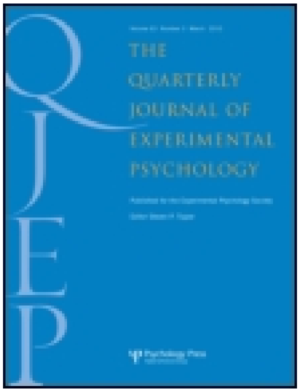

The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Colour and spatial cueing in low-prevalence visual search

Nicholas C. C. Russell [a] & Melina A. Kunar [a]
[a] Department of Psychology, The University of Warwick, Coventry, UK
Accepted author version posted online: 13 Jan 2012.Version of record first published: 12 Apr 2012.

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2012, 65 (7), 1327–1344

Psychology Press
Taylor & Francis Group

Colour and spatial cueing in low-prevalence visual search

Nicholas C. C. Russell and Melina A. Kunar

Department of Psychology, The University of Warwick, Coventry, UK

In visual search, 30–40% of targets with a prevalence rate of 2% are missed, compared to 7% of targets with a prevalence rate of 50% (Wolfe, Horowitz, & Kenner, 2005). This "low-prevalence" (LP) effect is thought to occur as participants are making motor errors, changing their response criteria, and/or quitting their search too soon. We investigate whether colour and spatial cues, known to improve visual search when the target has a high prevalence (HP), benefit search when the target is rare. Experiments 1 and 2 showed that although knowledge of the target's colour reduces miss errors overall, it does not eliminate the LP effect as more targets were missed at LP than at HP. Furthermore, detection of a rare target is significantly impaired if it appears in an unexpected colour—more so than if the prevalence of the target is high (Experiment 2). Experiment 3 showed that, if a rare target is exogenously cued, target detection is improved but still impaired relative to high-prevalence conditions. Furthermore, if the cue is absent or invalid, the percentage of missed targets increases. Participants were given the option to correct motor errors in all three experiments, which reduced but did not eliminate the LP effect. The results suggest that although valid colour and spatial cues improve target detection, participants still miss more targets at LP than at HP. Furthermore, invalid cues at LP are very costly in terms of miss errors. We discuss our findings in relation to current theories and applications of LP search.

Keywords: Visual search; Low prevalence.

In everyday life, people frequently perform visual search tasks, such as searching for a car in a car park or keys on a cluttered coffee table. Scientists study how well people perform these search tasks, in the laboratory, by asking participants to search for a prespecified target among distractor items and recording their reaction times (RTs) and error rates. Although RTs can vary depending on the difficulty of the search display (e.g., Treisman & Gelade, 1980), for the most part, error rates are low (typically less than 10%; Wolfe, 1998). Recent research has shown that the frequency, or prevalence, of the target has an interesting effect on the number of errors made. Wolfe, Horowitz, and Kenner (2005) found that miss errors (stating that a target is absent when it is physically present) rose from 7% to 30% when the prevalence rate of the target dropped from 50% ("high prevalence") to 2% ("low prevalence"). This increase in miss errors with a decrease in target prevalence is known as the low-prevalence (LP) effect. The LP effect has been replicated on several occasions with both complex and simple stimuli (e.g., Fleck & Mitroff, 2007; Kunar, Rich & Wolfe, 2010;

Correspondence should be addressed to Melina A. Kunar, Department of Psychology, The University of Warwick, Coventry, CV4 7AL, UK. E-mail: m.a.kunar@warwick.ac.uk

http://www.psypress.com/qjep
http://dx.doi.org/10.1080/17470218.2012.656662

Rich et al., 2008; Van Wert, Horowitz, & Wolfe, 2009; Wolfe et al., 2007) and is also found to be robust across a large number of experimental designs (e.g., Kunar et al., 2010; Rich et al., 2008; Wolfe et al., 2007; however, see Navalpakkam, Koch, & Perona, 2009, who found increasing a participant's expected reward outcome reduced the LP effect).

Several theories have been put forward for why the LP effect occurs. Wolfe et al. (2007) found that RTs were reduced at low prevalence. On face value, this could appear to indicate a speed–accuracy trade-off, where participants were merely responding too quickly without appropriately searching for the target. Some support to this account is shown by Fleck and Mitroff (2007) who introduced a correction facility to their experiments, in which participants were able to correct any mistakes during a previous trial. The aim of this manipulation was to help correct the errors of participants who may have either accidentally entered the wrong response or identified the target too late to suppress their "absent" reaction (indicating a motor error). Indeed, Fleck and Mitroff obtained a low-prevalence effect before the correction was included but this effect was removed when participants were given the option to correct their mistakes (although, see Kunar et al., 2010; Van Wert et al., 2009, who found that the LP effect was reduced but not removed after self-correction). These results were attributed to the elimination of "errors in action"—finding the target but responding hastily (Fleck & Mitroff, 2007, p. 943)—and not "errors of perception"—simply failing to find a target at low prevalence (Fleck & Mitroff, 2007, p. 943).

Although part of the LP effect may occur due to these motor errors, further research suggests that the LP effect cannot be entirely explained by a simple speed–accuracy trade-off. Wolfe et al. (2007) slowed participants' responses by alerting them if they were responding too quickly. If the increase in miss errors was due to participants responding prematurely, then asking participants to slow down should eliminate the LP effect. However, asking participants to respond more slowly did not improve target detection. Neither did forcing participants to wait a minimum period before responding (Rich et al., 2008). In complex search tasks, a simple speed–accuracy trade-off account cannot explain the full range of data.

Another way to examine the LP effect is with signal detection theory (SDT). SDT uses hit rates and false alarms to calculate participants' sensitivity (d') and decision criterion (c) to better understand why the LP effect occurs. At low prevalence, people could either be less sensitive to target detection or, with a vast increase in target-absent trials, less willing to decide that a display contains a target. Using this measure, Wolfe et al. (2007) and Van Wert et al. (2009) showed little effect of prevalence on sensitivity but instead found that, with low-prevalence targets, participants adopted a more conservative criterion. That is, when the target is rare, observers are more likely to adopt a criterion that is biased to respond to a display as "target absent". Wolfe and Van Wert (2010) have since proposed a two-stage model of LP search. In this account, they suggest that under low-prevalence conditions, both the decision criteria and the quitting threshold change so that participants are less willing to respond that an item is a target and also terminate their search sooner than when the target has a high prevalence.

Investigating low-prevalence search is important as it occurs in a number of applied settings—for example, detecting tumours in mammograms (Fenton et al., 2007; Leung et al., 2007) and detecting prohibited items at airport security checkpoints (e.g., Rubenstein, 2001). Taking this latter search task, in order to try and minimize miss errors, computer algorithms are applied to screened images to help officers identify targets ("screener assistant technology"; Parks, 2007). One technique assigns particular colours to items made of different materials, with the premise that screeners will be able to identify certain items more easily. For example, metallic objects appear in blue (Evans, Wang, Chan, Downes, & Liu, 2006). Thus, when searching for a knife, screeners restrict their search to the blue items. Another technology cues specific areas that are likely to contain a threat and which the screener should pay attention to

(Bretz, 2002; Guardian Technologies International, 2008). This cue takes the form of an exogenous red box which outlines potentially suspicious items.

Intuition would suggest that these cues have a beneficial effect on LP search. For example, colour is known to be a useful tool in visual search. Egeth, Virzi, and Garbart (1984) and Wolfe, Cave, and Franzel (1989) found that people use colour to guide their search to a subset of items and away from distracting stimuli. Colour cues can also help guide saccades (e.g., Hannus, van den Berg, Bekkering, Roerdink, & Cornelissen, 2006; Luria & Strauss, 1975; Motter & Belky, 1998; Williams & Reingold, 2001), and the use of colour has been implemented in a number of search models (e.g., Itti & Koch, 2000; Wolfe, 2007; Zelinsky, 2003, 2008). Michel, Koller, Ruh, and Schwaninger (2007) also found that colour is an effective tool for baggage screening tasks at high prevalence. Given this evidence, one would think that adding colour to the display may help participants find LP targets. However, other evidence suggests that this may not necessarily be the case. One of the reasons why colour improves target detection is that it, in part, increases sensitivity (or d') to that target (see Eckstein, Thomas, Palmer, & Shimozaki, 2000; Verghese, 2001, for evidence of signal detection in visual search). However, if the low-prevalence effect is a result of a criterion shift rather than a change in d' (as suggested by Wolfe et al., 2007), then adding colour to the display may not help in LP conditions. We investigate this in Experiment 1.

Experiments 2 and 3 investigate how LP search is affected by both valid and invalid cues. The algorithms used in screener assistant technologies are designed to highlight potential target colours or locations; however, they are fallible in that the target may appear in an unexpected colour (for example, a nonmetal knife would not appear in the colour blue) or outside the cued area. Furthermore, the algorithm may not detect the presence of a target at all. From previous visual search tasks, we know that validly cueing the target location improves target detection (e.g., Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000; Posner, 1980; Posner & Cohen, 1984). This is particularly true of exogenous cues, which are thought to automatically cue a target's location regardless of the goal state of the participant (e.g., Jonides, 1981; Remington, Johnston, & Yantis, 1992; Theeuwes, 1992). Invalid cues, however, may introduce a cost to search behaviour (Corbetta et al., 2000; Posner, 1980; Posner & Cohen, 1984). The effect of invalid cues becomes even more important at low prevalence as the potential cost of missing a rare target in an applied setting has serious implications (e.g., missing a weapon in screened baggage at an airport security checkpoint). Experiments 2 and 3 investigate the benefits of having the target appear in an expected colour or being cued by a valid exogenous cue (a red box), while also recording the costs when these cues are invalid. To preview the results, both colour and exogenous cues aid target detection if they are valid; however, miss errors are dramatically increased at LP on invalidly cued trials.

Of final note, the experiments here investigate how effective valid cues are under LP conditions. Previous work has suggested that search behaviour under low-prevalence conditions differs from that under high-prevalence conditions. For example, visual cues separating a search display across space and time, often helpful in search at high prevalence, fail to benefit search at low prevalence (Kunar et al., 2010). Furthermore, features that are trivially easy to find at high prevalence are missed more often at low prevalence (Rich et al., 2008). This then leads to the interesting question: Are cues that are useful at HP also useful at LP? There are three possible outcomes. First, colour and spatial cues that help HP search do not provide any benefit under LP conditions. There is some precedent in this as other experiments have shown that both temporal and spatial cues do little to reduce the LP effect (Kunar et al., 2010). Please note, however, that the temporal/spatial cues used by Kunar et al. (2010) only served to reduce the area to be searched at any given time and so were not used to highlight the target's location/features. Second, having the target appear either in a known colour or in a cued area will eliminate the LP effect, so that miss errors at LP are the same as those at HP. For example, as the red box cue in Experiment 3 is highly salient and captures attention, then any target falling inside the cue will be found with equal ease, regardless of its prevalence. Third, having the target appear in a cued area or colour will reduce the LP effect, but not eliminate it. In this case, search performance would be improved with the addition of the cue but miss errors would still be higher at LP than at HP. Here, the benefit of cueing the target would be constrained by target prevalence. To preview the results further, the data suggest that this latter option is correct. These results have implications for the use of cues in low-prevalence search.

EXPERIMENT 1

Experiment 1 investigated whether searching for a target of a known colour improved target detection within LP search. That is, do colour cues, which aid search at high prevalence, also benefit search when the target has a low prevalence?

Method

Participants
Ten participants ($M = 21.9$ years, $SD = 0.86$, 4 female) were recruited from the University of Warwick. Each received £6/hour remuneration. Informed consent was obtained, and all had normal or corrected-to-normal vision.

Stimuli and procedure
The experiment was generated using custom written programs and was presented on a PC. The stimuli were rotated Ts and Ls. Each stimulus had a visual angle of 1.7° × 1.7° at a viewing distance of 57 cm, and the vertical lines of the Ls were slightly offset from its horizontal line (see Figure 1). All stimuli were presented on a grey background. On each trial, there were always 12 stimuli presented (on "target-present" trials—11 distractors and 1 target; on "target-absent" trials —12 distractors).

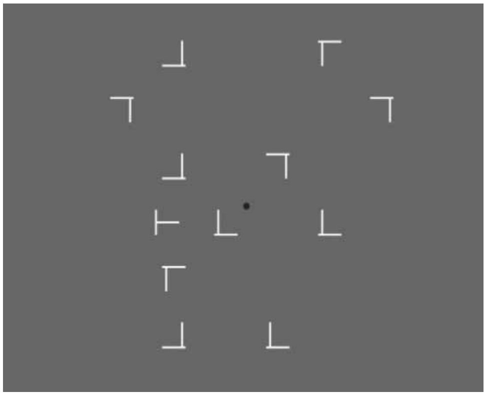

Figure 1. *Example of a target-present display in Experiment 1. Participants searched for the "T" (here, rotated 90° anticlockwise).*

There were four blocks of trials: two high-prevalence conditions—one with white stimuli and one with colour stimuli—and two low-prevalence conditions—one for white stimuli and another for colour stimuli. Participants completed all four conditions in one session. In the colour blocks, four of the stimuli were red, four were green, and four were blue. For each participant, in the colour blocks, the colour of the target was always the same and was told to the participant before the experiment. The order of the blocks was randomized, as was the assigned target colour each participant was given in the colour blocks.

To begin each trial, a blank screen appeared for 500 ms and was followed by a central fixation point for 500 ms. Stimuli were then presented randomly within an invisible 6 × 6 matrix. The target was a "T" and was presented on either 2% (low prevalence) or 50% (high prevalence) of trials. The distractor items were offset L shapes. All stimuli were presented randomly in one of four orientations (0°, 90°, 180°, or 270°) with equal probability. Participants were asked to indicate whether the target was "present" or "absent" by pressing either the "m" or the "z" key, respectively. Participants were instructed to respond as quickly but as accurately as possible. If no response was made within 10 s, the trial "timed out", and the next trial automatically started. Following a response or "timeout", the blank screen was again displayed before the next fixation point and trial.

In line with Fleck and Mitroff (2007), and to correct for motor errors, participants had the option of correcting their response to a trial during the succeeding trial. If a participant recognized that they had made an error, they were able to correct it on the following trial, by pressing the "Escape" key. They would then proceed with the current trial as normal. No feedback was given after any response or correction was made.

For each of the high-prevalence conditions, where the target was present 50% of the time, there were 80 trials (40 present and 40 absent). For each of the low-prevalence conditions, where the target was present 2% of the time, there were 1,000 trials (20 present and 980 absent). Participants were given a block of 20 practice trials before each experimental block. RTs and error rates (both the initial error made and the corrected errors, if participants pressed the "Escape" key) were recorded. Breaks were given every 200 trials

Results and discussion

RTs less than 200 ms and greater than 5,000 ms were removed as outliers. This led to the removal of 1% of the data at high prevalence and 1.63% of the data at low prevalence. The remaining error rates were arcsine transformed (both here and in all subsequent experiments) prior to analysis to compensate for unequal variances present in binomial data (Hogg & Craig, 1995). Values reported and plotted in the figures are the back-transformed means. Following the trend in the literature, false-alarm rates (where participants responded "target present" when the target was absent) were low (less than 1.6% in all experiments) and are not reported further. There are a number of possible statistics that we could report but to save journal space we only report the findings related to the question at hand. Our primary interest was what happens to miss errors at low prevalence. Miss errors, for Experiment 1, are shown in Figure 2. Independent t tests showed that there was no difference depending on which colour target participants searched for (all ts < 2.4, ps $> .07$). As such, all data were pooled for analysis.

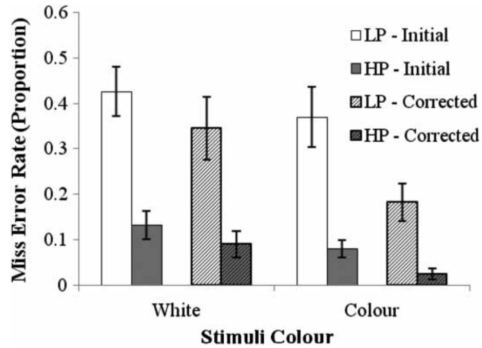

Figure 2. *Proportion of initial and corrected miss errors for high prevalence (HP, 50%) and low prevalence (LP, 2%) for both the white and the colour stimuli of Experiment 1. Error bars represent the standard error.*

Errors

Initial errors. A repeated measures analysis of variance (ANOVA) on the initial miss errors, with factors of prevalence (high or low) and colour (white or coloured), showed that there was a main effect of prevalence, where there were more miss errors at low prevalence (LP) than at high prevalence (HP), $F(1,\ 9) = 49.67$, $p < .01$. There was no main effect of colour, neither was there a significant Prevalence × Colour interaction.

Corrected errors. Looking at the miss errors after participants were given the option to self-correct, there was again a main effect of prevalence. $F(1,\ 9) = 22.75$, $p < .01$, and a main effect of colour, $F(1,\ 9) = 9.41$, $p < .05$. There were more miss errors at LP than at HP, and there were fewer miss errors in the colour condition than in the white condition. The Prevalence × Colour interaction was not significant.

The option to self-correct reduced the number of observed miss errors. Error rates were lower after correction in all conditions (ts > 2.7, ps $< .03$). For example, looking at the difference in LP trials, miss errors were reduced by 8% in the white condition and 19% in the colour condition after self-correction. This reduction suggests that some of the LP miss errors were likely to be due to errors in motor responses (Fleck & Mitroff, 2007). However, although the LP effect was reduced after self-correction, it was not eliminated, as miss rates were still higher in the LP conditions than in the HP conditions (26% vs. 6%, respectively). This supports previous findings that, with nontrivial searches, miss errors arise from more than just motor errors (Kunar et al., 2010; Rich et al., 2008; Van Wert et al., 2009).

Sensitivity (d′) and response criterion (c)

"False alarm" and "hit" data were used to calculate whether the LP effect occurred due to a change in d' (reflecting a change in sensitivity) or c (reflecting a criterion shift). Table 1 shows the d' and c values for each condition.

Initial d′. A repeated measures ANOVA on the initial d', with factors of prevalence (high or low) and colour (white or coloured), showed that there was no main effect of prevalence nor a main effect of colour. The Prevalence × Colour interaction was not significant.

Corrected d′. Analysis of the corrected d' data showed that there was no main effect of prevalence ($F < 2.44$, $p = .15$). However, there was a main effect of colour, $F(1,\ 9) = 7.27$, $p = .025$. d' was greater in the colour condition than in the white condition. There was no Prevalence × Colour interaction. Sensitivity in search was greater overall when participants could use colour to help find the target.

Initial criterion. A repeated measures ANOVA on the initial criterion, with factors of prevalence (high or low) and colour (white or coloured) showed there to be a main effect of prevalence, $F(1,\ 9) = 260.70$, $p < .001$. Similar to previous work, criterion was more conservative at LP than at HP (see also Van Wert et al., 2009; Wolfe et al., 2007). There was no main effect of stimuli colour and no significant Prevalence × Colour interaction.

Corrected criterion. Analysis of the corrected criterion data showed that there was a main effect of prevalence, $F(1,\ 9) = 50.27$, $p < .001$, where participants were more conservative at LP than at HP (see also Van Wert et al., 2009; Wolfe et al., 2007). There was also a main effect of stimuli colour, $F(1,\ 9) = 9.21$, $p = .014$, where criterion was more liberal for the colour condition than for the white condition. The Prevalence × Colour interaction was not significant.

The results of Experiment 1 concur with previous findings in the literature. Overall, there was little effect of prevalence on d'. However, criterion became more conservative at LP (see also Van Wert et al., 2009; Wolfe et al., 2007). In addition, there was a beneficial effect of colour on d' and c. Participants were more sensitive to detection of a coloured target than a noncoloured target and more willing to respond target present in the colour condition. This benefit in d' and criterion for colour stimuli supports the general understanding that colour aids response to a target in visual search (Egeth et al., 1984; Treisman & Gelade, 1980; Wolfe et al., 1989).

RTs

Reaction times for correct trials were analysed (based on a participant's initial response) and are

Table 1. Means and standard deviations for initial and corrected sensitivity and criterion, for high-prevalence and low-prevalence trials in Experiment 1

	Initial				Corrected			
Experimental condition	d′	SD	c	SD	d′	SD	c	SD
High prevalence/white stimuli	3.32	0.71	0.39	0.27	3.56	0.79	0.27	0.23
Low prevalence/white stimuli	3.42	0.50	1.51	0.25	3.33	0.65	1.17	0.47
High prevalence/colour stimuli	3.59	0.55	0.28	0.26	4.20	0.43	0.11	0.14
Low prevalence/colour stimuli	3.47	0.72	1.34	0.32	3.71	0.76	0.84	0.30

Note: d' = sensitivity. c = criterion.

shown in Table 2. As previously reported by Wolfe et al. (2007), RTs for correct target-absent trials (correct rejections) were considerably quicker at low prevalence. A repeated measures ANOVA on RTs for correct rejections, with factors of prevalence (high or low) and colour (white or coloured), showed a main effect of prevalence, $F(1, 9) = 16.30$, $p < .01$, where RTs were significantly quicker at LP than at HP, and a main effect of colour, $F(1, 9) = 65.41$, $p < .01$, where RTs for the colour condition were faster than those for the white condition. The Prevalence × Colour interaction was also significant, $F(1, 9) = 11.11$, $p < .01$. There was a greater difference in RTs between the HP and LP conditions for the white condition than for the colour condition.

In contrast, RTs for "hits" (correct target-present trials) were slower at LP than at HP, as indicated by a main effect of prevalence, $F(1, 9) = 8.64$, $p < .05$. This is similar to the data reported by Wolfe et al. (2007). Furthermore, there was a main effect of colour, $F(1, 9) = 28.88$, $p < .01$. RTs for the colour stimuli were quicker than those for the white stimuli. There was no Prevalence × Colour interaction. The reaction time data for Experiment 1 followed the general pattern found in the LP literature, where there is a speeding of target-absent trials at LP. The same general pattern was also found in Experiments 2 and 3. As such, although we report these data in the following experiments, we do not discuss them further.

Table 2. Mean reaction times for high-prevalence and low-prevalence trials in Experiments 1, 2, and 3

Experiment	*Experimental condition*	*Hits*	*Correct rejections*
Experiment 1	HP/white stimuli	1,530	2,641
	LP/white stimuli	1,761	1,897
	HP/colour stimuli	1,021	1,351
	LP/colour stimuli	1,381	1,181
Experiment 2	HP/majority colour	1,135	1,715
	LP/majority colour	1,477	1,280
	HP/minority colour	1,260	n/a
	LP/minority colour	1,376	n/a
Experiment 3	HP/TBox	903	n/a
	LP/TBox	1,232	n/a
	HP/DBox	1,733	n/a
	LP/DBox	1,940	n/a
	HP/NB	1,409	2,131
	LP/NB	1,680	1,658
	HP/Box–no target	n/a	2,459

Note: HP = high prevalence. LP = low prevalence. TBox trials: target located within the box. DBox trials: target present but outside the box. Box–no target trials: target absent, and the box appeared in a random position on the screen. NB = no box. n/a = not applicable. Reaction times in milliseconds. For Experiment 2, as there could be no minority or majority colour distinction on target-absent trials, the correct rejection reaction time reflects the overall reaction time for all target-absent HP and LP trials.

To investigate whether the elevated miss errors at LP were due to a speed–accuracy trade-off, we calculated the difference between HP and LP RTs for correct target-absent trials. This RT difference provided a measure of speed by showing how much faster participants were at LP than at HP (the larger the difference in RT, the more speeded the responses at LP; see also Kunar et al., 2010). This difference was then correlated with the miss errors at LP, to examine whether miss errors increased with speed. Two separate correlations for the white condition and the colour condition, however, showed there to be little evidence of this: $r^2 = .03$, $t(8) = 0.5$, $p = ns$, and $r^2 = .11$, $t(8) = 1.0$, $p = ns$, for the white and colour conditions, respectively. The results concur with previous work in the literature, which have suggested that a speed–accuracy trade-off was not responsible for the LP effect (e.g., Kunar et al., 2010; Rich et al., 2008; Wolfe et al., 2007).

The results of Experiment 1 demonstrate a number of factors. First, in line with previous studies, participants missed more targets in the LP condition than in the HP condition. Second, giving participants a chance to self-correct reduced the LP effect. For example, in the LP colour condition, the self-correct option approximately halved the number of miss errors observed. Nevertheless, even after self-correction there was still an LP effect as miss errors at LP were higher than those at HP (see also Kunar et al., 2010; Van Wert et al., 2009). Third, knowledge of a target's colour helped target detection, overall.

Examining the effect of colour on both d' and c, we see that, when participants were given the option to self-correct, colour helped target detection as it increased both sensitivity (d') and criterion. People were more sensitive to targets in the colour condition and more willing to respond to a target's presence if they knew its colour in advance. However, looking at the interaction between prevalence and colour type on miss errors, having a target appear in a known colour did not eliminate, nor statistically reduce, the LP effect. For example, examining the data, participants missed 29% more targets at LP than at HP in the white condition and in the colour condition before correction. After the opportunity to self-correct, participants still missed 26% more white targets and 16% more colour targets at low prevalence. When the target was rare, participants missed a large number of targets, regardless of whether they knew the colour or not, in comparison to when the target appeared more frequently.

Although our results showed that colour may not eliminate the LP effect (as miss errors at LP are higher than those at HP) they do show that there are fewer miss errors when the target is colour defined. Feeding these results back into the applied setting, where it is crucial to miss as few of the rare targets as possible, adding colour to the display, although not eliminating the LP effect, appears to be worthwhile. Within these applied settings, however, a target may not necessarily appear in the predicted colour. What happens to the LP effect when the target appears in an unexpected colour? This was investigated in Experiment 2.

EXPERIMENT 2

Method

Participants

Eight participants ($M = 21.8$ years, $SD = 1.1$, 5 female) were recruited from the University of Warwick. Each received £6/hour remuneration. Informed consent was obtained, and all had normal or corrected-to-normal vision.

Stimuli and procedure

The stimuli and procedure were similar to those for the colour stimuli in Experiment 1. There were no white trials in Experiment 2. However, here, the target (when present) appeared in one colour for 75% of the trials (the majority colour) and in a single different colour for the other 25% of trials (the minority colour). Participants were told that the target was likely to appear in the majority colour set but that it could also appear in the minority colour and were informed in advance what those colour sets would be. The assigned target colours that each participant was given were randomized and remained constant through each of the participant's experimental blocks.

There were five experimental blocks containing two conditions: high prevalence and low prevalence. The five blocks were presented over four sessions, with at least 30 minutes between sessions. The high-prevalence (50% target present) condition consisted of 160 trials (80 target present, 80 target absent). The low-prevalence (2% target present) condition consisted of 4,000 trials (80 target present, 3,920 target absent) split into four sessions of 1,000 trials each. The order of the four low-prevalence sessions and the position of the high-prevalence condition, within one of those sessions, were randomized. Participants were given 20 practice trials before each experimental block.

Results and discussion

RTs less than 200 ms and greater than 5,000 ms were removed as outliers. This led to the removal of less than 1% of the data at high prevalence and 1% of the data at low prevalence. Independent t tests showed that there was no difference depending on which colour target participants searched for (all ts < 1.3, ps $> .3$). As such, all data were pooled for analysis. Figure 3 shows the miss errors.

Errors

Initial errors. A repeated measures ANOVA on the initial miss errors, with factors of prevalence (high or low) and colour proportion (majority or minority) showed that there was a main effect of

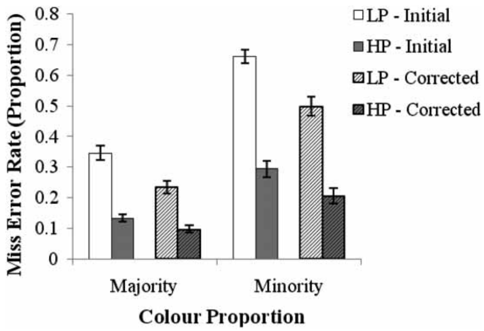

Figure 3. *Proportion of initial and corrected miss errors for the majority and minority colour conditions at high prevalence (HP) and low prevalence (LP) in Experiment 2. Error bars represent the standard error.*

prevalence, $F(1, 7) = 75.01$, $p < .01$, where more targets were missed at LP than at HP. There was also a main effect of proportion, $F(1, 7) = 37.21$, $p < .01$, where there were more miss errors when the target appeared in the minority colour than when it appeared in the majority colour. The Prevalence × Proportion interaction also approached significance, $F(1, 7) = 5.56$, $p = .051$. The LP effect was more pronounced when the target was in the minority colour than when it was in the majority colour.

Corrected errors. There was a main effect of prevalence, $F(1, 7) = 57.52$, $p < .01$, where more targets were missed at LP than at HP, and a main effect of proportion, $F(1, 7) = 15.32$, $p < .01$, where more targets were missed in the minority colour than in the majority colour. The Prevalence × Proportion interaction was also significant, $F(1, 7) = 12.29$, $p < .05$. The LP effect was more pronounced when the target was in the minority colour than when it was in the majority colour.

The self-correction option improved participant's performance (Fleck & Mitroff, 2007). In all conditions, the overall miss errors were significantly reduced after correction (all ts > 5.1, ps $< .01$). For example, at LP, miss errors were reduced by 11% in the majority colour targets and 16% for the minority colour targets. However, even after self-correction there was still an LP effect, as miss errors were higher at LP than at HP (37% vs. 15%, respectively).

Sensitivity (d') and response criterion (c)

Table 3 shows the d' and c values for each condition.

Initial d'. A repeated measures ANOVA on the initial d', with factors of prevalence (high or low) and colour proportion (majority or minority), showed that there was no main effect of prevalence. However there was a main effect of proportion, $F(1, 7) = 34.01$, $p = .001$. There was a lower sensitivity for targets appearing in the minority colour than for those in the majority colour. There was no Prevalence × Proportion interaction.

Corrected d'. Analysis of the corrected d' data showed that there was a main effect of prevalence, $F(1, 7) = 5.91$, $p = .045$, where sensitivity decreased at LP, and a main effect of proportion, $F(1, 7) = 16.92$, $p = .004$, where sensitivity was

Table 3. Means and standard deviations for initial and corrected sensitivity and criterion, for high-prevalence and low-prevalence trials in Experiment 2

	Initial				Corrected			
Experimental condition	d′	SD	c	SD	d′	SD	c	SD
High prevalence/majority colour	3.63	0.80	0.53	0.29	3.85	0.75	0.43	0.31
Low prevalence/majority colour	3.56	0.97	1.35	0.32	3.47	0.83	0.93	0.40
High prevalence/minority colour	2.95	0.81	0.86	0.24	3.32	0.82	0.70	0.30
Low prevalence/minority colour	2.59	0.95	1.83	0.45	2.61	1.08	1.37	0.44

Note: d' = sensitivity. c = criterion.

lower for the targets in the minority colour. The Prevalence × Proportion interaction was also significant, $F(1, 7) = 6.93$, $p = .034$. Sensitivity decreased more for the minority colour targets at LP than for the majority colour targets.

Initial criterion. A repeated measures ANOVA on the initial criterion, with factors of prevalence (high or low) and colour proportion (majority or minority), showed there to be a main effect of prevalence, $F(1, 7) = 42.73$, $p < .001$, where criterion was more conservative at LP. There was also a main effect of proportion, $F(1, 7) = 34.01$, $p < .001$, where criterion was more conservative for the minority targets than for the majority targets. The Prevalence × Proportion interaction was not significant.

Corrected criterion. Analysis of the corrected criterion showed there to be a main effect of prevalence, $F(1, 7) = 39.91$, $p < .001$, where criterion was more conservative at LP, and a main effect of proportion, $F(1, 7) = 16.92$, $p = .004$, where criterion was more conservative for the minority targets than for the majority targets. There was also a significant Prevalence × Proportion interaction, $F(1, 7) = 6.93$, $p = .034$, where the increase in criterion from HP to LP was greater for the minority targets than for the majority targets.

RTs

The RT results followed the same pattern as that observed in the literature (see Table 2). Correct target-absent trials (correct rejections) were significantly faster at LP than at HP, $t(7) = 3.48$, $p < .05$. As there could be no minority or majority colour distinction on target-absent trials, there was no comparison to make between proportions. There was a main effect of prevalence for correct target-present responses (hits), $F(1, 7) = 29.33$, $p < .01$, where RTs were slower at LP. However, there was no main effect of proportion or a Prevalence × Proportion interaction.

To investigate whether the elevated miss errors at LP were due to a speed–accuracy trade-off, we correlated the LP miss errors for each condition with the difference between HP and LP RTs for correct target-absent trials, as in Experiment 1. Two separate correlations for majority colour targets and minority colour targets suggested that there was little evidence for a speed–accuracy trade-off: $r^2 = .08$, $t(8) = 0.8$, $p = ns$, and $r^2 = .09$, $t(8) = 0.9$, $p = ns$, for the majority and minority conditions, respectively. This concurs with the results of previous work (e.g., Kunar et al., 2010; Wolfe et al., 2007).

An LP effect occurred in all conditions. Having a self-correct response reduced miss errors significantly for both the majority and minority colour targets. However, an LP effect was still observed, even after self-correction. Participants missed 21% more majority colour targets and 37% more minority colour targets at LP than at HP before correction. After the opportunity to self-correct, participants missed 14% more majority colour targets and 29% more minority colour targets at low prevalence. More crucially, having the target appear in a minority (or, unexpected) colour affected the magnitude of miss errors across conditions: The LP deficit was markedly worse for the minority colour targets than for the majority colour targets. Before correction, participants missed almost 70% of targets that appeared in the unexpected colour, numerically a much higher percentage than that usually observed in this task (e.g., 30–40%, Wolfe et al., 2005; see also Fleck & Mitroff, 2007; Kunar et al., 2010; Rich et al., 2008; Van Wert et al., 2009). After correction, participants missed approximately 50% of the targets—again a greater percentage than that observed elsewhere (e.g., Fleck & Mitroff, 2007; Kunar et al., 2010; Van Wert et al., 2009). Participants also showed a lower sensitivity (d') to minority colour targets and were less willing to respond to them (showing a criterion shift) than they were to majority colour targets. These effects were magnified even further at LP (especially after self-correction) and can explain why participants made more miss errors under LP minority target colour conditions. From Experiment 1, we know that the use of colour in LP search can improve target detection, overall, if the target appears in a known colour set, but Experiment 2 shows that it can also dramatically impair performance if the target is an unexpected colour.

Godwin et al. (2010) recently showed that when searching for two separate targets (e.g., a gun and an improvised explosive device, IED) with different prevalence rates, participants showed a deficit in performance in finding the less prevalent target. Our results are similar. However, in Godwin et al.'s study, participants were searching for two entirely separate complex items—with different shapes, sizes, and components. If observers had the goal state of searching for a gun, for example, then items falling outside of that target template's shape (e.g., an IED) may be easily missed. In our experiment, however, people could search for the target using two templates: shape (which was always valid) and colour (which was mostly valid). Even with an *always* valid shape, the imbalance in miss errors suggests that participants put more emphasis on the use of colour when looking for the target. This leads to a greater number of missed targets appearing in the minority colour, which may otherwise have been detected without the use of colour cues.

Experiment 2 investigated the use of valid and invalid colour cues in LP search. Experiment 3 investigates cueing further by using an exogenous red-box cue that could be either valid or invalid at predicting the target's location.

EXPERIMENT 3

Method

Participants

Nine participants ($M = 26.22$ years, $SD = 3.31$, 4 female) were recruited from the University of Warwick. Each received £6/hour remuneration. Informed consent was obtained, and all had normal or corrected-to-normal vision.

Stimuli and procedure

The stimuli were the same as those in the white stimuli trials from Experiment 1. On some critical trials, the outline of a red box with a visual angle of 9° × 8.5° (at a viewing distance of 57 cm) would be presented (see Figure 4).There were three conditions where the red box would appear: (a) TBox trials—with the target located within the bounds of the box; (b) DBox trials—with the target present but outside the box; or (c) box–no target trials—where the target was absent, and the box appeared in a random position on the screen (but always containing at least one distractor item—these trials were used to mimic false-alarm trials in the applied setting). In the case of the target appearing outside the red box (DBox trials), the box would, instead, be focused around a distractor. It was possible for the box to contain, or overlap, more than the target or a single distractor as the remaining distractors also appeared randomly, but it was ensured that the target could not be presented inside, or be obscured by, the box on DBox trials.

There were two experimental conditions: high prevalence and low prevalence. The HP condition was used as a baseline to analyse the specific effects of using a red box to cue the possible target locations. In this condition, there were 240 trials with 120 (50%) being target-present trials. Of these target-present trials, 40 would include the red box containing the target (TBox), 40 would have the red box with the target somewhere outside the box (DBox), and a further 40 would

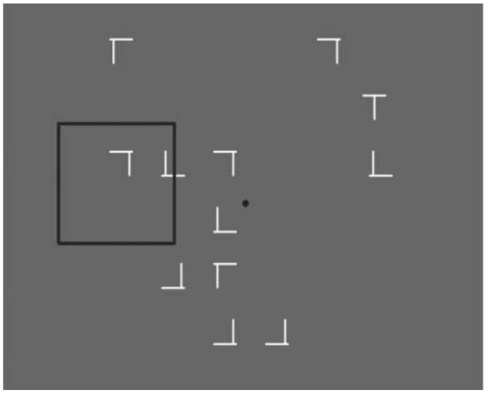

Figure 4. *Example of a target-present trial (DBox trial) where the target is located outside the red box (shown in dark grey) in Experiment 3.*

have no box at all (NB). The remaining 120 trials (target-absent) would include 40 with the red box randomly positioned in the display (box–no target). The HP condition demonstrated how effective the red-box cue was at capturing attention when the target was frequent (see also Jonides, 1981; Remington et al., 1992).

The LP condition had a total of 5,000 trials with 100 (2%) target-present trials—60 with the target inside the box (TBox), 20 with the target outside the box (DBox), and 20 with no box at all (NB). Also, 1,170 of the target-absent trials had a red box (box–no target) to bring the total trials with a box present to 1,250 (25% of total trials). Thus, for LP trials with these proportions, the target, if present, is likely to appear within the cued area (with the idea that screener assistant technologies are likely to catch prohibited items on the majority of bags they appear in). We also included "false-alarm" trials where the red box, but no target, would be present. These mimic baggage screening displays where the red box highlights a nonthreatening item (e.g., confusing a walkman for an IED). As the prevalence of the target appearing in the LP and HP trials needed to be kept at 2% and 50%, respectively, the absolute numbers of the different trial types could not be equated. There are a number of issues with this methodology. First, we see that the overall percentage of trials that contains a red box also differs between conditions. If people are simply using the presence of a red box to signify that a target may appear somewhere in the display, then this signal varies between HP and LP. Second, given the nature of this design, the probability of the target appearing within the cued region also differs between conditions. To address these concerns, we ran two baseline conditions that equated the proportion of HP trial types (e.g., TBox, DBox, NB trials) to that of the HP and LP conditions presented in Experiment 3, respectively. The results from both of these HP conditions showed a similar pattern of results to that presented here.[1] The data suggest that, despite the proportional differences in trial types, the HP condition is useful in giving a measure of how effective these cues are when the target appears frequently. We discuss these differences further in the General Discussion.

The total number of LP trials, and each trial type variation, was split equally into five sessions of 1,000 trials.[2] Participants completed the HP condition during one of these sessions, the order of which—including the position of the HP condition within one of these sessions—was randomized. There was at least 30 min between each session, and they could be spread over a number of days. Participants were given a 20-trial practice session before each experimental block. They were asked to respond as quickly but as accurately as possible and were also told that the red box was there to assist them and might highlight the most likely position of the target, if it was present.

Results and discussion

RTs less than 200 ms and greater than 5,000 ms were removed as outliers. This led to the removal of less than 1% of the data at both high and low prevalence. Figure 5 shows the miss errors.

Errors

Initial errors. A repeated measures ANOVA on miss errors with factors of prevalence (high and low) and target location (NB, TBox, or DBox) showed that there was a main effect of prevalence, $F(1, 8) = 159.66$, $p < .01$, with more miss errors

[1] Two baseline HP conditions were run: one that replicated the proportions of each trial in the HP condition in Experiment 3 ("HP-replication") and one that equated the proportions of each trial to that of the LP condition in Experiment 3 ("HP-equated"). The results showed that there was no difference in initial errors between each trial type across HP conditions (all ps $> .2$). Although RTs were faster, and sensitivity was lower for TBox trials in the HP-equated condition than in the HP-replication condition, there were no other differences between the conditions in terms of RTs or d'. More importantly, between-participant comparisons of the LP condition of Experiment 3 and the HP-equated condition showed that the same pattern of miss errors was observed as that reported in Experiment 3. Using different proportions of each trial type did little to change the pattern of results.

[2] One of the five sessions contained 1,001 trials (an extra target-absent trial), bringing the total to 5,001 trials. This extra trial, however, did not affect the overall pattern of results. The prevalence rate of the target was still 2.0%.

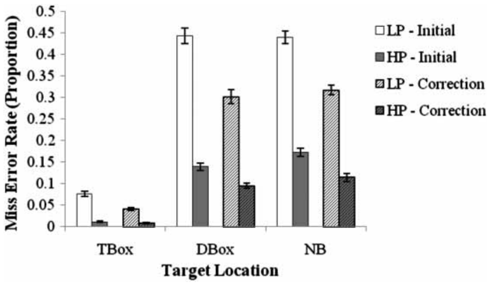

Figure 5. *Proportion of initial and corrected miss errors as a function of target location (TBox, DBox, and NB) at high prevalence (HP) and low prevalence (LP) in Experiment 3. TBox trials: target located within the box. DBox trials: target present but outside the box. NB = no box. Error bars represent the standard error.*

for LP than for HP, and a main effect of target location, $F(2,\ 16) = 79.01$, $p < .01$, where there were fewer errors in the TBox trials than in the DBox or NB trials. There was also a Prevalence × Target Location interaction, $F(2,\ 16) = 3.79$, $p < .05$, where the LP effect was bigger for the DBox and NB trials than for the TBox trials.

Corrected errors. There was a main effect of prevalence, $F(1,\ 8) = 74.39$, $p < .01$, where there were more errors in the LP conditions than in the HP conditions. There was also a main effect of target location, $F(2,\ 16) = 37.09$, $p < .01$, where there were fewer errors in the TBox trials than in the DBox or NB trials. The Prevalence × Target Location interaction was significant, $F(2,\ 16) = 4.63$, $p < .05$, where the LP effect was bigger for the DBox and NB trials than for the TBox trials.

As in Experiments 1 and 2, giving participants the option to correct their responses tended to reduce miss errors. Miss errors were reduced for the NB, $t(8) = 3.11$, $p < .05$, and DBox conditions, $t(8) = 3.92$, $p < .01$, at HP. There was no difference in miss errors in the TBox conditions at HP, presumably as performance was already at ceiling in this condition initially. Miss errors were significantly reduced by 4%, 14%, and 12% in the TBox, DBox, and NB trials, respectively, at LP (all ts > 2.91, ps < .03). However, even after self-correction, an LP effect was observed, as miss errors were higher at LP than at HP (22% vs. 7%, respectively).

Sensitivity (d') and response criterion (c)

Table 4 shows the d' and c values for each condition.

Initial d'. A repeated measures ANOVA on the initial d', with factors of prevalence (high or low) and target location (NB, TBox, or DBox), showed that there was a main effect of prevalence, $F(1,\ 8) = 11.93$, $p = .009$, where overall d' increased at LP, and a main effect of target location, $F(2,\ 16) = 63.38$, $p < .001$, where d' was greater for TBox trials than for DBox and NB trials. There was no Prevalence × Target Location interaction.

Table 4. Means and standard deviations for initial and corrected sensitivity and criterion, for high-prevalence and low-prevalence trials in Experiment 3

	Initial				Corrected			
Experimental condition	d′	SD	c	SD	d′	SD	c	SD
High prevalence/TBox	4.26	0.31	0.01	0.09	4.29	0.29	−0.004	0.11
Low prevalence/TBox	4.59	0.49	0.77	0.21	4.48	0.49	0.40	0.26
High prevalence/DBox	3.32	0.51	0.49	0.22	3.53	0.43	0.38	0.20
Low prevalence/DBox	3.23	0.60	1.45	0.21	3.23	0.53	1.02	0.36
High prevalence/NB	3.27	0.50	0.64	0.18	3.75	0.50	0.54	0.29
Low prevalence/NB	3.60	0.31	1.64	0.25	3.46	0.36	1.24	0.33

Note: d' = sensitivity. c = criterion. TBox trials: target located within the box. DBox trials: target present but outside the box. NB = no box.

Corrected d′. Analysis of the corrected d' data showed that there was no main effect of prevalence but there was a main effect of target location, $F(2, 16) = 30.79$, $p < .001$. The value of d' was greater for TBox trials than for DBox and NB trials. The Prevalence × Target Location interaction was significant, $F(2, 16) = 6.07$, $p = .011$, where d' was greater at LP than at HP for the TBox trials but not for the DBox or NB trials. Replicating Wolfe et al. (2007), d' did not change across prevalence but did vary between target location. Participants are more able to distinguish the target from distractors on the TBox trials than in the other target-present trials.

Initial criterion. A repeated measures ANOVA on the initial criterion, with factors of prevalence (high or low) and target location (NB, TBox, or DBox), showed that there was a main effect of prevalence, $F(1, 8) = 11.20$, $p < .001$, where people were more conservative in their response at LP, and that there was a main effect of target location, $F(2, 16) = 67.25$, $p < .001$, where responses to TBox trials were more liberal than responses to DBox and NB trials. The Prevalence × Target Location interaction was significant, $F(2, 16) = 5.28$, $p = .017$. People were less willing to respond target present at LP than at HP, although this difference was greater in the DBox trials and NB trials than in the TBox trials.

Corrected criterion. Analysis of the corrected criterion data showed there to be a main effect of prevalence, $F(1, 8) = 47.96$, $p < .001$, where people were more conservative at LP, and a main effect of target location, $F(2, 16) = 32.55$, $p < .001$, where responses to TBox trials were more liberal than responses to DBox and NB trials. The Prevalence × Target Location interaction was significant, $F(2, 16) = 6.49$, $p = .009$. All target location conditions showed that people were less willing to respond target present at LP than at HP, although this difference was greater in the DBox trials and NB trials than in the TBox trials.

RTs

Correct target-absent (correct rejection) responses were faster at LP than at HP, $F(1, 8) = 11.29$, $p < .05$ (see Table 2). This result is again consistent with the miss RT findings reported by Wolfe et al. (2007). Also, there was a main effect of target location, $F(1, 8) = 26.20$, $p < .05$, where NB responses were faster than the box–no target trials. This could be due to participants expecting a target to appear, or at least believing that it may be more likely, and allocating more search time to these trials. The "hit" RT data provided a marginal main effect of prevalence, $F(1, 8) = 4.74$, $p = .061$, where there was a trend for correct target-present responses to be slower at LP, as reported by Wolfe et al. (2007). Also, there was a main effect of target location, $F(2, 16) = 39.55$, $p < .01$, showing that RTs were fastest for TBox targets and slowest for DBox targets. There was no significant Prevalence × Target Location interaction.

To investigate whether the elevated miss errors at LP were due to speed–accuracy trade-offs, as in Experiments 1 and 2, we correlated the LP miss errors for each condition with the difference between HP and LP RTs for correct target-absent trials. Three separate correlations for TBox trials, DBox trials, and NB trials suggested that there was little evidence for a speed–accuracy trade-off in any of these conditions: $r^2 = .00$, $t(7) = 0.1$, $p = ns$; $r^2 = .01$, $t(7) = 0.3$, $p = ns$; and $r^2 = .01$, $t(7) = 0.2$, $p = ns$, for the TBox, Dbox, and NB trials, respectively. These results concur with those of previous studies (e.g., Kunar et al., 2010; Rich et al., 2008; Wolfe et al., 2007).

Overall, the data showed that there was an LP effect in every condition. Self-correction reduced the percentage of targets missed in each condition. Participants missed 6% more TBox targets at LP than at HP, while they missed 30% and 27% more DBox and NB targets, respectively, before correction. After the opportunity to self-correct, participants missed 3% more TBox targets at LP than at HP, while they missed 21% and 20% more DBox and NB targets, respectively. However, similar to Experiments 1 and 2, the option to self-correct did not eliminate the LP effect in any of the conditions, as there were still more miss errors at LP than at HP. The exogenous cue, on the other hand, did have an effect on miss errors across conditions. The LP effect was quantitatively reduced when the target appeared in the cued area (TBox trials) than when it appeared outside the cued area or when there was no box.

Similar to Experiment 2, miss errors for targets in the invalid conditions (where the target appeared outside the highlighted area, or if no red box was presented at all) were higher than those in the valid conditions. The data also showed that people were more sensitive to targets that were highlighted by the box (as d' was higher) and that people were less willing to respond to a target's presence if it fell outside the box or when no box was presented. This shift in criterion was greater for LP trials than for HP trials. Participants were even less willing to respond to a target's presence when it was rare *and* when the highlighting box either was absent or did not predict the target's location. We discuss this further in the General Discussion.

Of final note, having the target appear in the cued area greatly minimized miss errors in HP conditions. Miss errors were close to 0%, even before correction, presumably as the red box acted as a very strong exogenous cue, pointing to the target position. Despite acting as such a salient cue at HP, these cues did not lead to the same observable behaviour at LP. Although the presence of the cue reduced the number of miss errors, participants still missed more highlighted and saliently cued targets when they had a LP. As the cue was physically identical in both conditions, the results suggest that the effectiveness of the exogenous cue varied depending on the target's prevalence rate. An alternative suggestion may be that in both HP and LP conditions, the presence of a red box indicated the likelihood of a target being present. However, given the necessary difference in the number of trials across conditions, the predictive nature of the red box was greater at HP than at LP.[3] If so, this may explain why participants still missed some targets at LP but virtually no targets at HP. However even with a red cue appearing somewhere in the display (indicating the likelihood of a target), miss errors in DBox trials were similar to those when no predicting red box appeared (i.e., in NB trials, both at HP and at LP). Thus, in both conditions, the predictive aspect of the red box did not help search, unless it directly cued the target location.

GENERAL DISCUSSION

Previous work has shown that, under low-prevalence conditions, participants miss a target more often than if it appears frequently (Van Wert et al., 2009; Wolfe et al., 2005; Wolfe et al., 2007). The experiments presented here examined how visual cues, such as colour (Experiments 1 and 2) and exogenous cues, highlighting the likely area of the target (Experiment 3), affected search under LP conditions.

Experiment 1 investigated whether knowing the colour of the target in advance improved detection of rare targets. Colour has been shown to improve target detection under high-prevalence conditions (e.g., Egeth et al., 1984; Michel et al., 2007; Treisman & Gelade, 1980) and has been used as a strong guidance signal in search (e.g., Hannus et al., 2006; Luria & Strauss, 1975; Motter & Belky, 1998; Williams & Reingold, 2001; Wolfe, 2007; Wolfe et al., 1989). In a similar manner, we found that adding colour to the display improved detection of the target, overall, and resulted in fewer miss errors (especially after participants were given the opportunity to self-correct). However, the overall LP effect when participants were given advanced knowledge of the target's colour was no different than when no colour information was given.

Experiments 2 and 3 assessed the use of valid and invalid cues in LP search. Experiment 2 investigated miss errors of targets that appeared in an expected or unexpected colour, while Experiment 3 investigated miss errors of targets that were validly or invalidly cued by an exogenous red box. The results showed that both valid colour cues

[3] We thank a reviewer for this suggestion.

and exogenous red box cues reduced the number of miss errors observed at both high prevalence and low prevalence. Furthermore, having a valid cue reduced the LP effect compared to an invalid cue (Experiments 2 and 3) and when there was no cue at all (Experiment 3). Within both these experiments, miss errors increased when the cue was invalid. Having the target appear in an unlikely colour or misdiagnosed location posed a severe cost to LP search. This is important if we extrapolate our results back to the applied setting, where baggage screeners make use of colour cues and red box cues when searching for prohibited items in X-rayed images. Although these cues are useful when the target appears in the expected feature/location, having them appear in an unexpected colour or location seems to be more costly then not having the cue at all.

Giving participants the option to correct their responses reduced the LP effect. In all of our experiments, fewer miss error rates were observed after correction than before. This agrees with the findings of Fleck and Mitroff (2007) who suggest that some of the miss errors in LP search are due to motor errors. However, it should be noted that, although the option to correct reduced the LP effect, it did not eliminate it, as miss errors were higher at LP than at HP even after self-correction (see also Kunar et al., 2010; Van Wert et al., 2009). It seems that along with prepotent motor biases, other accounts are needed to explain the high miss errors observed in LP search, such as a shift in criterion and/or a reduced quitting threshold in search (Rich et al., 2008; Wolfe et al., 2007; Wolfe & Van Wert, 2010).

Let us now examine our results in terms of signal detection. Previous work suggests that the LP effect is due to a shift in criterion. Wolfe et al. (2007) identified this criterion shift as a conservative shift (a higher value of c, above zero) or an unwillingness to respond target present at LP. The same criterion shift was observed in all our experiments here. Furthermore, in Experiment 1, colour stimuli were found to be accompanied by a more liberal criterion than white stimuli at low prevalence. Participants would more readily accept an LP trial as target present when the stimuli were coloured than when they were white. Similar results were found for expected (majority colour) over unexpected (minority colour) target colour, where participants were more willing to respond target present when the target appeared in the expected colour (Experiment 2) and when the target appeared in a red box than when it appeared outside the red box or when no red box was present (Experiment 3). This change in response criterion may explain why these visual cues (if valid) were effective in reducing the number of miss errors observed.

While criterion shifted, Wolfe et al. (2007) reported that d' did not change across prevalence. For the most part, our data are consistent with this (although we did observe a decrease in sensitivity at LP for Experiment 2). However, the different manipulations had differential effects on sensitivity at LP and at HP. Unsurprisingly, knowing the target colour (or expected target colour) improved sensitivity. Likewise, having the target appear in a cued area also improved sensitivity. However, for predictive majority colour cues, the observed sensitivity was greater at HP than at LP (Experiment 2). Conversely, sensitivity was greater when the target appeared in a highlighted cued area at LP than at HP (Experiment 3). At first glance it may appear strange that sensitivity increased with a reduction in target frequency in this latter condition. However, this may be explained if we examine the percentage of trials across conditions that contained a red box cue. In the LP condition, 25% of trials contained a cue, while in the HP condition, 50% of trials contained a cue. Presenting the cue less frequently in the LP condition may have resulted in it having a greater attentional weight when it actually appeared, leading to an increase in d'. Please note, however, that even with this increased sensitivity at LP, the red box cue did not entirely eliminate the LP effect. Further research is needed to investigate this. However, more importantly, for present purposes, the data here show that, under LP conditions, screener assistant technologies that correctly highlight a potentially dangerous area improve the ability to distinguish target-present from target-absent trials.

In line with previous research, the results of these experiments point to the robustness of the LP effect and that finding ways to eliminate LP search is not trivial (e.g., Wolfe et al., 2007). Although manipulating a person's reward scheme has been shown to reduce the LP effect (Navalpakkam et al., 2009), other more physical manipulations have not been so successful. The results reported here show that adding colour cues and even highly salient exogenous cues, which effectively eliminate miss errors at HP (Experiment 3), fail to eliminate miss errors entirely at LP. Other work has also shown that search manipulations that are known to benefit search performance at high prevalence may not be as effective at low prevalence. Rich et al. (2008) showed that participants missed more LP targets than HP targets even in a single feature search, where the target was shown to "pop out" of the display. Furthermore, Kunar et al. (2010) found that separating the presentation of the display into two discrete spatial and/or temporal sections did not remove the LP effect. Given that LP search occurs in many important applied settings, further work is needed to understand and find ways to eradicate the search cost of having the target appear rarely.

Original manuscript received 10 February 2011
Accepted revision received 09 December 2011
First published online 13 April 2012

REFERENCES

Bretz, E. A. (2002). Delayed arrival for US baggage screening. *IEEE Spectrum*, *39*(5), 16–19.

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, *3*(3), 292–297.

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts effects of set size on visual search accuracy for feature, conjunction, triple conjunction and disjunction displays. *Perception & Psychophysics*, *62,* 425–451.

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(1), 32–39.

Evans, J. P. O., Wang, X., Chan, J. W., Downes, D., & Liu, Y. (2006). Color 3D x-ray imaging for security screening, In *IET—The Crime and Security Conference: Imaging for Crime Detection and Prevention* (pp. 372–377). London, June 2006.

Fenton, J. J., Taplin, S. H., Carney, P. A., Linn, A., Sickles, E. A., D'Orsi, C., et al. (2007). Influence of computer-aided detection on performance of screening mammography. *The New England Journal of Medicine*, *356*(14), 1399–1409.

Fleck, M. S., & Mitroff, S. R. (2007). Rare targets are rarely missed in correctable search. *Psychological Science*, *18*(11), 943–947.

Godwin, H. J., Menneer, T., Cave, K. R., Helman, S., Way, R. L., & Donnelly, N. (2010). The impact of relative prevalence on dual-target search for threat items from airport X-ray screening. *Acta Psychologica*, *134*(1), 79–84.

Guardian Technologies International (2008). *Security Screening*. Retrieved July 7, 2009, from Guardian Technologies International: http://www.guardiantechintl.com/security.php?npage=pinpoint

Hannus, A., van den Berg, R., Bekkering, H., Roerdink, J. B. T. M., & Cornelissen, F. W. (2006). Visual search near threshold: Some features are more equal than others. *Journal of Vision*, *6,* 523–540.

Hogg, R. V., & Craig, A. T. (1995). *Introduction to mathematical statistics* (5th ed.). Englewood Cliffs, NJ: Prentice Hall International.

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

Jonides, J. (1981). Voluntary vs. automatic control over the mind's eye's movement. In J. B. Long & A. D. Baddeley (Eds.), *Attention and Performance IX.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Kunar, M. A., Rich, A. N., & Wolfe, J. M. (2010). Spatial and temporal separation fails to counteract the effects of low prevalence in visual search. *Visual Cognition*, *18,* 881–897.

Leung, J. W., Margolin, F. R., Dee, K. E., Jacobs, R. P., Denny, S. R., & Schrumpf, J. D. (2007). Performance parameters for screening and diagnostic mammography in a community practice: Are there differences between specialists and general radiologists? *American Journal of Roentgenology*, *188*(1), 236–341.

Luria, S. M., & Strauss, M. S. (1975). Eye movements during search for coded and uncoded targets. *Perception & Psychophysics*, *17,* 303–308.

Michel, S., Koller, S. M., Ruh, M., & Schwaninger, A. (2007). Do “image enhancement” functions really enhance x-ray image interpretation? *Proceedings of the 29th Annual Cognitive Science Society* (pp. 1301–1306). Austin, TX: Cognitive Science Society.

Motter, B. C., & Belky, E. J. (1998). The guidance of eye movements during active visual search. *Vision Research*, *38,* 1805–1815.

Navalpakkam, V., Koch, C., & Perona, P. (2009). Homo economicus in visual search. *Journal of Vision*, *9*(1), 1–16.

Parks, L. (2007). Points of departure: The culture of US airport screening. *Journal of Visual Culture*, *6*(2), 183–200.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32,* 3–25.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bouwhuis (Eds.), *Attention & Performance X* (pp. 531–556). Hillsdale, NJ: Lawrence Erlbaum Associates.

Remington, R., Johnston, J. C., & Yantis, S. (1992). Attentional capture by abrupt onsets. *Perception & Psychophysics*, *51,* 279–290.

Rich, A. N., Kunar, M. A., Van Wert, M. J., Hidalgo-Sotelo, B., Horowitz, T. S., & Wolfe, J. M. (2008). Why do we miss rare targets? Exploring the boundaries of the low prevalence effect. *Journal of Vision*, *8* (15), 1–17.

Rubenstein, J. (2001). *Test and evaluation plan: X-Ray Image Screener Selection Test*, (Report No. DOT/FAA/AR-01/47). Washington, DC: Office of Aviation Research.

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

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

Van Wert, M. J., Horowitz, T. S., & Wolfe, J. M. (2009). Even in correctable search, some types of rare targets are frequently missed. *Attention, Perception & Psychophysics*, *71*(3), 541–553.

Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, *31*(4), 523–535.

Williams, D. E., & Reingold, E. M. (2001). Preattentive guidance of eye movements during triple conjunction search tasks: The effects of feature discriminability and saccadic amplitude. *Psychonomic Bulletin & Review*, *8*(3), 476–488.

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, *9*(1), 33–39.

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

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419–433.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual search. *Nature*, *435,* 439–440.

Wolfe, J. M., Horowitz, T. S., Ven Wert, M. J., Kenner, N. M., Place, S. S., & Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental Psychology*, *136* (4), 623–638.

Wolfe, J. M., & Van Wert, M. J. (2010). Varying target prevalence reveals two, dissociable decision criteria in visual search. *Current Biology*, *20,* 121–124.

Zelinsky, G. J. (2003). Detecting changes between real-world objects using spatiochromatic filters. *Psychonomic Bulletin & Review*, *10,* 533–555.

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, *115* (4), 787–835.