More is better: Relative prevalence of multiple targets affects search accuracy

Phillip (Xin) Cheng

Perception in Action Research Centre (PARC), Department of Cognitive Science, Faculty of Human Sciences, Macquarie University, Sydney, NSW, Australia

✉

Anina N. Rich

Perception in Action Research Centre (PARC), Department of Cognitive Science, Faculty of Human Sciences, Macquarie University, Sydney, NSW, Australia
ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney, NSW, Australia
Centre for Elite Performance, Expertise and Training (CEPET), Macquarie University, Sydney, NSW, Australia

⌂ ✉

In real-world searches such as airport baggage screening and radiological examinations, miss errors can be life threatening. Misses increase for additional targets after detecting an initial target, termed "subsequent search misses" (SSMs), and also when targets are more often absent than present, termed the low-prevalence effect. Real-world search tasks often contain more than one target, but the prevalence of these multitarget occasions varies. For example, a cancerous tumor sometimes coexists with a benign tumor and sometimes exists alone. This study aims to investigate how the relative prevalence of multiple targets affects search accuracy. Naive observers searched for all Ts (zero, one, or two) among Ls. In Experiment 1, SSMs occurred in small but not large set sizes, which may be explained by classic capacity limit effects such as the attentional blink and repetition blindness. Experiment 2 showed an interaction between SSMs and the relative prevalence of dual-target trials: Low prevalence of dual-target trials increased SSMs relative to high prevalence dual-target trials. The prevalence of dual-target trials did not affect accuracy on single-target trials. These results may provide a novel avenue for reducing misses by increasing the prevalence of instances with multiple targets. Future efforts should take into account the relative prevalence of multiple targets to effectively reduce life-threatening miss errors.

Introduction

Visual search is ubiquitous and seems effortless most of the time. For example, we can easily find milk in the refrigerator using a combination of our knowledge of the organization of the refrigerator and what the milk looks like; if we accidentally select another item or miss the presence of the milk, there is typically no harm done. Although these everyday searches seem mundane, there are real-world situations where accuracy in search is vital. These searches include airport baggage screening and radiological examinations. Missing a gun in x-rayed baggage or a cancer in a radiograph can have life-and-death consequences. Unfortunately, these high-stake targets have two features that make them particularly error-prone—coexistence with other targets (Berbaum et al., 1991) and infrequent occurrence (Wolfe, Horowitz, & Kenner, 2005). Here, we explore the extent to which the prevalence of multiple targets (rare vs. frequent) affects search, and offer important clues for reducing life-threatening search errors.

In most laboratory-based searches observers look for a single target among distractors (e.g., Chun & Wolfe, 1996; Nakayama & Martini, 2011), but real-world searches often have more than one target. For example, an x-rayed bag may contain more than one threat and a radiograph more than one potential tumor. Unfortunately, misses increase after detecting a target, an effect termed "subsequent search misses" (SSMs; Adamo, Cain, & Mitroff, 2013), and originally termed "satisfaction of search" (Smith, 1967). SSMs were originally proposed by Tuddenham (1962) as the failure to continue searching after an initial detection. However, later studies did not support this theory and showed that observers kept searching after finding a target (e.g., Berbaum et al., 1991). Berbaum et al. (1991) suggested that additional targets more similar to the first target

Received November 21, 2017; published April 3, 2018

are less likely to be missed than less similar targets. However, Fleck, Samei, and Mitroff (2010) showed that SSMs still occur with identical targets, suggesting a more complicated story. Cain and Mitroff (2013) suggested that the location and identity of the found target consume working memory, which can otherwise benefit further search. In support of this theory, SSMs are reduced by removing the found target (Cain & Mitroff, 2013) or splitting a multitarget search into single-target searches (Cain, Biggs, Darling, & Mitroff, 2014). However, neither approach eliminates SSMs, indicating the contribution of other factors. Factors affecting SSMs include spatial closeness of stimuli (Adamo, Cain, & Mitroff, 2015), anxiety (Cain, Dunsmoor, LaBar, & Mitroff, 2011), and decision criteria (Biggs & Mitroff, 2015). Despite being studied for over 50 years, a clear solution is still lacking to prevent elevated misses in multitarget searches (Biggs, 2017).

The other key factor affecting miss rates in real-world searches is target prevalence. A well-documented example is the low prevalence effect (LPE; Wolfe et al., 2005): A rare target is far more likely to be missed. In real-world searches dangerous targets are often rare. For example, a tiny proportion of items in x-rayed bags or radiographs would be guns or cancers, respectively. Studies using realistic x-rayed bag stimuli found that not only naive searchers (Wolfe et al., 2007) but also airport professionals (Wolfe, Brunelli, Rubinstein, & Horowitz, 2013) show elevated misses in low prevalence conditions. Because misses in real-world searches can be catastrophic, researchers have been looking for countermeasures. Although attempts such as instructions and rewards have proven unsuccessful (Ishibashi, Kita, & Wolfe, 2012; Maddox, 2002), Wolfe and colleagues (Wolfe et al., 2013; Wolfe et al., 2007) showed that bursts of high prevalence trials can reduce subsequent low-prevalence misses. Here, we do not directly examine the LPE, but instead manipulate the relative prevalence of single- versus dual-target trials. By understanding multitarget search errors more thoroughly, we offer new avenues for designing countermeasures to ameliorate the LPE.

Previous studies have identified two types of search errors: selection errors (targets not fixated) and identification errors (targets fixated but not recognized; Peltier & Becker, 2016; Wolfe & van Wert, 2010). Both types of errors occur in single-target (Peltier & Becker, 2016) as well as multitarget searches (Cain, Adamo, & Mitroff, 2013), and vary with prevalence (Godwin, Menneer, Cave, Thaibsyah, & Donnelly, 2015). Selection errors seem to explain a larger proportion of misses than identification errors in both single- (Hout, Walenchok, Goldinger, & Wolfe, 2015; Peltier & Becker, 2016) and multitarget searches (Cain et al., 2013). In single-target searches, selection errors seem to

result mainly from a speed-accuracy tradeoff (i.e., faster responses at the expense of accuracy; Peltier & Becker, 2017a; Rich et al., 2008; Schwark, Sandry, & Dolgov, 2013), and the identification errors seem to result from criterion shifts (i.e., the likelihood of incorrectly dismissing a target after fixating it increases as prevalence decreases; Hout et al., 2015; Peltier & Becker, 2016). The cause of both types of errors seems to be more heterogeneous in multitarget searches than in single-target searches (Cain et al., 2013).

Most multitarget studies have employed either only target-present trials or equal prevalence of single- versus dual-target trials (see Fleck et al., 2010 for an exception). In real-world searches, however, targets are often absent and the prevalence of single- versus dual-targets is often unequal. For instance, in radiographs, a rare cancerous tumor often coexists with a more prevalent tumor (Smith & Turnbull, 1997). Fleck et al. (2010) used a T-among-L search with target-absent (TA; 50% of total trials), single-target, and dual-target trials, and showed elevated misses in dual-target trials when they were rare (10% of total trials) but not when dual- and single-target trials were equiprobable, suggesting that these misses result from low prevalence of dual-target trials rather than SSMs. Here, we extended this study by comparing performance on *rare* dual-target trials (10% of total trials) with that of *frequent* dual-target trials (40% of total trials). Note, we are manipulating the relative prevalence of single- and dual-target trials, rather than overall target prevalence per se. We have a greater difference in prevalence (30% for both single- and dual-target trials [40%–10%] compared to Fleck et al.'s 15% [40%–25%]), which, combined with two set sizes, allows us to more thoroughly explore the source of misses in single- versus multitarget searches.

Here, we use simple search stimuli (Ts and offset Ls; Figure 1a) to investigate the effect of the relative prevalence of multiple targets on search accuracy. In Experiment 1, we tested whether SSMs occur in our perceptually simple search to form a baseline. Observers searched for Ts among Ls (50% target-absent, 50% target-present). The probability of single- or dual-target trials was equal within the TP trials (Figure 1b). We found SSMs in the small set size (10 letters in a display) but not the large set size (20), which aligns with Fleck et al.'s (2010) finding that SSMs did not occur in a set size of 25. Thus, in simple searches the occurrence of SSMs seems to depend on the number of distractors. In Experiment 2, we tested whether and how the prevalence of multiple targets affects accuracy. Single-target trials were 4 times more common than dual-target trials (rare dual-target condition; Figure 1c) or the reverse (rare single-target condition; Figure 1d). Prevalence exerted a stronger influence on the additional target in dual-target trials than on the target in
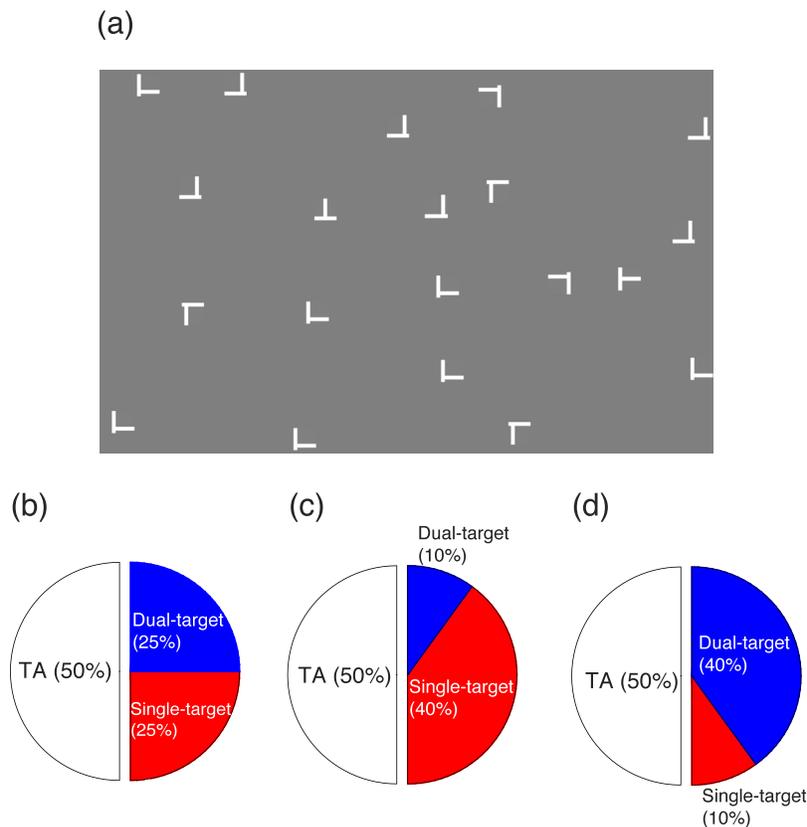
(a)

(b)　　　　　　　　　(c)　　　　　　　　　(d)

Figure 1. Example search display and prevalence conditions. (a) An example dual-target target-present trial in the large set size. There could be zero, one, or two Ts in a small (10 items) or large (20 items) set size; (b) prevalence conditions for Experiment 1; (c) rare dual-target prevalence condition for Experiment 2; (d) rare single-target prevalence condition for Experiment 2.

single-target trials, showing an interaction between SSMs and prevalence. Observers were just as accurate in finding one and two targets in the rare single-target condition as they were in finding one target in the rare dual-target condition, whereas their errors for two targets in the latter condition (SSMs) increased. This suggests that a high prevalence of dual-target trials prevented additional errors in multitarget searches. Real-world search situations may benefit from the present finding to counteract life-threatening errors.

Experiment 1: SSMs in perceptually simple search

The aim of Experiment 1 was to investigate whether SSMs occur in perceptually simple searches. The task of finding Ts among offset Ls requires attention to each item, so reaction time (RT) increases with each additional distractor, rendering it inefficient (Wolfe, Cave, & Franzel, 1989). The search is simple because one can easily distinguish a T from an offset L once attending it (see Figure 1a). Fleck et al. (2010) used similar stimuli and did not find SSMs. However, they

only employed one set size of 25. We hypothesized that the number of items in the display might be an important factor in whether SSMs occur. We therefore included two set sizes (10 and 20) to explore how set size influences miss rates in dual-target trials.

Methods

Participants

There were 12 observers, all of whom reported normal or corrected-to-normal vision and reported no history of mental or neurological disorders. They gave informed consent and received course credits or AU$15 per hour for a ~50-min. experiment. This study was approved by the Macquarie University Human Research Ethics Committee.

Apparatus

We presented displays on a 27-in. Samsung SA950 monitor (resolution: 1,920 × 1,080 pixels, refresh rate: 120 Hz). The monitor was controlled by a Dell Optiplex 9010 Intel, Pentium Quad-Core desktop. We programmed our experiment in MATLAB Release

Prevalence	Response	TA (%)	Number of targets	
			Single target (%)	Dual target (%)
1T:2T = 1:1	TA	**99.37**	*13.95*	*4.18*
	Single target	*0.49*	**84.94**	***19.22***
	Dual target	*0.14*	*1.12*	**76.60**
1T:2T = 4:1	TA	**98.99**	*19.62*	*7.87*
	Single target	*0.87*	**79.79**	***27.11***
	Dual target	*0.14*	*0.59*	**65.03**
1T:2T = 1:4	TA	**99.28**	*15.86*	*5.94*
	Single target	*0.39*	**79.83**	***12.85***
	Dual target	*0.33*	*4.31*	**81.21**

Table 1. Percentages of responses to displays of different trial types. *Notes*: Numbers in bold are correct response rates, in italic are false alarm rates, and in bold-italic are miss rates. In both experiments, there were 50% target absent (TA) trials. In Experiment 1, the remaining 50% target-present (TP) trials had equiprobable single-target and dual-target trials. In Experiment 2, dual-target trials were rare on half of TP trials and single-target trials were rare on the other half of TP trials. Note that the accuracy for TA trials in all prevalence conditions is near ceiling. Underlined values are those analyzed because we focused on subsequent search misses (SSM; see Results in Experiment 1).

2016a, using the Psychtoolbox-3 extensions (Kleiner, Brainard, & Pelli, 2007).

Stimuli

The stimuli (Figure 1a) were white (RGB: [255 255 255]) rotated Ts and offset Ls of 1.2° × 1.2° visual angle at an approximate viewing distance of 90 cm, presented on a gray (RGB: [128 128 128]) background, which filled the screen. The Ts and Ls had four possible orientations: 0°, 90°, 180°, and 270°. The stroke width was 0.2°. The background was separated into an imaginary 5 × 4 grid for large set size trials and 4 × 3 for small set size trials (see Procedure and design). Each stimulus appeared randomly within the cell.

Procedure and design

Each trial started with a fixation cross appearing at the center of the screen for 0.5 s. The cross was then replaced by the search display containing either 10 (the small set size) or 20 (the large set size) items. The two set sizes were randomly interleaved within the block. There could be zero, one, or two Ts randomly appearing among Ls in each display. We instructed observers that it was important to find all the Ts, and that they had 12 s to do so. This extended display time ensured accuracy was prioritized over speed. It was considered a time-out if they did not respond in 12 s. The intertrial interval was 1 s.

Half of the trials were TA and the other half TP, in which the single- and dual-target trial were equally likely (Figure 1b). Observers did not know beforehand the number of targets in each trial or the prevalence of each trial type. Each observer did 24 practice trials before completing 240 experimental trials. Observers

could take a break after every 80 trials. After each trial they received feedback of correct or wrong based on their accuracy. In both experiments, we planned to exclude observers who did not reach an overall accuracy of 70%.

Results

All observers in Experiment 1 performed well (>75% correct) and were included for analyses. We excluded trials with an RT smaller than 200 ms (Wolfe & van Wert, 2010) and greater than 12 s, leading to a removal of 0.35% of trials.

We focused analyses on comparing search accuracy for single-target trials to that for dual-target ones (i.e., whether the SSMs occurred), which does not require the TA trials. However, we present percentages of all types of errors in Table 1. Because SSMs are about the accuracy of finding the second target *after finding the first*, we excluded dual-target trials with TA responses to avoid inflating the SSMs (Biggs, 2017). We also analyzed RTs to examine whether the speed-accuracy tradeoff could explain the possible difference in accuracy.

We used a generalized linear mixed-effect model (GLMM; MATLAB and the Statistical and Machine Learning Toolbox Release 2017a, MathWorks, Natick, MA) to estimate effects on search accuracy and RTs. We used GLMM because it offers higher power than standard statistical tests (Baayen, 2008) and is appropriate for prevalence studies that inevitably end up with unbalanced data. For all measures, we specified effects as random at the observer level and reported them at the population level. We chose the effects based on Bayesian information criterion (BIC) to balance

goodness of fit and parsimony (Madigan & Raftery, 1994). BIC has been shown to perform well with small samples (Heskes, 1998; Preacher, Zhang, Kim, & Mels, 2013). In the Supplementary Material (S1), we present statistics from models as well as the odds ratio (OR) and Cohen's d as the standardized effect size for accuracy and RTs, respectively (Hedges, 2007; Zhang & Yu, 1998). Here odds ratios are standardized because accuracy is a binary predictor. We also present 95% confidence intervals (CIs) to better estimate true means (Cumming, 2014).

We used a Bayesian approach when a p value was greater than 0.05 to distinguish between insufficient sensitivity and true evidence for the null (Dienes, 2014). Bayes factors (BF) were derived from the Bayesian calculation for binomial data (Rouder, Morey, & Province, 2013) and Bayesian t test (Rouder, Speckman, Sun, Morey, & Iverson, 2009) implemented in a Matlab toolbox (Schwarzkopf, 2014). Here, a BF quantified the evidence of an alternative hypothesis (H1: $\mu \neq 0$) relative to the null hypothesis (H0: $\mu = 0$), and the prior probability was set as 1. In line with recommendations in the literature (Jeffreys, 1961), we interpreted $BF < 1/3$ as evidence for H0.

Performance was better for the small set size relative to the large set size (i.e., higher accuracy: $\beta = -0.07$, $p < 0.001$; shorter RT: $\beta = -0.01$, $p < 0.001$). We therefore separated data based on set sizes for further analyses.

Figure 2a shows that search accuracy is numerically higher for single- rather than dual-target trials in both set sizes. However, a model estimating the effect of number of targets indicated that the difference reached statistical significance for the small ($\beta = -0.75$, $p = 0.002$) but not large set size ($\beta = -0.31$, $p = 0.12$). The BF for the effect in the large set size favored the H0 but only slightly ($BF = 0.62$).

We examined RT data from both correct and incorrect trials to obtain more information about the accuracy difference between single- and dual-target trials. Our RT distributions were positively skewed, consistent with previous studies involving simple judgments (e.g., Yap & Balota, 2007). Lo and Andrews (2015) argued that raw RTs should be used to maintain the ratio scale in such circumstances. We used gamma distributions in GLMMs to realize the fitting. In addition to the statistical appropriateness, gamma theoretically fits the distribution—it is linked to waiting time until a response (a keypress in this study; Lo & Andrews, 2015).

To test whether observers spent an equal amount of time searching in single- and dual-target displays, we separated correct and incorrect trials and analyzed the effect of number of targets on RTs. Correct RTs were longer for single-target (Figure 2b, left light gray in each panel) than dual-target (left dark gray in each panel) trials in small ($\beta = 0.06$, $p < 0.001$) and large set sizes ($\beta = 0.06$, p
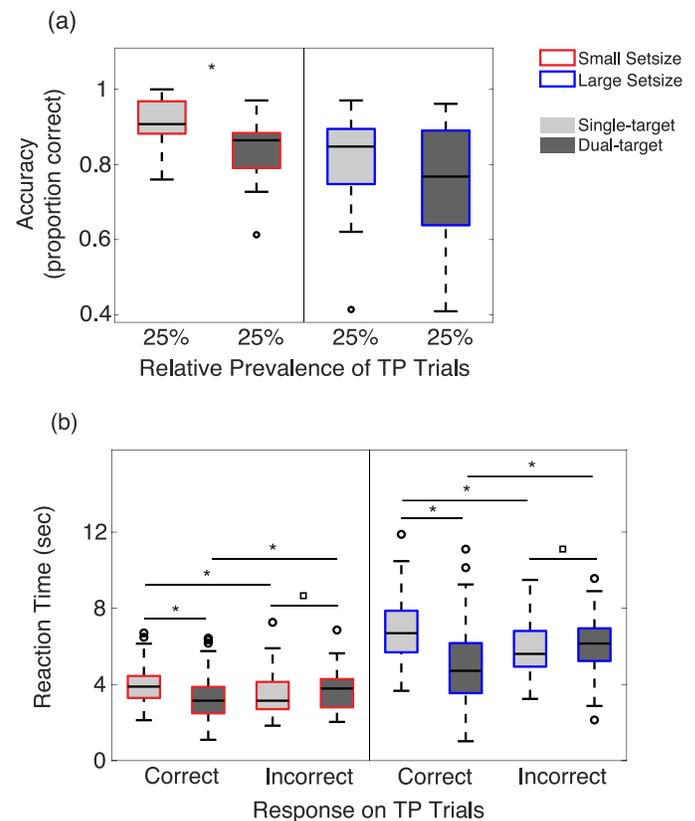

Figure 2. Accuracy and RTs of TP trials for Experiment 1, wherein single- and dual-target trials are equally likely. Each box covers data from the first to third quartile of a specified condition, and the central line is the median. The whiskers cover extreme data within 1.5 times the interquartile range (IQR). The dots represent individual scores outside 1.5× IQR. (a) Accuracy in small (left panels; red outline) and large (right panels; blue outline) set sizes. Single-target is the accuracy of finding the target in single-target trials, and dual-target is that of finding the second target in dual-target trials. The accuracy of each observer is averaged within condition, so each box plot represents 12 dots (i.e., 12 observers). (b) RTs from correct and incorrect trials in small and large set sizes. Incorrect trials are single-target trials with TA responses and dual-target trials with single-target responses. Single-target is RT on single-target trials, and dual-target is RT on dual-target trials. * indicates $p < 0.05$; □ indicates $BF < 0.3$.

< 0.001). In contrast, incorrect RTs (right two data points within each panel) did not differ for either set size (small: $\beta = -0.009$, $p = 0.53$, $BF = 0.26$; large: $\beta = -0.005$, $p = 0.49$, $BF = 0.27$). Thus, observers searched longer for a possible yet nonexistent additional target after finding the first target in single-target trials.

We then examined whether a speed-accuracy trade-off explains the pattern of accuracy. If it does, RTs for correct trials should be longer than those for incorrect trials. Indeed, single-target RTs (Figure 2b, light gray) were longer for correct than incorrect trials (small: $\beta = $

-0.02, $p = 0.02$; large: $\beta = -0.02$, $p < 0.001$). However, dual-target RTs (dark gray) were *shorter* for correct than incorrect trials (small: $\beta = 0.04$, $p < 0.001$; large: $\beta = 0.04$, $p < 0.001$). Thus, these results suggest that a speed-accuracy tradeoff can account for the accuracy pattern for single- but not dual-target response patterns.

Discussion

Experiment 1 demonstrates that in perceptually simple searches, SSMs can impair accuracy depending on the number of items in a display. Compared to the target in single-target trials, observers made more misses on the additional target in displays of 10 items but achieved the same level of accuracy in displays of 20 items, which is consistent with Fleck et al. (2010), who employed a set size of 25 without finding SSMs.

Different mechanisms seem to underlie the accuracy pattern for single- and dual-target trials in Experiment 1. RTs suggest that a speed-accuracy tradeoff could account for misses in single-target trials because observers searched a *shorter* time in incorrect than correct single-target trials. In other words, observers seemed to inspect fewer items in incorrect single-target trials, consistent with a previous finding that selection errors account for most single-target misses (Peltier & Becker, 2016). However, the *longer* RTs for incorrect than correct dual-target trials align with a previous finding that distractors are more likely to be refixated in incorrect dual-target trials (Cain et al., 2013). Thus, most misses in single- but not dual-target trials can be accounted for by a speed-accuracy tradeoff account.

The current set size–dependent SSM indicates that a smaller number of distractors exacerbates misses of an additional target. Because the current search requires attention to each item (Treisman & Gelade, 1980), the timing of attending the second target may be a factor that affects its detection. In a small set size, it might be more likely that the second target is fixated while attentional resources are still engaged with processing the first target (Sergent, Baillet, & Dehaene, 2005), a spatial version of the attentional blink (AB). In the classic AB, there is a decrease in accuracy in detecting a second target presented $\sim$200–500 ms after a found target (Raymond, Shapiro, & Arnell, 1992). Although an AB is usually elicited by a rapid serial visual presentation (RSVP) wherein items briefly appear one after another, Adamo et al. (2013) reported a self-generated AB-like effect in a search task (also see Cain et al., 2013), suggesting an AB may occur with simultaneously presented stimuli. Consistent with this finding, these authors have also shown that observers who recover more slowly from the AB in an RSVP task tend to make more SSMs in a search task (Adamo, Cain, & Mitroff, 2017). Based on the number of

stimuli in our task, in the small set size, the additional target would be twice as likely to be fixated shortly after detecting the first target than in the large set size, consistent with an increased potential for an AB.

Another potential contributor to our set size–dependent SSM is repetition blindness (RB; Kanwisher, 1987), which is a difficulty in detecting a second target after an identical target. When targets are letters, as is the case here, different orientations of the first and second target do not seem to reduce occurrences of RBs (Corballis & Armstrong, 2007). After the first target, RB is strongest for immediately subsequent targets and diminishes until it disappears after approximately five items (Kanwisher, 1987). In our task, in the small set size, it is more likely that the second target is seen within five items of the first target than in the large set size. RB can occur at a perceptual, memory or decision-making level (Friedenberg, 2012), and seem to represent separate mechanisms from the AB (Chun, 1997). Here, then, we may be seeing a combination of AB and RB effects on search accuracy.

Taken together, these results show that SSMs can occur in perceptually simple searches, potentially depending on the average number of distractors between finding the first and additional target. Experiment 1 lays the foundation for Experiment 2 to examine the effect of the relative prevalence of multiple targets.

Experiment 2: Relative prevalence of multiple targets

Research has shown that increasing target prevalence reduces misses in single-target trials (Godwin et al., 2015; Wolfe et al., 2005). The aim of Experiment 2 was to examine whether the relative prevalence of dual- versus single-target trials affects the accuracy of finding an additional target (i.e., influences the occurrence of SSMs).

Methods

Participants

There were 12 observers, all of whom reported normal or corrected-to-normal vision and no history of mental or neurological disorders. They gave informed consent and received course credits or AU\$15 per hour for a $\sim$110-min. experiment. The study was approved by the Macquarie University Human Research Ethics Committee.

Procedure and design

The apparatus and stimuli in Experiment 2 were essentially identical to those in Experiment 1, with the

exception of the ratio of the single- and dual-target displays. Each participant completed two conditions of different target prevalence without being informed of the prevalence, in separate blocks. Both conditions had 50% TA trials, but the proportions of single- and dual-target trials differed within the 50% TP trials. In the rare dual-target condition, there were 40% single-target trials and 10% dual-target trials (Figure 1c). In the rare single-target condition, there were 10% single-target trials and 40% dual-target trials (Figure 1d). The order of the conditions was counterbalanced across observers. Each participant completed 24 practice trials and 1,200 experimental trials across both conditions (i.e., 600 trials per condition). Observers could take a break after every 100 trials and received feedback of correct or wrong based on their accuracy after every trial.

Results

All observers in Experiment 2 performed well ($>75\%$ correct) and were included for analyses. With exclusion criteria identical to those of Experiment 1 we removed 0.33% of trials.

We focused our analyses on the influence of dual-target trial prevalence on SSMs. We also looked at RT data to explain accuracy differences. Because the duration of Experiment 2 was much longer than that of Experiment 1, we analyzed RT data within-experiment. As the difference in performance between set sizes was robust, as in Experiment 1, we analyzed the data separately by set size (rare dual-target condition accuracy for small vs. large: $\beta = -0.08$, $p < 0.001$; RT: $\beta = -0.01$, $p < 0.001$; rare single-target condition accuracy for small vs. large: $\beta = -0.08$, $p < 0.001$; RT: $\beta = -0.01$, $p < 0.001$). The justification for GLMM fitting was the same as in Experiment 1.

Half of the observers did the rare dual-target condition first and the other half did the rare single-target condition first. Thus, we first checked whether order affected model fittings and generalizability. We compared the model containing number of targets, prevalence, order, and their interactions ($df = 9$) with the nested model without order and its interactions with other factors ($df = 5$), to examine their effects on accuracy. BIC favors the model without order over that with order, suggesting that order is not a major influence on the results. We therefore analyzed the effect of number of targets, prevalence, and their interaction on accuracy. The main effect of number of targets was significant for both the small ($\beta = -2.12$, $p < 0.001$) and the large ($\beta = -1.26$, $p < 0.001$) set size. The main effect of prevalence was significant for the small ($\beta = -1.11$, $p = 0.007$) but not the large ($\beta = -0.35$, $p = 0.32$, $BF = 0.44$) set size. The interaction is significant for both small ($\beta = 1.22$, $p < 0.001$; see Figure 3a, red outline) and

large ($\beta = 0.68$, $p = 0.002$) set sizes (Figure 3a, blue outline).

The interaction between number of targets and prevalence is due to a larger effect of prevalence on the accuracy for the additional targets in dual-target trials than for the targets in single-target trials (Figure 3a). In the rare dual-target condition, observers were more accurate on single- than dual-target trials in both small ($\beta = -0.89$, $p < 0.001$, $OR = 0.41$) and large ($\beta = -0.61$, $p < 0.001$, $OR = 0.55$) set sizes. In the rare single-target condition, however, their accuracy did not differ between single- and dual-target trials in either small ($\beta = 0.32$, $p = 0.10$; $BF = 0.69$) or large set size ($\beta = 0.10$, $p = 0.53$; $BF = 0.26$). Note the evidence for the null is weak in the small set size case.

We then tested how prevalence of dual-target trials affects accuracy in single-target trials, and found that accuracy for rare single-target trials was *not* lower than that for prevalent single-target trials (Figure 3a, light gray boxplots). Compared to prevalent single-target trials, accuracy on rare single-target trials did not differ for the small set size ($\beta = 0.12$, $p = 0.51$, $BF = 0.26$) and it was actually higher for the large set size ($\beta = 0.32$, $p = 0.04$). This finding seems inconsistent with a speed-accuracy tradeoff account, which would predict lower accuracy for rare single-target trials due to a bias towards fast responses for the more dual-target prevalent trials.

As in Experiment 1, we examined whether the SSMs were the result of changes in RTs (Figure 3b). If this were the case, observers should have searched longer in single- than dual-target trials in the rare dual-target condition but not in the rare single-target condition. We did not find this pattern. Correct RTs for the rare dual-target condition were longer for single- than dual-target trials (small: $\beta = 0.06$, $p < 0.001$; large: $\beta = 0.06$, $p < 0.001$; left two data points in Figure 3b, Panels 1 and 3). Correct RTs for the rare single-target condition were also longer for single- than dual-target trials (small: $\beta = 0.12$, $p < 0.001$; large: $\beta = 0.09$, $p < 0.001$; left two data points in Panels 2 and 4). However, incorrect RTs for the rare dual-target condition in the small set size were shorter for single- than dual-target trials ($\beta = -0.03$, $p = 0.01$; right two data points in Panel 1), but those in the large set size did not differ ($\beta = 0.0002$, $p = 0.98$, $BF = 0.21$; right two data points in Panel 3). Incorrect RTs for the rare single-target condition were shorter for single- than dual-target trials (small: $\beta = -0.08$, $p < 0.001$; large: $\beta = -0.06$, $p < 0.001$; right two data points in Panels 2 and 4). Consistent with Experiment 1, SSMs did not seem to result from differences in RTs. Observers looked longer when it was likely there was a second target present than when it was highly unlikely there was another target.

The unaffected accuracy in the rare single-target condition might result from extended RTs, because observers might have searched longer when dual-target trials were more common versus single-target trials
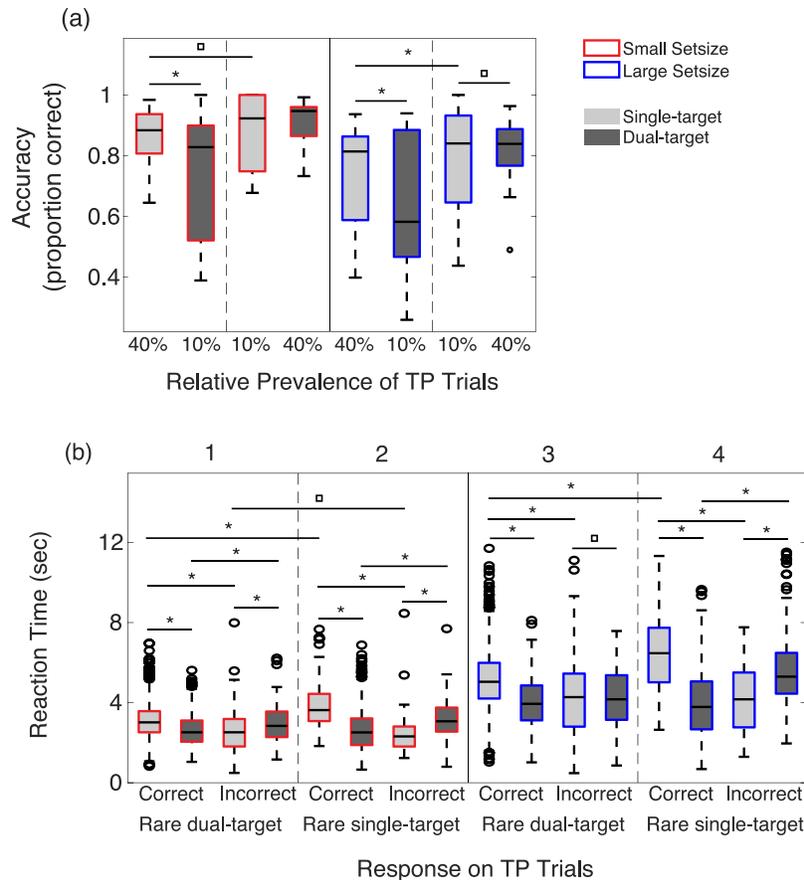
Figure 3. Accuracy and RTs of TP trials for Experiment 2, wherein dual-target trials are rare on half of TP trials and single-target trials are rare on the other half of TP trials. The boxplots have the same features as in Figure 2. (a) Accuracy in small (left panels; red outline) and large (right panels; blue outline) set sizes in rare dual-target and rare single-target conditions. Single-target is the accuracy of finding the target in single-target trials, and dual-target is that of finding the second target in dual-target trials. (b) RTs from correct and incorrect trials in small and large set sizes in rare dual-target and rare single-target conditions (blocked). Incorrect trials are single-target trials with TA responses and dual-target trials with single-target responses. Single-target is RT on single-target trials, and dual-target is RT on dual-target trials.

(i.e., more likely to find a target). Consistent with this, RTs for correct single-target trials differed between prevalence conditions (Figure 3b). Correct single-target RTs were shorter for the rare dual-target than rare single-target condition (small: $\beta = -0.06$, $p < 0.001$; large: $\beta = -0.04$, $p < 0.001$; light gray left data point in Panels 1 vs. 2 and 3 vs. 4). However, incorrect single-target RTs did not differ between conditions (small: $\beta = 0.001$, $p = 0.97$, $BF = 0.21$; large: $\beta = 0.01$, $p = 0.35$, $BF = 0.32$; light gray right data point in Panels 1 vs. 2 and 3 vs. 4). Thus, when there were more dual-target trials than single-target trials, observers tended to search longer and miss fewer targets.

We then compared RTs from correct and incorrect trials to test whether a speed-accuracy tradeoff could account for misses (Figure 3b). Single-target RTs for the rare dual-target condition were longer for correct than incorrect trials (small: $\beta = -0.06$, $p < 0.001$; large: $\beta = -0.03$, $p < 0.001$; light gray, Panels 1 and 3). Single-target RTs for the rare single-target condition were longer for

correct than incorrect trials (small: $\beta = -0.11$, $p < 0.001$; large: $\beta = -0.07$, $p < 0.001$; light gray, Panels 2 and 4). However, dual-target RTs for the rare dual-target condition in the small set size were shorter for correct than incorrect trials ($\beta = 0.03$, $p = 0.02$; dark gray, Panel 1), but those in the large set size did not differ ($\beta = 0.02$, $p = 0.06$, $BF = 0.92$; dark gray, Panel 3). Dual-target RTs for the rare single-target condition were shorter for correct than incorrect trials (small: $\beta = 0.05$, $p < 0.001$; large: $\beta = 0.07$, $p < 0.001$; dark gray, Panels 2 and 4). Thus, these results indicate that a speed-accuracy trade-off can explain the accuracy pattern for single-target trials but not that for dual-target trials.

Discussion

Experiment 2 reveals an effect of the prevalence of multiple targets on search accuracy. Trial-type preva-

lence has a stronger influence on the accuracy for dual-target trials than for single-target trials.

When dual-target trials were rare, observers missed more additional targets on dual-target trials (increased SSMs) than when dual-target trials were more prevalent. When single-target trials were rare, observers achieved the same level of accuracy as when single-target trials were more prevalent.

As in Experiment 1, a speed-accuracy tradeoff seems to explain most misses in single-target trials but not the misses in dual-target trials, since observers searched longer in correct single-target trials than incorrect single-target trials, but shorter in correct dual-target trials than incorrect dual-target trials. The *longer* RTs in incorrect than correct dual-target trials suggest that the search in incorrect dual-target trials is more thorough and should have been sufficient to find both targets. Thus, these SSMs seem unlikely to result from selection errors, which seem to be the main contributor to missing a single target (Peltier & Becker, 2016). Instead, these SSMs may reflect identification errors, perhaps due partially to the limits on deploying attention over time (i.e., a limit on temporal attention). We suggest this might be a combination of AB and RB phenomena, but it requires more targeted studies to explore this fully.

Our data show that the prevalence of multiple targets affects accuracy. More SSMs occur in the rare than frequent dual-target trials; in single-target trials, performance is not impaired by low prevalence of this trial type (miss rates of rare single targets were not higher than those of frequent single targets; Figure 3a). This is in contrast to the finding of Fleck et al. (2010), who did not find an interaction when comparing their *rare* dual-target condition with an *equiprobable* (single-vs. dual-target) condition. Their rare dual-target condition was identical to ours (10% total trials), but the interaction became evident when comparing the *rare* dual-target condition with the *frequent* dual-target condition (the rare single-target condition). Therefore, it seems likely that their difference in prevalence between dual-target trials may have been too small to reveal the interaction. Accuracy for the frequent dual-target trials in the rare single-target condition seemed unaffected by SSMs, whereas the rare dual-target trials showed higher SSM rates. This is particularly surprising given the difficulties in countering SSMs (Biggs, 2017). RTs suggest that these frequent dual-target trials raise the quitting threshold because observers searched longer after finding the first target than in the rare dual-target condition, leading presumably to a more thorough search (Wolfe & van Wert, 2010) and so fewer misses. Future studies may be able to exploit this interaction to reduce search errors.

Generally, Experiment 2 demonstrates that low prevalence of multiple targets impairs search accuracy.

Fortunately, high prevalence of multiple targets prevents additional misses, which might provide a potential avenue for reducing life-threatening search misses.

General discussion

The present study demonstrates that the relative prevalence of single- and dual-target trials affects search accuracy. Subsequent search misses increased when dual-target trials were rarer than single-target trials, but not when the reverse was true. With appropriate applications, this might suggest an avenue to reduce errors in real-world searches.

Although carefully inspecting every item would minimize misses, it is impractical in most situations (Cain, Vul, Clark, & Mitroff, 2012; Chun & Wolfe, 1996). In the current data, misses in single-target trials seem attributable mainly to selection errors, because observers quit earlier on incorrect relative to correct single-target trials, consistent with the finding that misses in single-target trials stem mainly from prematurely terminating search (Chun & Wolfe, 1996; Peltier & Becker, 2017a; Rich et al., 2008). In contrast, misses on dual-target trials do not seem to result from selection errors or premature quitting. Observers searched longer in incorrect than correct dual-target trials, suggesting their SSMs reflect a different cause.

The present SSMs cannot be fully explained by existing theories. The satisfaction theory predicts the failure to continue search after finding the first target (Tuddenham, 1962), but here, after the initial detection, observers kept searching longer than the theoretical requirement for detecting both targets. The perceptual-set theory proposes more subsequent misses if two targets share fewer perceptual similarities with the prediction that identical targets should not be subject to SSMs (Berbaum et al., 1991). Here, however, using identical targets (albeit sometimes rotated) did not eliminate SSMs. The resource-depletion theory suggests that a found target consumes working memory that otherwise benefits subsequent search (Cain & Mitroff, 2013). This would predict more SSMs in larger versus smaller set sizes, because memorizing a target among more distractors is more cognitively consuming (Cain & Mitroff, 2013). However, we found SSMs only in our small set size when single- and dual-target trials are equally likely (Experiment 1), and the effect size is larger for the small than the large set size when dual-target trials are rare, speaking against the resource-depletion theory. Thus, the current SSMs call for additional explanations.

The current study and the resource-depletion theory ostensibly disagree on the cognitive stage the first target

may occupy that hinders the detection of the second target. This is likely, however, to result from methodological differences. Specifically, Cain and Mitroff's (2013) study cannot assess the contribution of temporal attention to SSMs because they used a single set size of 25 and could not assess the average timing of attending an additional target relative to the first. The current study cannot assess working memory because we did not manipulate memory load. We propose that some missed targets here were never attended and so never reached working memory. Accordingly, we suggest that attention complements working memory to contribute to SSMs. Converging evidence shows that attention modulates the accessibility of information in the working memory (Gazzaley & Nobre, 2012), and the availability of working memory can determine the extent to which an item is attended (de Fockert, Rees, Frith, & Lavie, 2001; Lavie & de Fockert, 2005). Attention and working memory also recruit overlapping cortical regions (Mayer et al., 2007). Thus, the limited capacity of *an attention-memory circuit* would be a more comprehensive cognitive source for elevated misses in multitarget searches.

Previous studies suggest that processing an initial target occupies attentional resources and hence temporarily impairs processing a second target (Adamo et al., 2013; Adamo et al., 2017; Berbaum et al., 1991). The current set size–dependent SSM is consistent with this view. Although processing a T is much less perceptually demanding than processing a cancerous tumor or an x-rayed bag, it still occupies attention for approximately 500 ms according to research on AB and RB (Kanwisher, 1987; Sergent et al., 2005). RB seems an additional difficulty to an AB for recognizing a repeated item (Chun, 1997). Here, we speculate that we see more SSMs in our small set size because there is a greater likelihood that the two targets are more likely to be fixated close together in time, which may induce an AB, RB, or both.

There are also neuroimaging data consistent with the finding that processing an initial target interferes with processing the second. Sergent et al. (2005) showed that an item is blinked when related cortical areas are processing the previously attended item. Williams, Visser, Cunnington, and Mattingley (2008) showed that AB deficits can modulate the earliest cortical processing in the primary visual cortex. If items are identical, the perceptual deficit becomes even worse (Koivisto & Revonsuo, 2008). Here, given the use of identical targets, the greater effect size of SSMs in the small than large set size may stem from combined effects of the AB and RB. In Experiment 1, SSMs only occurred in the small set size, but in Experiment 2, we saw SSMs in both set sizes for the rare dual-target condition. This would be consistent with decreased prevalence exacerbating attentional capacity limit effects. We propose

that the low prevalence of dual-target trials sets an expectation on the number of targets in a display, which exerts a top-down influence further impairing the detection of additional targets in close succession.

In addition to the temporal allocation of attention to items on a display, the spatial arrangement of items can also affect the likelihood of detecting a target. A target cluttered by nearby distractors can more greatly tax attention, rendering the next target more likely to be missed (Adamo et al., 2015). Such a crowding effect is unlikely to cause the current SSMs, because there were *more* misses of the additional targets in the *less* cluttered small set size. The current SSMs, therefore, seem likely to result from the temporal limit of attention without particular impacts of spatial clutter.

Compared to the stimuli here, real-world targets are likely to demand greater attentional resources, perhaps for longer, due to their greater perceptual complexity. They can also contain identical targets. Additionally, experts such as radiologists tend to allocate a consistent amount of time to each trial regardless of the number of targets (Berbaum et al., 1991; Biggs & Mitroff, 2013). Here, nonprofessional observers searched longer on incorrect than correct dual-target trials. Given misses in multitarget searches do not reduce with expertise, experts could be more subject to ABs and RB when the number of targets increases, necessitating future research to investigate how ABs and RB affect expert searchers.

Search errors were not elevated when there was a high likelihood of finding an additional target. Specifically, when dual-target trials were more prevalent, the accuracy on single-target and dual-target trials was as high as that of the more prevalent single-target trials. This seems likely to stem from the effect of global prevalence on search behavior. When observers expected more dual-target trials, they searched more thoroughly on each trial. In single-target search, the global prevalence of the last ∼40 trials seems to affect search behavior (Ishibashi et al., 2012; Wolfe & van Wert, 2010). In multitarget search, global prevalence can affect foraging for a varying number of targets (e.g., zero to seven; Cain et al., 2012; Fougnie, Cormiea, Zhang, Alvarez, & Wolfe, 2015). However, the quitting rule differs in foraging versus the present study (Biggs, 2017). During foraging, observers voluntarily terminate a search knowing that there may be targets left (Hills et al., 2014), whereas here observers were instructed to find *all* targets. Although we instructed observers to search exhaustively in all conditions, they still searched longer in higher dual-target prevalence. Note that we provided feedback on *all* trials, which would facilitate incorporating the higher global prevalence in the rare single-target condition, leading to a more thorough search and fewer misses. The higher expectation of finding a target also

seemed to increase strategic errors (i.e., guessing that there is another target without actually detecting it; Peltier & Becker, 2017b; Schwark, MacDonald, Sandry, & Dolgov, 2013), here reflected in the higher number of false alarms in the rare single-target condition than in other conditions (Table 1). This also is likely to contribute to the reduced misses in the rare single-target condition. Observers noticed the higher global prevalence when dual-target trials were more common versus single-target trials and so they searched longer and missed fewer targets.

This study may provide an avenue for improving real-world search through increasing the prevalence of multiple targets. When dual-target trials were more common, participants tended to search more thoroughly, which may counter the selection errors that are a major component of both SSMs and low-prevalence misses. Thus, our findings may help design countermeasures against both types of misses. Wolfe and colleagues (Wolfe et al., 2013; Wolfe et al., 2007) used single-target among TA trials, and showed that high-prevalence blocks reduce low-prevalence misses. They suggested practicing in high prevalence before work could reduce life-threatening misses. With our findings, it is reasonable to suggest that practicing a combination of more common multitarget and rarer single-target trials may assist in reducing additional errors due to low-prevalence and multiple targets in the field.

There are three caveats here. First, our stimuli were simpler than those in real-world searches. Second, the observers were nonprofessional. Third, the study was run in a laboratory not the field. However, expertise seems not to reduce low-prevalence (Wolfe et al., 2013) or multitarget misses (Biggs & Mitroff, 2013). Additionally, professional searchers make low-prevalence misses during performance evaluation, suggesting that searching in the field does not reduce low-prevalence misses (Wolfe et al., 2013). Finally, misses tend to increase with perceptual complexity (Wolfe et al., 2007), so using real-world stimuli would presumably increase misses relative to our findings here. Taken together, the prevalence of multiple targets affects search accuracy. Future studies should take advantage of this finding to improve real-world search.

Conclusions

Elevated misses in multitarget searches can occur with perceptually simple stimuli, due presumably to limitations on temporal attention. Low prevalence of multiple targets relative to a single target impairs search accuracy. In contrast, high prevalence of multiple targets counteracts miss errors. Future efforts on designing countermeasures should take into account

the relative prevalence of multiple targets to effectively reduce life-threatening search errors.

Acknowledgments

References

Adamo, S. H., Cain, M. S., & Mitroff, S. R. (2013). Self-induced attentional blink: A cause of errors in multiple-target search. *Psychological Science*, *24*, 2569–2574, https://doi.org/10.1177/0956797613497970.

Adamo, S. H., Cain, M. S., & Mitroff, S. R. (2015). Targets need their own personal space: Effects of clutter on multiple-target search accuracy. *Perception*, *44*, 1203–1214, https://doi.org/10.1177/0301006615594921.

Adamo, S. H., Cain, M. S., & Mitroff, S. R. (2017). An individual differences approach to multiple-target visual search errors: How search errors relate to different characteristics of attention. *Vision Research*, *141*, 258–265, https://doi.org/10.1016/j.visres.2016.10.010.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge: Cambridge University Press.

Berbaum, K. S., Franken E. A., Jr., Dorfman, D. D., Rooholamini, S. A., Coffman, C. E., Cornell, S. H., . . . Smith, T. P. (1991). Time course of satisfaction of search. *Investigative Radiology*, *26*, 640–648.

Biggs, A. T. (2017). Getting satisfied with "satisfaction of search": How to measure errors during multiple-target visual search. *Attention, Perception, & Psychophysics*, https://doi.org/10.3758/s13414-017-1300-2.

Biggs, A. T., & Mitroff, S. R. (2013). Different

predictors of multiple-target search accuracy between nonprofessional and professional visual searchers. *The Quarterly Journal of Experimental Psychology*, *67*, 1335–1348, https://doi.org/10.1080/17470218.2013.859715.

Biggs, A. T., & Mitroff, S. R. (2015). Improving the efficacy of security screening tasks: A review of visual search challenges and ways to mitigate their adverse effects. *Applied Cognitive Psychology*, *29*, 142–148.

Cain, M. S., Adamo, S. H., & Mitroff, S. R. (2013). A taxonomy of errors in multiple-target visual search. *Visual Cognition*, *21*, 899–921, https://doi.org/10.1080/13506285.2013.843627.

Cain, M. S., Biggs, A. T., Darling, E. F., & Mitroff, S. R. (2014). A little bit of history repeating: Splitting up multiple-target visual searches decreases second-target miss errors. *Journal of Experimental Psychology: Applied*, *20*, 112–125, https://doi.org/10.1037/xap0000014.

Cain, M. S., Dunsmoor, J. E., LaBar, K. S., & Mitroff, S. R. (2011). Anticipatory anxiety hinders detection of a second target in dual-target search. *Psychological Science*, *22*, 866–871, https://doi.org/10.1177/0956797611412393.

Cain, M. S., & Mitroff, S. R. (2013). Memory for found targets interferes with subsequent performance in multiple-target visual search. *Journal of Experimental Psychology: Human Perception & Performance*, *39*, 1398–1408, https://doi.org/10.1037/a0030726.

Cain, M. S., Vul, E., Clark, K., & Mitroff, S. R. (2012). A Bayesian optimal foraging model of human visual search. *Psychological Science*, *23*, 1047–1054, https://doi.org/10.1177/0956797612440460.

Chun, M. M. (1997). Types and tokens in visual processing: A double dissociation between the attentional blink and repetition blindness. *Journal of Experimental Psychology: Human Perception & Performance*, *23*, 738–755.

Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology*, *30*, 39–78.

Corballis, M. C., & Armstrong, C. (2007). Repetition blindness is orientation blind. *Memory and Cognition*, *35*, 372–380, https://doi.org/10.3758/BF03193458.

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*, 7–29, https://doi.org/10.1177/0956797613504966.

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001, March 2). The role of working memory in visual selective attention. *Science*, *291*, 1083–1086.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontier in Psychology*, *5*, 1–17, https://doi.org/10.3389/fpsyg.2014.00781.

Fleck, M. S., Samei, E., & Mitroff, S. R. (2010). Generalized 'satisfaction of search': Adverse influences on dual-target search accuracy. *Journal of Experimental Psychology: Applied*, *16*, 60–71.

Fougnie, D., Cormiea, S. M., Zhang, J., Alvarez, G. A., & Wolfe, J. M. (2015). Winter is coming: How humans forage in a temporally structured environment. *Journal of Vision*, *15*(11):1, 1–11, https://doi.org/10.1167/15.11.1. [PubMed] [Article]

Friedenberg, J. (2012). *Visual attention and consciousness*. New York, NY: Psychological Press.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*, 129–135, https://doi.org/10.1016/j.tics.2011.11.014.

Godwin, H. J., Menneer, T., Cave, K. R., Thaibsyah, M., & Donnelly, N. (2015). The effects of increasing target prevalence on information processing during visual search. *Psychonomic Bulletin and Review*, *22*, 469–475.

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, *32*, 41–370, https://doi.org/10.3102/1076998606298043.

Heskes, T. (1998). Bias/variance decompositions for likelihood-based estimators. *Neural Computation*, *10*, 1425–1433, https://doi.org/10.1162/089976698300017232.

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., Couzin, I. D., & Group, C. S. R. (2014). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences*, *19*, 46–54, https://doi.org/10.1016/j.tics.2014.10.004.

Hout, M. C., Walenchok, S. C., Goldinger, S. D., & Wolfe, J. M. (2015). Failures of perception in the low-prevalence effect: Evidence from active and passive visual search. *Journal of Experimental Psychology: Human Perception & Performance*, *41*, 977–994, https://doi.org/10.1037/xhp0000053.

Ishibashi, K., Kita, S., & Wolfe, J. M. (2012). The effects of local prevalence and explicit expectations on search termination times. *Attention, Perception, & Psychophysics*, *74*, 115–123, https://doi.org/10.3758/s13414-011-0225-4.

Jeffreys, H. (1961). *The theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Kanwisher, N. G. (1987). Repetition blindness: Type recognition without token individuation. *Cognition*, *27*, 117–143.

Kleiner, M., Brainard, D., & Pelli, D. G. (2007). What's new in Psychtoolbox-3? *Perception 36 ECVP Abstract Supplement*.

Koivisto, M., & Revonsuo, A. (2008). Comparison of event-related potentials in attentional blink and repetition blindness. *Brain Research*, *1189*, 115–126, https://doi.org/10.1016/j.brainres.2007.10.082.

Lavie, N., & de Fockert, J. W. (2005). The role of working memory in attentional capture. *Psychological Bulletin & Review*, *12*, 669–674.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontier in Psychology*, *6*, 1–16, https://doi.org/10.3389/fpsyg.2015.01171.

Maddox, W. T. (2002). Toward a unified theory of decision criterion learning in perceptual categorization. *Journal of the Experimental Analysis of Behavior*, *78*, 567–595, https://doi.org/10.1901/jeab.2002.78-567.

Madigan, D., & Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using Occams window. *Journal of the American Statistical Association*, *89*, 1535–1546, https://doi.org/10.2307/2291017.

Mayer, J. S., Bittner, R. A., Nikolic, D., Bledowski, C., Goebel, R., & Linden, D. E. (2007). Common neural substrates for visual working memory and attention. *Neuroimage*, *36*, 441–453, https://doi.org/10.1016/j.neuroimage.2007.03.007.

Nakayama, K., & Martini, P. (2011). Situating visual search. *Vision Research*, *51*, 1526–1537, https://doi.org/10.1016/j.visres.2010.09.003.

Peltier, C., & Becker, M. W. (2016). Decision processes in visual search as a function of target prevalence. *Journal of Experimental Psychology: Human Perception & Performance*, *42*, 1466–1476, https://doi.org/10.1037/xhp0000248.

Peltier, C., & Becker, M. W. (2017a). Individual differences predict low prevalence visual search performance. *Cognitive Research: Principles and Implications*, *2*, 1–11, https://doi.org/10.1186/s41235-016-0042-3.

Peltier, C., & Becker, M. W. (2017b). Target-present guessing as a function of target prevalence and accumulated information in visual search. *Attention, Perception, & Psychophysics*, *79*, 1064–1069, https://doi.org/10.3758/s13414-017-1297-6.

Preacher, K. J., Zhang, G. J., Kim, C., & Mels, G. (2013). Choosing the optimal number of factors in exploratory factor analysis: A model selection perspective. *Multivariate Behavioral Research*, *48*, 28–56, https://doi.org/10.1080/00273171.2012.710386.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 849–860, http://dx.doi.org.simsrad.net.ocs.mq.edu.au/10.1037/0096-1523.18.3.849.

Rich, A. N., Kunar, M. A., Wert, M. J. V., 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):15, 1–17, https://doi.org/10.1167/8.15.15. [PubMed] [Article]

Rouder, J. N., Morey, R. D., & Province, J. M. (2013). A Bayes factor meta-analysis of recent extrasensory perception experiments: Comment on Storm, Tressoldi, and Di Risio (2010). *Psychological Bulletin*, *1*, 241–247, https://doi.org/10.1037/a0029008.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*, 225–237, https://doi.org/10.3758/PBR.16.2.225.

Schwark, J., MacDonald, J., Sandry, J., & Dolgov, I. (2013). Prevalence-based decision undermine visual search. *Visual Cognition*, *21*, 541–568, https://doi.org/10.1080/13506285.2013.811135.

Schwark, J., Sandry, J., & Dolgov, I. (2013). Evidence for a positive relationship between working-memory capacity and detection of low-prevalence targets in visual search. *Perception*, *42*, 112–114, https://doi.org/10.1068/p7386.

Schwarzkopf, S. (2014). Bayes factors Matlab functions. Retrieved from https://figshare.com/articles/Bayes_Factors_Matlab_functions/1357917

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, *8*, 1391–1400, https://doi.org/10.1038/nn1549.

Smith, M. J. (1967). *Error and variation in diagnostic radiology*. Springfield, IL: Charles C. Thomas.

Smith, P. A., & Turnbull, L. S. (1997). Small cell and 'pale' dyskaryosis. *Cytopathology*, *8*, 3–8.

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

Tuddenham, W. J. (1962). Visual search, image organization, and reader error in roentgen diagnosis. Studies of the psychophysiology of roentgen

image perception. *Radiology*, *78*, 694–704, https://doi.org/10.1148/78.5.694.

Williams, M. A., Visser, T. A. W., Cunnington, R., & Mattingley, J. B. (2008). Attenuation of neural responses in primary visual cortex during the attentional blink. *The Journal of Neuroscience*, *28*, 9890–9894, https://doi.org/10.1523/JNEUROSCI.3057-08.2008.

Wolfe, J. M., Brunelli, D. N., Rubinstein, J., & Horowitz, T. S. (2013). Prevalence effects in newly trained airport checkpoint screeners: Trained observers miss rare targets, too. *Journal of Vision*, *13*(3):33, 1–9, https://doi.org/10.1167/13.3.33. [PubMed] [Article]

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 & Performance*, *15*, 419–433, https://doi.org/10.1037/0096-1523.15.3.419.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005, May 26). Cognitive psychology: Rare targets are often missed in visual search. *Nature*, *435*, 439–440, https://doi.org/10.1038/435439a.

Wolfe, J. M., Horowitz, T. S., Van 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: General*, *136*, 623–638, https://doi.org/10.1037/0096-3445.136.4.623.

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

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 274–296, https://doi.org/10.1037/0278-7393.33.2.274.

Zhang, J., & Yu, K. F. (1998). What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *Journal of the American Medical Asssociation*, *280*, 1690–1691, https://doi.org/10.1001/jama.280.19.1690.