That is, the precritical trials were orientation heterogeneous but color homogeneous. In the critical trial, a color singleton was presented at the position of the target. Thus, if an expectation of a singleton can cause the absence of an attentional shift to the color singleton in the critical trial, the set-size effect in the critical trial should be equal to that

---

[2] The variable singleton/nonsingelton color, if included in the ANOVA, revealed no significant main effect or interactions with the other variables, with all Fs < 1.

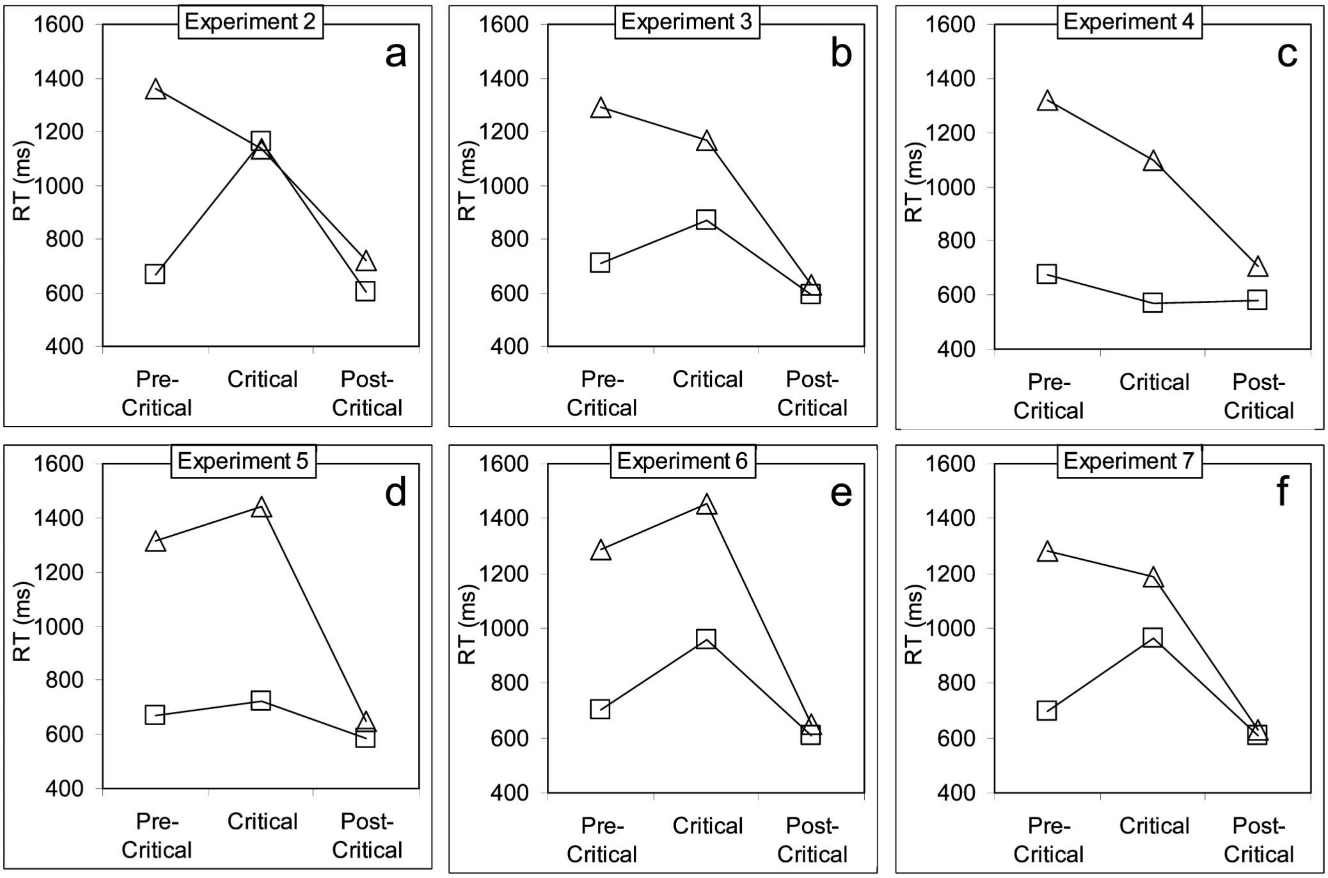

Figure 3. Mean correct reaction times (RTs) for both set-size conditions (set-size 12: triangles; set-size 4: squares) in the precritical, critical, and postcritical trials of Experiments 2–7.

in the precritical trials. If, conversely, color homogeneity in the precritical trials is the more important condition for attention to a new color in the critical trial, the set-size effect in that trial should be equal to that in the postcritical trials.

Table 2
Error rates (in percentages) for Experiments 2–7

Experiment/Condition	Precritical	Critical	Postcritical
Experiment 2			
Set size 4	3.1	0.0	4.6
Set size 12	2.7	9.1	3.4
Experiment 3			
Set size 4	2.6	0.0	3.8
Set size 12	3.9	6.3	3.8
Experiment 4			
Set size 4	2.9	0.0	4.3
Set size 12	4.3	0.0	2.4
Experiment 5			
Set size 4	2.4	7.7	1.3
Set size 12	2.7	0.0	3.2
Experiment 6			
Set size 4	4.4	6.7	3.3
Set size 12	4.5	0.0	3.3
Experiment 7			
Set size 4	3.8	0.0	2.5
Set size 12	4.3	0.0	3.9

Method

Participants. Participants were 6 men and 26 women with a mean age of 24.4 years (*SD* = 5.8). Compensation was €1.

Stimuli, procedure, and design. These were the same as in the preceding experiment, except for the composition of the precritical trials. In the precritical trials, all but one square (the orientation-singleton square) were presented in upward orientation, whereas the orientation-singleton square was rotated by 45° (see Figure 4). The position of the orientation singleton was determined randomly, with all positions being used equally often. On each trial, each position had an equal probability of containing the target, such that the position of the orientation singleton did not predict the position of the target. The critical trial and the following 23 postcritical trials were the same as in Experiment 2, with a color-singleton square that always contained the target.

Results

Errors and RTs longer than 3,000 ms were excluded from the RT analysis, resulting in a loss of 4.6% of the experimental trials, one in the critical trial; this reduced the number of participants to 31, 16 in the set-size 4 condition and 15 in the set-size 12 condition. Figure 3 (panel b) shows the main results. There was a large set-size effect in the precritical trials (73 ms/letter) but only a small set-size effect in the postcritical trials (6 ms/letter). The set-size effect was reduced in the critical trial (38 ms/letter), although it was clearly not zero.

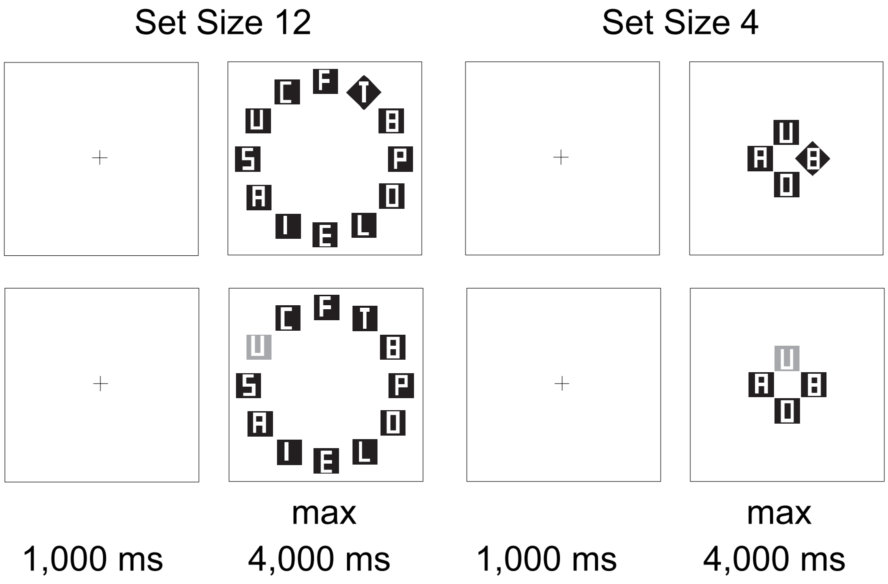

Figure 4. Trial structure and display types in Experiment 3. All trials were either form heterogeneous but color homogeneous (upper row) or color heterogeneous but form homogeneous (lower row). Display samples are shown for both set sizes (left: 12 letters; right: 4 letters). Different colors are symbolized by different shades of gray. max = maximum.

An overall ANOVA with the variables set size (4 vs. 12) and trial type (precritical vs. critical vs. postcritical) revealed significant effects of set size, $F(1, 29) = 30.0$, $p < .01$, and of trial type, $F(2, 58) = 62.7$, $p < .01$, and a significant Set Size × Trial Type interaction, $F(2, 58) = 22.2$, $p < .01$. A corresponding error analysis revealed no significant effects whatsoever (Fs < 1.4). Mean error rates appear in Table 2.

The main test determined whether the effect of set size was different for the precritical trials versus the critical trial. A 2 (set size: 4 vs. 12) × 2 (trial type: precritical vs. critical trial) ANOVA revealed the predicted Set Size × Trial Type interaction, $F(1, 29) = 9.0$, $p < .01$, and a significant set-size effect, $F(1, 29) = 36.1$, $p < .01$. The main effect of trial type was not significant ($F < 1$). The significant interaction indicates that RTs changed differently in the two set-size conditions; they were 159 ms longer in the set-size 4 condition, $t(15) = 3.2$, $p < .01$, but tended to be 123 ms shorter in the set-size 12 condition, $t(14) = 1.5$, $p = .15$ (two-tailed).

I further tested whether the set-size effect was different in the critical and the postcritical trials. A 2 (set size: 4 vs. 12) × 2 (trial type: critical vs. postcritical trial) ANOVA revealed main effects of set size, $F(1, 29) = 7.1$, $p < .05$, and trial type, $F(1, 29) = 75.4$, $p < .01$, as well as a significant Set Size × Trial Type interaction, $F(1, 29) = 7.9$, $p < .01$. The Set Size × Trial Type interaction indicates that set size had some effect in the critical trial but virtually no effect in the postcritical trials.

Discussion

Experiments 1A and 1B showed that the new color in the critical trial did not capture attention after repeated presentation of nonpredictive color singletons in the precritical trials. The present experiment reveals that it was not the presence of a nonpredictive singleton in the precritical trials per se that prevented attentional capture from occurring. The set-size effect in the critical trial of the present experiment was clearly reduced as compared with the precritical trials. That is, the change from single-colored displays in the precritical trials to the multicolored display in the critical trial appears to be sufficient to draw attention. However, attention capture was not as strong as in the present Experiment 2 (or as in Horstmann, 2002, Experiment 3). This indicates that the combination of a new color and the first presentation of a singleton was an especially potent attractor of attention.

In summary, the present experiment shows that the first presentation of a new color, which need not necessarily be the first presentation of a singleton, was sufficient to attract attention. In the next experiment, I test whether the first presentation of a singleton is sufficient to capture attention—that is, whether it captures attention in the absence of a change from single-colored to multicolored displays.

Experiment 4: Heterogeneous Nonsingleton Displays

The findings of the present and of prior experiments (Horstmann, 2002, 2004, in press) are consistent with the view that the first and unannounced presentation of a color singleton captures attention to the extent that the singleton color sufficiently deviates from expectations and does not capture attention if the color is largely consistent with expectations. However, the findings may be seen as being consistent with a quite different view on attentional capture: Singletons may always capture attention, but participants may immediately reorient away from the singleton if they believe that it is nonpredictive (e.g., Kim & Cave, 1999; Theeuwes & Godijn, 2001). According to this reorientation account, an unannounced singleton captures attention when preceded by homogeneous trials because, as participants do not expect a distracting singleton, they have no control setting to reorient away from the singleton. In contrast, an unannounced feature change of the singleton does not capture attention because participants know that color singletons are nonpredictive and hence ignore them. Expectation is perhaps a superfluous concept to explain the present results.

The following three experiments were designed to deal with this contention. In principle, the singleton-capture account can explain all cases of attentional capture by an unexpected singleton. Therefore, the experiments focus on conditions in which the singleton-capture hypothesis, but not the surprise-capture hypothesis, predicts that the singleton will capture attention in the critical trial. According to the singleton-capture account, a singleton in the critical trial should reveal attentional capture unless the precritical trials misguide the participants in ignoring singletons of any kind. In Experiment 4, heterogeneous displays are presented in the precritical trials that do not contain singletons; rather, each display contains red and green squares that are arranged such that the colors alternate. Because of the regular arrangement of the two colors, salience (sensu Wolfe, 1994) should be equal at all positions. More precisely, because there is a feature contrast for each of the squares with respect to the neighboring squares, no single position should attract attention more strongly than any other. As a consequence, there is no need to suppress an attentional response to a singleton. If color singletons always capture attention, given no intention to the contrary, a single red square among green squares (or vice versa) should capture attention in the critical trial.

The conditions in the present Experiment 4 were optimized to provide a clear test of a singleton-capture account for the present experiments; they were less optimal in testing the expectation-discrepancy account. As a result, the predictions are a bit more complicated. An expectation-discrepancy account predicts no initial capture by the singleton, because the singleton's color is consistent with expectations. However, this does not imply that the observers will not become aware of the changed composition of the critical trial display. However, detecting the difference in composition amounts to comparing specific combinations of features, and the detection of this discrepancy requires attention according to feature integration theory (e.g., Treisman & Gelade, 1980). Thus, the discrepancy will probably be detected only during serial search for the target letter. However, in this case, it is the color homogeneity of most items that is expectancy discrepant, and attending to this aspect of the display does not help in finding the target, because the homogeneous items are all distractors. In sum, I predict that the set-size effect in the critical trial will not be reduced, although the RTs may be longer than in the precritical trials, because of a detection of the discrepancy.

Method

Participants. Participants were 9 men and 15 women with a mean age of 24.5 years ($SD = 5.5$). Compensation was €1.

Apparatus, stimuli, procedure, and design. These were the same as in Experiment 2, except for the composition of the precritical trials. In these trials, heterogeneous displays were presented in which red and green squares occupied alternating positions. The two possible presentation positions for the red and green squares were used equally often. The sequence of the two positions was randomized. Target position was determined at random, and color was nonpredictive of the position of the target. The critical trial and the following 23 postcritical trials were identical to the corresponding trials of Experiments 2 and 3. That is, there was a singleton square (either red or green) among the nonsingleton squares at a random position that always contained the target.

Results

Errors and RTs longer than 3,000 ms were excluded from the analysis, resulting in a loss of 4.5% of the experimental trials. Figure 3 (panel c) presents mean RTs; error rates are given in Table 2. There was a large set-size effect in the precritical trials (81 ms/letter) and in the critical trial (66 ms/letter) but a small set-size effect in the postcritical trials (16 ms/letter).

An overall ANOVA with the variables set size (4 vs. 12) and trial type (precritical vs. critical vs. postcritical) revealed significant effects of set size, $F(1, 22) = 59.1$, $p < .01$, and of trial type, $F(2, 44) = 17.1$, $p < .01$, and a significant Set Size × Trial Type interaction, $F(2, 44) = 10.0$, $p < .01$. A corresponding error analysis revealed a significant trial type effect, $F(2, 44) = 9.0$, $p < .01$, indicating that participants traded speed for accuracy in the critical trial, in which no errors occurred, relative to the precritical and the postcritical trials (see Table 2). However, the main effect for trial type is not of prime interest for the present examination.

The main test concerned whether the effect of set size was different for precritical trials versus the critical trial. A 2 (set size: 4 vs. 12) × 2 (trial type: precritical vs. critical trial) ANOVA revealed main effects of set size, $F(1, 22) = 56.0$, $p < .01$, and trial type, $F(1, 22) = 5.2$, $p < .05$. The Set Size × Trial Type interaction was not significant ($F < 1$).

A corresponding ANOVA on the critical trial versus the postcritical trials tested whether the set-size effect in the critical trial was larger than in the postcritical trials. The ANOVA revealed significant main effects of set size, $F(1, 22) = 21.4$, $p < .01$, and trial type, $F(1, 22) = 7.9$, $p < .01$, and a significant Set Size × Trial Type interaction, $F(1, 22) = 8.8$, $p < .01$. That is, search in the critical trial was not as efficient as in the postcritical trials.

As before, the trials immediately following the critical trial were analyzed, and, similar to the prior experiments, the set-size effect declined rapidly; it was 66 ms/item in the critical trial and 30, 19, and 13 ms/item in the following three trials.

Discussion

The first presentation of a singleton in a visual-search task did not capture attention if this presentation was preceded by heterogeneous displays without a singleton. Because of the construction of the heterogeneous displays, there is no reason to assume that participants were set to ignore singletons. The experiment thus indicates that a singleton-capture account of attentional capture by surprising color singletons is difficult to defend. Conversely, the results are consistent with an expectation-discrepancy account of surprise capture: If the first presentation of a singleton is preceded by heterogeneous displays containing the very two colors presented in the critical trial, expectation deviation of the singleton's color should be zero.

As outlined in the introduction, a central assumption of the present model of surprise capture is that an expectancy discrepancy can be detected either preattentively or postattentively, depending on whether the expectation pertains to a simple feature or a feature combination. Accordingly, the discrepancy of the critical trial display cannot be determined preattentively in this experiment. However, after attention had been deployed to some of the stimuli in the display, the discrepancy should have become available. Consistent with this reasoning, most of the participants apparently used the singleton to locate the target in the trials immediately following the critical trial. It remains an open question, however, as to at what point in time the discrepancy was detected during the critical trial. Further experiments, possibly using eye movements as a more direct indicator of movements of attention during the critical trial, are required to answer this question.

Experiment 5: Homogeneous Displays of Varying Colors

Experiment 4 indicates that the presentation of a singleton was not sufficient to capture attention, even if the observers had no intention of ignoring the singleton. This result is consistent with the hypothesis that it is mainly the expectation of (task-irrelevant) color variation that eliminates surprise capture by the new color. This assumption suggests that color variation should eliminate surprise capture by an unannounced singleton even if the critical trial is preceded by homogeneous displays only, given that the homogeneous displays presented in different trials vary in color. Therefore, in Experiment 5 the precritical trials were composed of red and green homogeneous displays that were presented in a random sequence. The critical and the postcritical trials were the same as in the previous experiment.

A singleton-capture account predicts that the singleton captures attention because it is a singleton and the observers are not set to reorient away from a singleton (because no singleton was presented in the precritical trials). In contrast, an expectancy-discrepancy account predicts no capture of attention by the singleton, because the singleton's' color is consistent with expectations. However, because the new combination of colors may be detected after attention is focused to the singleton's position in the course of the visual-search task, RTs in the critical trial may be longer than in the precritical trials.

Method

Participants. Participants were 7 men and 19 women with a mean age of 23.8 years ($SD = 5.6$). Compensation was €1.

Apparatus, stimuli, procedure, and design. The only difference from Experiment 4 concerned the composition of the precritical trials, which comprised red and green homogeneous displays that appeared in a randomly determined order.

Results

Errors and RTs longer than 3,000 ms were excluded from analysis, which resulted in a loss of 3.5% of the experimental trials, two in the critical trials, and reduced the number of participants to 24, 12 per set-size condition. Figure 3 (panel d) shows the mean RTs; Table 2 displays the corresponding error rates. There was a large set-size effect in the precritical trials (80 ms/letter) and in the critical trial (90 ms/letter) but only a small set-size effect in the postcritical trials (8 ms/letter).

An overall ANOVA of the RTs with the variables' set size (4 vs. 12) and trial type (precritical vs. critical vs. postcritical) revealed significant effects of set size, $F(1, 22) = 25.5$, $p < .01$, and of trial type, $F(2, 44) = 30.5$, $p < .01$, and a significant Set Size × Trial Type interaction, $F(2, 44) = 15.9$, $p < .01$. A corresponding error analysis revealed no significant main effects or interactions (Fs $<$ 1.2). Mean error rates appear in Table 2.

The main prediction concerned whether the effect of set size was different for precritical trials versus the critical trial. A 2 (set size: 4 vs. 12) × 2 (trial type: precritical vs. critical trial) ANOVA revealed a main effect of set size only, $F(1, 22) = 30.8$, $p < .01$. The main effect of trial type, $F(1, 22) = 1.6$, and the Set Size × Trial Type interaction ($F < 1$) were not significant.

Conversely, the set-size effect was different for the critical trial and the postcritical trials. A 2 (set size: 4 vs. 12) × 2 (trial type: critical vs. postcritical) ANOVA revealed significant main effect of set size, $F(1, 22) = 24.1$, $p < .01$, and trial type, $F(1, 22) = 51.9$, $p < .01$, and a significant Set Size × Trial Type interaction, $F(1, 22) = 47.2$, $p < .01$.

As before, the trials immediately following the critical trial were inspected. The set-size effect declined rapidly in these trials; it was 90 ms/item in the critical trial and 10, 13, and 10 ms/item in the following three trials.

Discussion

The present experiment provides convergent evidence that presenting a singleton in the absence of an intention to ignore singletons is not sufficient to elicit attentional capture. Only homogeneous displays were presented in precritical trials, such that there was no reason whatsoever to adopt any set toward singletons. Nevertheless, the singleton did not capture attention.

The present results are consistent with an expectation-discrepancy account of surprise capture. Because more than one color was used for the homogenous displays, the singleton's color in the critical trial was consistent with expectations. Again, this does not mean that the presentation of the singleton in the critical trial following homogeneous displays was not at all expectation discrepant. However, the present results indicate that this latter kind of expectation discrepancy (i.e., concerning the composition of the display by schema-congruent elements) either was very weak or was detected only after attention had been allocated to the singleton's location. The latter alternative is quite plausible, given that the conjunction of features is achieved by focused attention (e.g., Treisman & Gelade, 1980). This implies that unless attention is focused on the location, each color is evaluated independently with respect to its expectation consistency. Consistent with these assumptions, the singleton was used by most or all participants even in the trial that followed the critical trial, as indicated by the very small set-size effect in this trial.

Experiment 6: Single Color Change

Experiments 4 and 5 yield evidence that a singleton that does not deviate from expectations does not capture attention, even if the participants are not set to ignore singletons. Experiment 6 is an extension of Experiment 5 in that it changes (only) one aspect—that is, the homogeneous displays in the precritical trials were presented not in a random sequence but blockwise. For example, Trials 1–24 were homogeneously green, Trials 25–48 were homogeneously red, and the critical trial presented a green singleton among red nonsingleton stimuli. With this change, the degree of color variation necessary to induce an expectation of color variation is tested. When color changes only once, there is obviously a very moderate manipulation of an expectation of color variation, which thus tests a lower limit for the reported effect. For this reason, Experiment 6 is more exploratory in nature than the preceding experiments: Given that relatively little is known about expectations, it may be that the more recent trials (Trials 25–48) have more impact than the earlier trials (Trials 1–24). Conversely, there was little more than 1 min between Trial 24 and the critical trial (Trial 49), which suggests that the earlier trials may still influence the content of expectations.

Experiment 6 also provides an additional test of the current model for the RTs in the critical trial. To recapitulate, I assume that the RTs in the critical trial can be analyzed into three additive components: (a) the time to find the target, (b) decision-level processing of the surprising event, and (c) S-R translation processes that eventually lead to the response. On the basis of the present account of surprise capture and previous results, I assume that the color change between Trials 1–24 and Trials 25–48 elicits surprise, because the expectation primarily concerns color (e.g., present Experiments 3–5), and the change is preceded by homogeneous displays of only one color, which has proved to be a condition in which surprise is reliably elicited (e.g., present Experiment 1; Horstmann, 2002, 2004, in press). I therefore predict that the decision-level processing of the surprising event will result in a prolongation of the RTs that is independent of set size. However, with a color change of the whole display, attending to the changed color does not help in finding the target. Therefore, the target can still only be found with serial search, as in the preceding trials. Thus, the set-size effect present in Trials 1–24 should also be present in Trial 25. Because of the additivity of these two components, I predict a general increase of RT for Trial 25, which does not interact with set size.

Method

Participants. Participants were 11 men and 18 women with a mean age of 25.8 years (*SD* = 9.2). Compensation was €1.

Apparatus, stimuli, procedure, and design. The only difference from Experiment 5 was that red and green homogeneous displays were presented blocked, with one color used in Experimental Trials 1–24 and the other color in Experimental Trials 25–48. The color in the 12 practice trials was the same as in Trials 1–24. The nonsingleton color in the 24 heterogeneous display trials (critical and postcritical trials) was the same as in the immediately preceding trials (Trials 25–48).

Results

Errors and RTs longer than 3,000 ms were excluded from analysis, resulting in a loss of 4.6% of the experimental trials, one in the critical trial, and reducing the number of participants to 28, 14 per set-size condition. Figure 3 (panel e) shows the RT data for the two set-size conditions in the precritical, critical, and postcritical trials. The corresponding error rates are listed in Table 2. There was a large set-size effect in the precritical trials (74 ms/letter) and in the critical trial (62 ms/letter) but only a small set-size effect in the postcritical trials (6 ms/letter).

Color change within the precritical trials. A 2 (set size: 4 vs. 12) × 2 (trial: Trial 1–24 vs. Trial 25 vs. Trial 26–48) ANOVA was conducted to test whether the first presentation of a homogeneous display of a different color elicited surprise. (Because two errors occurred in the 25th trial, the analysis was conducted with only 26 participants.) The analysis revealed significant main effects of set size, $F(1, 24) = 40.9$, $p < .01$, and trial type, $F(2, 48) = 14.0$, $p < .01$, but no significant Set Size × Trial Type interaction ($F < 1$). A corresponding error analysis yielded no significant effects (Fs < 2.7, ps $> .10$). Planned comparisons revealed that the mean RT—collapsed over the set-size variable—was longer in Trial 25 (1,282 ms) than in Trials 1–24 (944 ms), $t(25) = 4.1$, $p < .01$, and longer than in Trials 26–48 (976 ms), $t(25) = 3.5$, $p < .01$.

Critical trial. An overall ANOVA with the variables set size (4 vs. 12) and trial type (precritical vs. critical vs. postcritical) revealed significant effects of set size, $F(1, 26) = 22.6$, $p < .01$, and of trial type, $F(2, 52) = 25.4$, $p < .01$, and a significant Set Size × Trial Type interaction, $F(2, 52) = 6.3$, $p < .05$. A corresponding error analysis revealed no significant main effects or interactions (Fs < 1.2). Mean error rates appear in Table 2.

The main test concerned whether the effect of set size was different for precritical trials versus the critical trial. A 2 (set size: 4 vs. 12) × 2 (trial type: precritical vs. critical trial) ANOVA revealed main effects of set size, $F(1, 26) = 23.1$, $p < .01$, and trial type, $F(1, 26) = 4.7$, $p < .05$, whereas the Set Size × Trial Type interaction was not significant ($F < 1$). The effect of trial type reflected the surprise-induced RT increase, with RTs being 214 ms longer in the critical trial than in the precritical trials.

However, the set-size effect was different for the critical trial and the postcritical trials. A 2 (set size: 4 vs. 12) × 2 (trial type: critical vs. postcritical) ANOVA revealed significant main effects of set size, $F(1, 26) = 6.5$, $p < .05$, and trial type, $F(1, 26) = 35.5$, $p < .01$, and a significant Set Size × Trial Type interaction, $F(1, 26) = 5.4$, $p < .05$.

As before, the trials immediately following the critical trial were inspected. The set-size effect declined rapidly in these trials; it was 62 ms/item in the critical trial and 3, 2, and 3 ms/item in the following three trials.

Discussion

Experiment 6 reveals three interesting results. First and most important, the results are at odds with a singleton-capture account that predicts attentional capture by a salient singleton given that observers do not intend to ignore it. In particular, although there was no reason for participants to intentionally ignore a singleton because they never saw a singleton in the precritical trials, no attentional capture by the singleton in the critical trial was observed. Results are thus more consistent with an expectation-mismatch explanation of surprise capture. According to this account, the change of color of the homogeneous displays during the precritical trials disconfirmed the observers' original expectation concerning color and led them to broaden the range of expected colors.

Second, it is remarkable that the acceptability of both presented colors persisted for the 24 trials that lay between the color change within the homogeneous displays and the critical trial. This result supports an assumption underlying the present experiments: Surprise can be induced only once via one type of surprising event (e.g., a color change), because after the surprising event, expectancies are changed (e.g., Schützwohl, 1998).

Third, the color change within the precritical trials resulted in longer RTs as compared with the preceding and the following trials, with the RT increase being additive to the set-size effect. This supports the RT model assumed for the present paradigm, in particular the assumption that the distraction responsible for the RT increase is independent of set size and thus additive to the other component processes of RT in the critical trial. In particular, the new color in the 25th trial mismatched expectations, and this incidence disrupted search and slowed RT. Because the expectancy discrepancy concerned the display as a whole and did not single out the position of the target, the set-size effect was not affected. Rather, participants had to resume the search for the target after the surprising color change had been processed.

Experiment 7: Form-Heterogeneous Nonsingleton Precritical Trials

Experiment 7 was designed to control for the possibility that the absence of attentional capture by the color singleton in Experiments 4–6 was due to unspecific heterogeneity of the displays in the precritical trials rather than to color heterogeneity. This possibility was already rendered implausible by Experiment 3. Experiment 3 revealed that presenting a form singleton in the precritical trials did not eliminate attentional capture by an unannounced color singleton in the critical trial, which implies that the expectation of variability does not strongly generalize over dimensions. However, attentional capture was not as strong in that experiment as in the experiments with entirely homogeneous displays in the precritical trials, as indicated by a significant difference of the set-size effects in the critical and the postcritical trials. To safeguard the conclusion that it is not heterogeneity per se that eliminates capture to an unannounced color singleton, I conducted the present Experiment 7. It combines features of Experiments 3 and 4, in that each display in the precritical trials was orientation heterogeneous, but unlike Experiment 3 the orientation-heterogeneous displays did not contain singletons. Rather, analogous to Experiment 4, the two orientations of the squares alternated in adjacent positions in a checkerboard fashion. The critical trial was the same as in Experiments 3–6—that is, all squares had the same orientation, but one had a different color.

Method

Participants. Participants were 14 men and 15 women with a mean age of 27.1 years ($SD = 8.1$). Compensation was €1.

Apparatus, stimuli, procedure, and design. These were the same as in the preceding experiment except for the composition of the precritical trials. In these trials, upright-oriented and 45°-rotated squares alternated in adjacent positions.

Results

Errors and RTs longer than 3,000 ms were excluded from RT analysis, resulting in the loss of 4.7% of the trials, one of them a critical trial. Mean RTs and error percentages are depicted in Figure 3 (panel f) and in Table 2, respectively. The set-size effect in the critical trial (28 ms/letter) was reduced relative to the precritical trials (73 ms/letter) but was somewhat larger than in the postcritical trials (3 ms/letter).

An overall ANOVA of the RTs yielded a significant main effect of trial type, $F(2, 52) = 28.3$, $p < .01$, and set size, $F(1, 26) = 17.7$, $p < .01$, and a significant Trial Type × Set Size interaction, $F(2, 52) = 9.8$, $p < .01$. A corresponding error analysis yielded a significant main effect for block only, $F(2, 52) = 11.3$, $p < .01$ (the other Fs < 1). That is, participants traded accuracy for speed in the critical trial (see Table 2). However, as in the previous experiments, such a trade-off is of minor importance for the present investigation, because the main prediction concerns the Trial Type × Set Size interaction.

An analysis including the precritical trials and the critical trial yielded a significant main effect of set size, $F(1, 26) = 21.1$, $p < .01$, and a significant Trial Type × Set Size interaction, $F(1, 26) = 5.0$, $p < .05$, reflecting a smaller set-size effect in the critical trial than in the precritical trials (224 ms vs. 585 ms). The main effect for trial type was not significant, $F(1, 26) = 1.2$.

When the analysis considered the critical trial and the postcritical trials, only a significant main effect for trial type was obtained, $F(1, 26) = 38.7$, $p < .01$, whereas the main effect of set size and the Set Size × Trial Type interaction were not significant (Fs $<$ 1.9).

Discussion

Presenting form-heterogeneous nonsingleton displays in the precritical trials gave results where the set-size effect reduction in the critical trial was between completely homogeneous displays (Experiment 1, homogeneous display condition) and form-heterogeneous displays containing a singleton (Experiment 3). More precisely, similar to Experiment 2, there was quite a strong reduction of the set-size effect in the critical trial relative to the precritical trials and no significant difference of the set-size effects in the critical trial and the postcritical trials. This result suggests that for a surprising color singleton to capture attention, color homogeneity in the precritical trials is of prime importance, whereas form heterogeneity is not. This result adds to the conclusions drawn from Experiments 1 and 3 that expectations are relatively specific.

However, as in Experiment 3, the set-size effect in the critical trial was somewhat larger than in Experiments 1 and 2, suggesting that form heterogeneity does have some influence (although not significantly different from the postcritical trials in this case). Thus, the data indicate that a feature-unspecific expectation of homogeneity is a second—although apparently less important—determinant of surprise capture, in addition to the feature-specific expectation of certain colors.

Experiment 8: Large Set Sizes

Experiment 8 returns to the original manipulation to induce surprise capture: the presentation of a color singleton after repeated presentations of color-homogeneous precritical trials (present Experiments 1 and 2; Horstmann, 2002). One possible objection to the previous experiments is that the pattern of results assumed to be indicative of surprise capture is not as simple as one might wish. In particular, although the predicted reduction of the set-size effect was obtained in the critical trial, the RTs in the critical trial were not always shorter than with serial search in the precritical trials, a result that is usually obtained with attentional capture. In the current experiments, a reduction in RT was obtained with set-size 12 but not with set-size 4, where RT increased. This pattern of results was explained with reference to the RT delay that is typically observed when a surprise stimulus is presented during an RT task.

If this explanation is valid, one should predict RT benefits in the critical trial as compared with the serial-search task in the precritical trials with larger set sizes. That is, if the set sizes are sufficiently large and RTs in the visual-search task are comparably long, then RTs should be shorter in the critical trial despite the "handicap" of additional processes concerned with the surprise stimulus. Thus, with sufficiently large set sizes, it should be possible to demonstrate an attentional shift to the surprise stimulus, exhibiting both characteristics of attention capture: a reduction of the set-size effect and an overall reduction of RTs in both the large and the small set-size conditions.

Method

Participants. Participants were 25 women and 11 men with a mean age of 25.3 years (*SD* = 4.3). They were paid €1 for their services.

Apparatus, stimuli, procedure, and design. These were basically the same as in the previous experiments, except that (a) a different monitor was used (a 17-in [43.18-cm] Sony), (b) the two set sizes tested contained 25 or 49 stimuli, and (c) the display duration was changed to 10.000 ms to take into account the increased difficulty of the serial search with the large set sizes.

Stimuli. Targets and distractors were the same as in the set-size 12 conditions of the previous experiments. In contrast to the previous experiments, however, the 11 distractors appeared several times in each display. More precisely, the distractor positions were filled by randomly drawing letters without replacement from sets of 11 distractors. The distractors were arranged in a 5 × 5 matrix for the set-size 25 condition and in a 7 × 7 matrix for the set-size 49 condition. The center-to-center distance of the letters was 1.5°. Similar to Experiments 1 and 2, the letters were themselves colored instead of being presented on colored patches, as in Experiments 3–7. Two color pairs were used: red–green, as in the previous experiments, and blue–yellow. Half of the participants had the letters presented in red or green, and the other half in blue or yellow. Which color of a pair served as the homogeneous display color and which served as the singleton color were balanced over participants.

Results

Errors and RTs longer than 8,000 ms were excluded from analysis, which concerned 7.9% of the experimental trials (although none in the critical trial). The RT criterion was adjusted in this experiment, because, as a result of the large set sizes, RTs were much longer in the present experiments. The criterion of 8,000 ms was chosen because it excluded about 2% of the RTs, which corresponds to the 3,000 ms criterion in the previous experiments. Figure 5 shows the mean RTs; error rates are given in Table 3. There was a large set-size effect in the precritical trials (42 ms/letter) but not in the critical trial (6 ms/letter) or in the postcritical trials (2 ms/letter; for the RT results, see Figure 5).

Table 3
Error rates (in percentages) for Experiment 8

Condition	Precritical	Critical	Postcritical
Set size 25	3.3	0.0	3.0
Set size 49	4.6	0.0	2.1

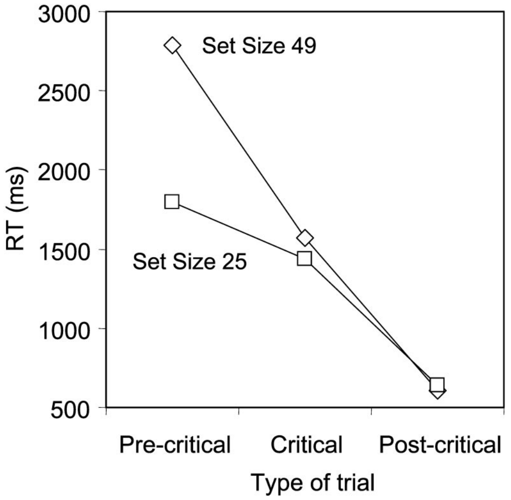

Figure 5. Mean correct reaction times (RTs) for both set-size conditions (set-size 25: diamonds; set-size 49: squares) in the precritical, critical, and postcritical trials of Experiment 8.

An overall ANOVA of the RTs yielded a significant main effect of trial type, $F(2, 52) = 91.8$, $p < .01$, and set size, $F(1, 26) = 6.3$, $p < .05$, and a significant Trial Type × Set Size interaction, $F(2, 52) = 9.98$, $p < .01$. (Including color pair [i.e., red–green vs. yellow–blue] in the analysis did not change the results, although this revealed shorter RTs with the yellow–blue color pair.) A corresponding error analysis yielded a significant main effect for block only, $F(2, 52) = 10.8$, $p < .01$ (the other *F*s < 1). That is, participants traded accuracy for speed in the critical trial. However, as in the previous experiments, such a trade-off is of minor importance for the present investigation, because the main prediction concerns the Trial Type × Set Size interaction.

The analysis considering the precritical trials and the critical trial yielded a significant main effect of set size, $F(1, 26) = 7.2$, $p < .01$, a significant main effect for trial type, $F(1, 26) = 37.1$, $p < .01$, and a significant Trial Type × Set Size interaction, $F(1, 26) = 11.0$, $p < .01$, reflecting a smaller set-size effect in the critical trial than in the precritical trials (1,008 ms vs. 138 ms). *t* tests revealed that RTs were shorter in the critical trial than in the precritical trials in the set-size 25 condition, $t(13) = 2.0$, $p < .05$, and in the set-size 49 condition, $t(13) = 6.5$, $p < .01$ (one-tailed).

When the analysis considered the critical trial and the postcritical trials, only a significant main effect for trial type was obtained, $F(1, 26) = 61.1$, $p < .01$, whereas the main effect of set size and the Set Size × Trial Type interaction were not significant (*F*s < 1).

Discussion

In agreement with the predictions, shorter RTs in the critical trial than in the precritical trials were obtained when large set sizes were tested. The remaining pattern of results is consistent with the previous experiments, revealing a large set-size effect in the precritical trials that was strongly reduced in the critical trial and in the postcritical trials. The present experiment thus shows that a "simple" pattern of results indicative of attentional capture (with both a reduction of the set-size effect and shorter RTs than with serial search) can be obtained for the first and unannounced presentation of a color singleton, when the set sizes used are sufficiently large.

General Discussion

The aim of the present study was to test the conditions of attentional capture to the first and unannounced presentation of a singleton during a visual-search task. The central hypothesis was that attention can be captured by an unannounced stimulus if the stimulus deviates from an expectation or schema and if the expectation concerns a stimulus aspect that is preattentively available. This hypothesis was tested with regard to expectations for the dimension of color. The experiments demonstrated that a singleton (e.g., a red patch among green patches) attracted attention on its very first presentation if its color was new and if it followed precritical trials without color variation (Experiments 1a, 2, 3, 7, and 8). In contrast, the color singleton did not capture attention if its color was not new but was repeatedly presented during the precritical trials (Experiments 4, 5, and 6). This result indicates that the expectation discrepancy is an important condition of attentional capture by an unannounced singleton.

Several results indicate that the expectancy discrepancy that triggers an attentional shift concerns the color of the singleton and not the presence of the singleton per se. Recall that in most experiments demonstrating surprise capture (e.g., present Experiments 1, 2, and 8), the singleton in the critical trial is the first object in the experiment with a given color, and it is the first singleton presented in the experiment. That is, both aspects could potentially be the instance that triggers attentional capture. Three results indicate that the color dimension is more important. First, a red or green singleton that followed red and green color-homogeneous displays in the precritical trials (Experiment 5 and 6) should have been discrepant from an expectation of a homogeneous display but not from expectations concerning color, yet the color singleton in the critical trial did not capture attention. This indicates that only when the first and unannounced presentation of a singleton was accompanied by the occurrence of an expectation-discrepant color did the singleton capture attention. Second, with the critical trial following form-heterogeneous trials (Experiment 3 and 7), attention was attracted by the color singleton (although attention capture was not as strong as in the most comparable experiment, Experiment 2, in which the precritical trials were both color and form homogeneous; this result is discussed shortly). Third, the unannounced color change within the (homogeneous) precritical trials in Experiment 6 induced an RT increase that was additive to the set-size effect. Although this pattern of results does not reveal a spatial shift of attention because the new color was an aspect of the whole display and thus did not single out the target's location, the RT increase indicates that the color change was attended to. It might be noted that distraction by a nonpredictive stimulus is an often used indicator of attentional capture (e.g., Theeuwes, 1992, 1994), although RT increases are susceptible to alternative explanations (e.g., Folk & Remington, 1998). In sum, evidence for attentional shifts covaried with the color change and not with changes from homogeneous to heterogeneous conditions.

Apart from testing the conditions of surprise capture, a second important aim of the present experiments was to test whether purely bottom-up-driven "singleton capture" can account for the capture effects in the present paradigm (e.g., Itti & Koch, 2000; Kim & Cave, 1999; Koch & Ullman, 1985; Theeuwes & Godijn, 2001; Wolfe, 1994). According to this account, singletons always capture attention, but participants quickly reorient away from the singleton if they believe that the singleton is task irrelevant. The results of Experiments 4–6 render this account unconvincing for the present paradigm. The checkerboard-like heterogeneous precritical trials in Experiment 4 did not contain singletons and thus cannot have induced an intention to ignore singletons, yet no attentional capture occurred with the first presentation of a singleton in the critical trial. Neither could the red and green homogeneous displays in the precritical trials of Experiments 5 and 6 induce an intention to ignore singletons, yet no attentional capture occurred in the critical trial. In the face of the very consistent results of Experiments 4–6, attentional capture by an unannounced singleton cannot be accounted for easily by purely bottom-up-driven singleton capture.

The present results allow some generalizations on the expectation mismatch that is assumed to enable or disable attentional capture in the present paradigm. First, expectations could be specific with respect to feature dimensions (e.g., color or form), as indicated by the fact that expecting form heterogeneity did not eliminate attentional capture by a color singleton in Experiments 3 and 7. Second, expectations could be quite specific with respect to feature values (e.g., that all stimuli are red), given that one feature value was consistently presented in the precritical trials. In contrast, if several feature values (e.g., red and green) were presented in the precritical trials, expectancies were broad and fuzzy within the respective feature dimension (cf. Schützwohl, 1998). Consequently, the introduction of a third color did not capture attention (Experiment 1A, heterogeneous display condition, and Experiment 1B). This result is reminiscent of Schützwohl's (1998) finding that an unannounced change in the physical appearance of a word was less surprising (as indicated by the amount of RT interference) if the word's font varied slightly in the precritical trials, as compared with a constant font in the precritical trials. Third, following even a single color change within homogeneous displays, the expected range of colors remained sufficiently broad over quite a number of trials that a singleton instantiating one of the previously presented colors did not capture attention. Fourth, expectations that led to surprise capture in the critical trial probably concerned basic features but not combinations of basic features (cf. Treisman & Gelade, 1980). This is not to say that combinations of features cannot be expectation discrepant; rather, this discrepancy would not be detected preattentively.

The present results support the notion that the detection of an expectancy discrepancy (including attention to the surprising objects) can result from preattentively and from attentively processed information. Because preattentive analysis provides separated but not integrated information (Treisman & Gelade, 1980), surprise capture can occur only if the presence of a feature is schema discrepant with respect to any part of the schema. That is, after repeated presentations of a green square and a green circle, a red square is discrepant because none of the elements of the schema is assumed to be red. In contrast, after repeated presentations of a green square and a red circle, a green circle is not schema discrepant on the basis of preattentively delivered feature information, because the color green is consistent with elements of the schema, and no surprise capture should occur. However, after the green circle has been attended to, the combination of form and color can be surprising and promote more attentive processing.

Some Methodological Considerations

As repeatedly noted before, other research on attentional capture with unannounced color singletons indicates that the attentional shift has a latency of about 300 ms. That is, within the present task, participants probably had already searched through one or two item positions when their attention was drawn to the color singleton. This implies that, in the present task, the color singleton was occasionally found during search and not by attentional capture. Further, this occurred more often with set-size 4 than with set-size 12. In the following, I discuss what implications this possibility has for the interpretation of the data, and I substantiate why this does not alter the implications of the present experiments.

A situation in which two independently running processes compete for the determination of some result is usually modeled as a race (e.g., Horstmann, 2003; Logan, 1994; Logan & Cowan, 1984). For the present situation, search processes and discrepancy-detection processes run in parallel and compete for the detection of the singleton. The search processes proceed serially, and the probability of finding the singleton increases with time. In contrast, the discrepancy-detection processes proceed in parallel and have a fixed singleton-detection time of about 300 ms. The time for the detection of the singleton in the critical trial is determined by the winner of the race, that is, by the process that succeeds first in detecting the singleton. Given that both—the surprise and the search processes—detect the singleton in a portion of the critical trials, the mean detection time is the weighted mean of the detection times of both processes. Because the probability that the singleton will be detected during search is higher with set-size 4 than with set-size 12, the weight given to the finishing time of the search processes is higher with set-size 4 than with set-size 12 (and, accordingly, the weight given to the finishing time of the discrepancy-detection processes is lower).

As a consequence, the mean singleton-detection time, and therefore the mean RT in the critical trial, is more affected by trials on which the singleton is found through search with set-size 4 than with set-size 12. The important point is that the mean singleton-detection time, and thus the mean RT in the critical trial, is shorter the more probable a detection through search is. This is because serial search beats discrepancy detection in the race only if it is faster.

For this reason, the possibility that the singleton was found during search, which is more probable in the set-size 4 than in the set-size 12 condition, cannot explain the reduction of the set-size effect in the critical trial. On the contrary, it can only explain the presence of a set-size effect with empirically obtained RTs, where, theoretically—on the basis of the assumption that mean RT is determined solely by discrepancy-detection processes but not by serial search in the critical trial—no set size effect is predicted. Thus, this possibility works against the presently tested hypothesis.

Alternative and Complementary Accounts

It may appear that habituation/sensitization processes can explain the presence and absence of surprise capture. That is, habituation of the response to the nonsingleton color could have occurred because of its repeated presentations, whereas sensitization of the response to the singleton color could have occurred because of the long time during which it was not used. This hypothesis, however, is difficult to reconcile with Experiment 6, because the critical trial was preceded by 24 homogeneous trials with a constant color, which should be sufficient to enable habituation/sensitization effects. That is, although it is possible that habituation/sensitization processes add to surprise capture in some instances, they are probably not the main mechanism.

The present investigation, in particular with respect to some of the reported effects, is reminiscent of the novel pop-out hypothesis that was discussed in the 1990s. According to this hypothesis, attention is rapidly and automatically drawn to a single novel item within a display of familiar items (e.g., Johnston, Hawley, Plewe, Elliott, & DeWitt, 1990; for a critical review see Christie & Klein, 1996). It is clear that the novel color in the critical trial presented along with familiar colors could be an example of novel pop-out, according to this definition. However, there are a number of differences from the research program of novel pop-out, both in method and in theory.

First, Johnston et al. (1990) used letter strings in their experiments that were either familiar in the experimental context (presented repeatedly during the experiment) or novel (presented for the first time). Four items were presented on each trial that could be either (a) all familiar, (b) all novel, or (c) mixed, with three familiar and one novel item. This four-word display was replaced by a probe display, in which one of the previously presented four words was repeated at the four positions that were previously occupied by the words in the four-word display (an intermediate mask of "*XXXXXX*" intervened between the four-word display and the probe display). The task was to indicate the position of the probe word in the four-word display. Novel pop-out was inferred from a localization advantage for the novel word ("the *novel pop-out effect*") and a localization disadvantage for the familiar words ("*familiar sink-in*"). A methodological and a corresponding theoretical difference from the present research are immediately obvious. In the novel pop-out paradigm, complex stimuli are used, and it is claimed that the unfamiliarity of complex stimuli is assessed preattentively, which is a controversial position. In contrast, the present experiments used simple stimulus features (e.g., color) for which preattentive processing is well documented in the visual search literature (e.g., Yantis & Egeth, 1999). Furthermore, although the present model assumes that attention is generally attracted by expectancy-discrepant events, attention capture as a spatial shift of attention following only a preattentive analysis is assumed to be possible only for discrepant simple features, not for complex configurations of features.

Second, the novel pop-out paradigm was designed specifically to test novel pop-out and has not been used by other researchers interested in attentional capture. In fact, as Christie and Klein (1996) pointed out, the particularities of the procedure used to assess novel pop-out render an unambiguous inference for fast attentional capture difficult. In contrast, the present experiments used one of the standard procedures to assess attentional capture. It might be noted that Christie and Klein's suggestion to avoid interpretational problems by using a probe task "that does not entail the use of identity or other information stored in memory about the probed items" (p. 207) was implemented in the present experiments. That is, the probe task (*H* vs. *U* discrimination) did not require the use of identity or other information about the probed items (the color patches).

Third and finally, the most compelling evidence for novel pop-out came from experiments that carried a confound: Because in these experiments the novel words were more likely the targets than the familiar words, attending to the novel items would be advantageous, thus undermining the claim of automatic capture. In contrast, the letter-search task was completely unrelated to the color of the patches in the precritical trials of the present experiments.

Surprise has been framed within a schema-theoretic framework (cf. Neisser, 1976; Rumelhart, 1984; Rumelhart, Smolensky, McClelland, & Hinton, 1986). According to this account, perception is essentially hypothesis testing. For example, Neisser's (1976) perceptual-cycle hypothesis assumes that observers have schemas or expectations for what belongs to the scene (i.e., which objects with what characteristics should be present), and these expectations are used to guide attention. Attention then picks up information from the scene, which, in turn, fleshes out or modifies the schema. That is, perception is viewed as a cyclic process of hypothesis building and testing. Neisser proposed that the incorporation into the perceptual cycle is essential for the information to be "seen" (see also Di Lollo, Enns, & Rensink, 2000; Most & Simons, 2001). That is, although some information outside the perceptual cycle can give rise to an orienting response, this information cannot acquire identity or meaning unless it is incorporated into the cycle of anticipation and information pick-up.

This model implies a fundamental difference between expected and unexpected events. Taking expected stimuli into consideration first, the organism is prepared to process them so that the meaning and the task relevance of each stimulus are readily available (see also Meyer et al., 1997). In contrast, the observer is not prepared to process unexpected stimuli. Rather, these stimuli have to be incorporated in a new perceptual cycle. What, then, determines whether they are incorporated? As I have said, one possibility is that certain stimulus conditions kindle an orienting response that triggers a new perceptual cycle. Among these conditions are probably high-intensity stimuli (loud noise, bright light, heavy touch) and perhaps the appearance of new objects and highly salient stimuli (cf. Most & Simons, 2001; remember, however, that the present results are at odds with the view that salience alone is sufficient to capture attention).

From the present theoretic viewpoint, experiments on surprise capture suggest a complement to this model: that expectation discrepancies are another condition to interrupt the ongoing perceptual cycle and instigate a new one. Note that schema discrepancy is apparently sufficient to trigger a new perceptual cycle: In Experiment 6, the color change between homogeneous displays was quite a strong distractor from the primary task activity (as indicated by the RT increase), although the new color was not a salient event but solely expectation discrepant. For the goal of extending the model, it suffices to assume that the testing of expectancies is not necessarily confined to attention but can proceed autonomously on both preattentive and attentive information.

Should the Attentional Shift to the Unannounced Singleton Be Called Attentional Capture?

Attentional capture and intentional orienting have inherited the meaning components of their respective parent concepts, being automatic versus voluntary processes (Jonides, 1981). That is, attentional capture is characterized by its independence from intentions, by its freedom from central capacity limitations, and by its reliance on nonconscious processing (Posner & Snyder, 1975). Speed is also often considered an important characteristic of automatic processes. The attentional shift to the unannounced singleton is independent from intentions by definition, in the sense that the capturing item is not related to the task performed by the observer (e.g., Yantis, 1993). Although the present experiments do not inform about the issue of conscious processing, the attention shift does not appear to be strongly dependent on central capacity mechanisms, as indicated by the rather weak influence of set size on its latency. A problematic aspect is probably speed or, rather, the lack of it. Experiments using the accuracy paradigm (Horstmann, 2002, 2004, in press) have characterized the effect as slower than, for example, the fast attentional shift to a singleton expected to be predictive of the target.[3] If speed is considered the most important criterion for attentional capture, then an attentional shift to an unannounced singleton will not be regarded as an instance of attentional capture. However, I argue that the most important criterion for attentional capture is that it occurs independently of the intentions in the sense that the capturing item is independent from the task performed by the observer (e.g., Yantis, 1993), and this clearly holds for the attention shift to the new color presented without prior warning.

Final Remarks

Expectation-discrepant events in the real world are certainly not always as simple as in the present experiments. Thus, surprise capture reveals a boundary condition of the working of a more general mechanism that orients attention to schema-discrepant stimuli. Frequently, however, expectation discrepancies are detected after attention has been directed to various objects in the scene in the course of schema testing or task pursuance. More precisely, schema discrepancies often are defined not by preattentively available basic features but rather by the identity or category-membership information that is available only after the deployment of attention. Readers should bear this in mind when generalizing the present results.

The present results also suggest a new look at the classical dichotic-listening experiments of Cherry (1953). Cherry presented participants with different auditory streams at each ear. The task was to shadow one of the streams, which always contained speech, while ignoring the other stream. When both streams contained speech, participants had some knowledge of the shadowed stream but very poor knowledge of the ignored stream. Cherry conducted

---

[3] One may even suggest that the latency of surprise capture is in the magnitude typical for endogenous shifts of attention. However, this hypothesis is difficult to reconcile with the fact that an endogenous shift presumes a reason, and because the new color is unrelated to the task, there is no reason to attend to it. For a more thorough discussion of this point, see Horstmann (in press).

several experiments to explore what kinds of changes in the ignored stream participants attended to. He found that physical changes of the material, such as a change from a male to a female voice or the change from speech to a pure 400-Hz tone were recognized by nearly all of the participants. In contrast, reversing the male speech (an operation that retains spectrum) or switching to another language were rarely discerned. Because Cherry tested naive participants who were not informed about possible changes, there are fairly strong parallels between Cherry's and the present experiments, in that stimuli that broadly shared physical properties with the known properties of the to-be-ignored stream did not receive attention (language change or speech reversal), whereas stimuli that deviated strongly from the anticipated physical properties apparently captured attention (pitch change introduced by the female speaker or pure tone). This suggests a common explanation for these two phenomena from the visual and the auditive domain: A mismatch between expectations and basic feature properties promotes an orientation of attention toward the unexpected stimulus.

References

Ansorge, U., & Heumann, M. (2003). Top-down contingencies in peripheral cuing: The roles of color and location. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 937–948.

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

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25,* 975–979.

Christie, J., & Klein, R. M. (1996). Assessing the evidence for novel popout. *Journal of Experimental Psychology: General, 125,* 201–207.

Darwin, C. (1872). *The expression of the emotions in man and animals.* London: Murray.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129,* 481–507.

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,* 847–858.

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.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1993). Contingent attentional capture: A reply to Yantis (1993). *Journal of Experimental Psychology: Human Perception and Performance, 19,* 682–685.

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 displaywide visual features. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 699–706.

Gibson, B. S., & Peterson, M. A. (2001). Inattentional blindness and attentional capture: Evidence for attention-based theories of visual salience. In C. Folk & B. Gibson (Eds.), *Attraction, distraction, and action: Multiple perspectives on attentional capture* (pp. 51–76). New York: Elsevier Science.

Horstmann, G. (2001). Die Unterbrechungsfunktion der Überraschung: Ein neues experimentelles Paradigma und ein Test der Automatizitätshypothese. [Unveröffentlichte Dissertation, Universität Bielefeld.] [The interrupt-function of surprise: A new experimental paradigm and a test of the automaticity hypothesis]. [Unpublished doctoral dissertation, Bielefeld University]. Universität Bielefeld.

Horstmann, G. (2002). Evidence for attentional capture by a surprising color singleton in visual search. *Psychological Science, 13,* 499–505.

Horstmann, G. (2003). The psychological refractory period of stopping. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 965–998.

Horstmann, G. (2004). *Predictive and non-predictive surprising singletons.* Manuscript submitted for publication.

Horstmann, G. (in press). The time course of intended and unintended allocation of attention. *Psychologische Forschung.*

Horstmann, G., & Ansorge, U. (in press). Attentional capture by rare singletons. *Visual Cognition.*

Horstmann, G., & Schützwohl, A. (1998). Zum Einfluss der Verknüpfungsstärke von Schemaelementen auf die Stärke der Überraschungsreaktion [Strength of interconnections between schema elements and intensity of the emotion of surprise]. *Zeitschrift für Experimentelle Psychologie, 45,* 203–217.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40,* 1489–1506.

Johnston, W. A., Hawley, K. J., Plewe, S. H., Elliott, J. M. G., & DeWitt, M. J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General, 119,* 397–411.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. Long & A. D. Baddeley (Eds.), *Attention & performance IX* (pp. 187–203). Hillsdale, NJ: Erlbaum.

Kim, M. S., & Cave, K. R. (1999). Grouping effects on spatial attention in visual search. *Journal of General Psychology, 126,* 326–352.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology, 4,* 219–227.

Krummenacher, J., Müller, H. J., & Heller, D. (2002). Visual search for dimensionally redundant pop-out targets: Parallel-coactive processing of dimensions is location specific. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1303–1322.

Logan, G. D. (1994). On the ability to inhibit thought and action. A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91,* 295–327.

Mack, A., & Rock, I. (1998). *Inattentional blindness.* Cambridge, MA: MIT Press.

Mandler, G. (1985). *Cognitive psychology: An essay in cognitive science.* Hillsdale, NJ: Erlbaum.

Mandler, J. M. (1984). *Stories, scripts, and scenes: Aspects of schema theory.* Hillsdale, NJ: Erlbaum.

Meyer, W.-U. (1988). Die Rolle von Überraschung im Attributionsproze [The role of surprise in the attribution process]. *Psychologische Rundschau, 39,* 136–147.

Meyer, W.-U., & Niepel, M. (1994). Surprise. In V. S. Ramachandran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 353–358). Orlando, FL: Academic Press.

Meyer, W.-U., Niepel, M., Rudolph, U., & Schützwohl, A. (1991). An experimental analysis of surprise. *Cognition and Emotion, 5,* 295–311.

Meyer, W.-U., Reisenzein, R., & Schützwohl, A. (1997). Toward a process analysis of emotions: The case of surprise. *Motivation and Emotion, 21,* 251–274.

Most, S. B., & Simons, D. J. (2001). Attention capture, orienting, and awareness. In C. Folk & B. Gibson (Eds.), *Attraction, distraction, and* *action: Multiple perspectives on attentional capture* (pp. 151–173). Amsterdam: Elsevier Science.

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2001). How not to be seen: The contribution of similarity and selective ignoring to sustained IB. *Psychological Science, 12,* 9–17.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology.* San Francisco: Freeman.

Neisser, U. (1979). The control of information pickup in selective looking. In A. D. Pick (Ed.), *Perception and its development: A tribute to Eleanor J. Gibson* (pp. 201–219). Hillsdale, NJ: Erlbaum.

Niepel, M., Rudolph, U., Schützwohl, A., & Meyer, W.-U. (1994). Temporal characteristics of the surprise reaction induced by schema-discrepant visual and auditory events. *Cognition and Emotion, 8,* 433–452.

Nothdurft, H.-C. (2000). Salience from feature contrast: Variations with texture density. *Vision Research, 40,* 3181–3200.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Prinz, W. (1983). *Wahrnehmung und Tätigkeitssteuerung.* Berlin: Springer.

Prinz, W. (1990). Unwillkürliche Aufmerksamkeit. In C. Meinecke & L. Kehrer (Eds.), *Bielefelder Beiträge zur Kognitionspsychologie* (pp. 49–57). Göttingen: Hogrefe.

Rumelhart, D. E. (1984). Schemata and the cognitive system. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 161–188). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 99–135). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In D. E. Rumelhart, J. L. McClelland, & PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2, pp. 7–57). Cambridge, MA: MIT Press.

Schützwohl, A. (1993). Schemata und Überraschung: Untersuchungen zum Zusammenwirken von Kognition und Emotion. Unveröffentlichte Dissertation, Universität Bielefeld [Schema and surprise: Investigations concerning the interaction of cognition and emotion. Unpublished doctoral dissertation, Bielefeld, University]. Universität Bielefeld.

Schützwohl, A. (1998). Surprise and schema strength. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1182–1199.

Selz, O. (1922). *Zur Psychologie des produktiven Denkens und des Irrtums: eine experimentelle Untersuchung.* [The psychology of productive thinking and of error: An experimental study.] Bonn, Germany: Cohen.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28,* 1059–1074.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30,* 276–315.

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

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.

Theeuwes, J., & Burger, R. (1998). Attentional control during visual search: The effect of irrelevant singletons. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1342–1353.

Theeuwes, J., & Godijn, R. (2001). Attentional and oculomotor capture. In C. Folk & B. Gibson (Eds.), *Attraction, distraction, and action: Multiple perspectives on attentional capture* (pp. 121–149). Amsterdam: Elsevier Science.

Todd, S., & Kramer, A. F. (1994). Attentional misguidance in visual search. *Perception & Psychophysics, 56,* 198–210.

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

Wilcocks, R.-W. (1928). The effect of an unexpected heterogeneity on attention. *Journal of General Psychology, 1,* 286–319.

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

Wolfe, J. M., O'Neill, P., & Bennett, S. C. (1998). Why are there eccentricity effects in visual search? Visual and attentional hypotheses. *Perception & Psychophysics, 60,* 140–156.

Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 676–681.

Yantis, S., & Egeth, H. E. (1999). On the distinction between visual salience and stimulus-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 661–676.

Received December 12, 2003
Revision received October 29, 2004
Accepted March 29, 2005 ■