Published in final edited form as:
J Exp Psychol Appl. 2010 March ; 16(1): . doi:10.1037/a0018629.

Generalized "Satisfaction of Search": Adverse Influences on Dual-Target Search Accuracy

Mathias S. Fleck,
Department of Psychology and Neuroscience, Center for Cognitive Neuroscience, Duke University

Ehsan Samei, and
Carl E. Ravin Advanced Imaging Laboratories, Department of Radiology, Duke University Medical Center

Stephen R. Mitroff
Department of Psychology and Neuroscience, Center for Cognitive Neuroscience, Duke University

Abstract

The successful detection of a target in a radiological search can reduce the detectability of a second target, a phenomenon termed *satisfaction of search* (SOS). Given the potential consequences, here we investigate the generality of SOS with the goal of simultaneously informing radiology, cognitive psychology, and nonmedical searches such as airport luggage screening. Ten experiments utilizing nonmedical searches and untrained searchers suggest that SOS is affected by a diverse array of factors, including (1) the relative frequency of different target types, (2) external pressures (reward and time), and (3) expectations about the number of targets present. Collectively, these experiments indicate that SOS arises when searchers have a biased expectation about the low likelihood of specific targets or events, and when they are under pressure to perform efficiently. This first demonstration of SOS outside of radiology implicates a general heuristic applicable to many kinds of searches. In an example like airport luggage screening, the current data suggest that the detection of an easy-to-spot target (e.g., a water bottle) might reduce detection of a hard-to-spot target (e.g., a box cutter).

Keywords

visual search; satisfaction of search; luggage screening; radiological search

Given the potentially harmful consequences of missing an abnormality in a radiological scan, radiological research has scrutinized the circumstances of successful and failed target detection. Among other findings, it has been shown that a specific target is more likely to be missed when it is accompanied by an additional abnormality than when it is the only target in a radiological examination (e.g., Tuddenham, 1962). That is, an abnormality has a lower rate of detection when presented in the same image as another problem spot than when presented alone (see Berbaum, Franken, Caldwell, & Schartz, 2009, for a recent review). This phenomenon was originally characterized as a visual search that is discontinued once the searcher finds a target and then becomes "satisfied" with the "meaning" of an image

Correspondence concerning this article should be addressed to Stephen R. Mitroff, Duke University, LSRC Building, Box 90999, Durham, NC 27708. mitroff@duke.edu.

(Tuddenham, 1962). Such “satisfaction of search” (SOS) errors remain an acknowledged problem in radiologic examinations and have been demonstrated in chest radiography (e.g., Berbaum et al., 1998; Samuel, Kundel, Nodine, & Toto, 1995), abdominal radiography (e.g., Franken et al., 1994), osteoradiology (e.g., Ashman, Yu, & Wolfman, 2000), and multiple trauma patients (e.g., Berbaum et al., 1994), although the original “discontinued-search” account for such effects has not seen strong support (Berbaum et al., 2009).

SOS generalizes across multiple radiological domains and extends, at a minimum, to certain cytological searches as well (Bowditch, 1996). But do SOS errors arise in other visual searches beyond the medical domain? The goal of the current article is to explore the parameters in which SOS occurs in nonmedical visual search in order to simultaneously inform multiple fields. First, expanding the study of SOS into nonmedical searches will contribute to the discussion about the cause of SOS within radiology. Recent radiography studies have suggested that SOS effects can arise from scanning errors (Berbaum et al., 1996; Samuel et al., 1995), recognition errors (Berbaum, Franken, Dorfman, Caldwell, & Krupinski, 2000), or decision errors (Franken et al., 1994). Understanding the parameters of SOS will be important in better specifying the source of such errors and developing methods to counteract them.

Second, establishing the scope of SOS errors in nonmedical searches can inform the nature of the basic cognitive processes broadly involved in visual search. SOS may reflect a general heuristic of the decision-making process (e.g., “satisificing”; see Simon, 1976) involved in any kind of visual search, and if so, such errors should be observable in nonradiological contexts. Recent studies incorporating searches with more than one type of target (e.g., Menneer, Barrett, Phillips, Donnelly, & Cave, 2007; Menneer, Cave, & Donnelly, 2009) have complemented and expanded general theories of visual search (e.g., Treisman & Gelade, 1980; Wolfe, 2007), and the current study can contribute further. Past research has examined various external factors on search such as item familiarity (e.g., Wang, Cavanagh, & Green, 1994) and the emotional salience of search items (e.g., Hanson & Hanson, 1988), and here we explore the effects of multiple targets.

Third, finding that SOS effects exist outside the medical realm would carry critical implications for tasks such as airport security searches. For example, does the presence of a water bottle in a luggage X-ray adversely affect the detectability of a pair of scissors also in the bag? While the commonalities between airport baggage screening and medical image searches have only briefly been considered together (e.g., Fleck & Mitroff, 2007; Gale, Mugglestone, Purdy, & McClumpha, 2000; Wolfe, Horowitz, & Kenner, 2005), given the dangerous implications, it is critical to determine whether multiple target errors might occur in airport security searches and to establish what properties of the search might be predictive of SOS in these critical situations.

In the present article, we explore the robustness of SOS in basic visual search tasks in which we systematically manipulated several diverse factors. Through 10 experiments we examine how accuracy in a dual-target search is affected by (1) the relative salience and frequency of different target types, (2) time and reward pressures, and (3) expectations about number of targets in a given display. We reveal that SOS errors (the reduced detection of a target when another target has also been detected in the same display) may reflect a default and generalized search heuristic that is sensitive to the interplay between several factors. Broadly, SOS appears to arise in time-pressured search tasks when the searchers have an expectation about the low likelihood of a specific target type appearing, in relation to a different target type.

Effect of Salience and Frequency on SOS (Experiments 1–3)

SOS in radiology is typically defined as an increased detection rate for a particular target (e.g., a lesion) when it is the only target in a radiographic image in comparison with when it is accompanied by an additional target (e.g., a pulmonary nodule). Usually the relative salience of the primary and secondary targets is not manipulated, and the targets are from different categories (e.g., a lesion and a pulmonary nodule). However, evidence from osteoradiology found that SOS was stronger when the added secondary target was more salient than the test target (Berbaum et al., 1994). Cytology research has additionally suggested that the frequency and conspicuity of particular targets may lead to missing subtle, smaller targets (Bowditch, 1996; DeMay, 1996, 1997). Here we are particularly interested in this situation with differing salience between the primary and secondary target given that, in response to the nature of recent terrorist threats to aircraft safety, airport baggage security regulations were broadened in August 2006 by adding new categories to the prohibited items list, including liquids and gels (Transportation Security Administration, 2009). These additions vastly increase both the frequency and salience of certain targets, as well as the possibility of multiple targets co-occurring in the same image. Here we ask how the addition of a highly salient and frequent target (e.g., a water bottle) might impact the detection of a less salient or infrequent target (e.g., a box cutter) that may be present in the same image. Manipulating the frequency of a particular target type may bias the searcher's expectation for that target and contribute to producing an SOS effect.

Method

All 10 experiments in this study utilized a similar search paradigm, detailed here. Any differences from this paradigm are noted in the Method subsection for each subsequent section. See Table 1 for a summary of all experiments and parameters.

Participants—Thirty individuals (Experiment 1: mean age = 18.9 years, *SD* = 1.2 years; Experiment 2: mean age = 18.4 years, *SD* = 1.0 years; Experiment 3: mean age = 18.7 years, *SD* = 0.5 years) from the Duke University community participated (10 in each experiment with 3, 3, and 4 male participants in each experiment, respectively). Each participant in this article completed only one experiment. All participants gave informed consent and received either course credit or $10. The experiments were conducted on a Dell Optiplex computer running Windows 2000 and programmed in Matlab 7.0 using the Psychophysics toolbox version 3.0 (Brainard, 1997).

Stimuli—The stimuli comprised two perpendicular lines slightly offset from each other (Ts and Ls, stroke width = 0.3°, subtending 1.3° × 1.3° total), with target Ts having a crossbar directly in the middle and with distractor Ls having the crossbar slid at variable distances away from the center (see Figure 1). Stimuli were presented on a rendered grayscale "cloud" background (brightness range = 10%–50% black) that differed on each trial (see Figure 1). The background served to increase overall difficulty and to model background noise that is commonly present in radiological searches. Distractor Ls were presented at varying shades of gray (range = 28%–66% black), and target Ts were presented at one or both of two visibility levels: high salience (range = 66%–70% black) or low salience (range = 28%–40% black). In this fashion, high-salience Ts were relatively easy to detect, and low-salience Ts were more difficult to detect. Each stimulus was placed with a slight spatial jitter within randomly selected cells of an invisible 8 × 7 grid subtending 25.4° × 19.1° at an approximate viewing distance of 60 cm.

Procedure and design—Each trial began with a cross appearing for 0.5 s at the center of the screen. The cross was replaced with the search array, which consisted of 25 items.

Participants were informed that there were 0, 1, or 2 target Ts to find within each display. Participants used the mouse to click on each detected target item and then clicked a blue button at the bottom of the screen labeled *DONE* to complete the trial. The DONE button appeared 3 s after the onset of each trial, and the mouse cursor was reset to the center of the screen after each trial. Participants could correct an error or a mis-click by clicking a yellow button at the bottom of the screen labeled *CLEAR* before completing the trial. Each trial had a time limit of 15 s, after which no further clicks were accepted and a message was displayed encouraging participants to try to finish searching and press the DONE button before time elapsed on subsequent trials. Responses made prior to the timeout were recorded and analyzed even if the DONE button was not pressed.

Trials were classified as one of four types on the basis of the number and the salience of the target Ts presented within the array of distractor Ls, resulting in trial types of no target, single high salience, single low salience, or dual target (both a high-salience target and a low-salience target present). Experiments 1–3 only varied in the relative proportion of trial types presented to each set of participants to bias their expectation about how frequently a high-salience or low-salience target might appear. High-salience single targets were as equally frequent as were low-salience single targets in Experiment 1, twice as frequent in Experiment 2, and three times as frequent in Experiment 3.

Each experiment began with one block of 20 practice trials that was matched to the trial-type frequency of the rest of the experiment. Participants were not informed about how often targets would be present, although the practice block gave some indication as to what to expect in the rest of the experiment. During practice, immediate feedback was provided on any false positive identification or missed targets. Participants then completed 250 test trials with no accuracy feedback, divided into 50-trial blocks.

Planned analyses—To assess SOS errors, the critical planned comparison for each experiment was how accurately the low-salience targets were detected when they were the only target in the display (low-salience single targets) versus how accurately the low-salience targets were detected in a dual-target trial, given that the high-salience target was also detected. Note that this calculation of SOS differs slightly from that typically used in radiology. Radiology studies have often used a small number of trials in which each specific image is presented twice, once with a single target and once with that same target accompanied by an additional target. Such studies typically examine SOS by looking for changes in receiver-operating characteristic (ROC) curves, which incorporate a reported confidence rating about each detected target. We did not utilize identical displays with or without an added target and instead compared a large number of single- and dual- target trials, which provide the power to look for generalized effects. We also did not include confidence ratings, given the reduced role of decision-making required with our simpler stimuli.

Results

Our primary measure of interest is detection accuracy for low-salience targets (response time data can be seen in Appendix A, and high-salience target accuracy data can be seen in Appendix B), and mean rates are presented in Table 1. Single-target accuracy was calculated as the number of single-target trial hits divided by the total number of single-target trials. Dual-target accuracy was calculated as the number of dual-target trials in which both targets were detected divided by the total number of dual-target trials in which the high-salience target was detected (and the low-salience target missed). This calculation of dual-target accuracy was chosen to focus specifically on the hypothesis that detection of a high-salience target interferes with detection of a low-salience target. Additionally, it offers a conservative measure of SOS; inclusion of dual-target trials in which both the low- and high-salience targets were missed would potentially serve to inflate any potential SOS effect by lowering the dual-target accuracy rate in relation to the single target rate.

A repeated-measures analysis of variance (ANOVA) was run on the low-salience target accuracy with the between-subjects factor of frequency (1×, 2×, and 3×) and the within-subject factor of number of targets (1 and 2). There was a main effect of number of targets, $F(1, 27) = 7.40$, $p = .01$, $\eta_p^2 = .22$, indicating that accuracy for low-salience targets was lower when coupled with a high-salience target. There was also an interaction between number of targets and frequency, $F(2, 27) = 5.02$, $p = .02$, $\eta_p^2 = .27$, such that single-target accuracy was increasingly better than dual-target accuracy as a function of increasing frequency.

Planned one-tailed *t* tests for low-salience targets revealed that the SOS effect was not significant at the 1 × frequency, $t(9) = 1.28$, $p = .23$, one-tailed, $d = -.23$, but was significant at the 2× frequency, $t(9) = 3.18$, $p = .01$, one-tailed, $d = .38$, and at the 3× frequency, $t(9) = 3.56$, $p < .01$, one-tailed, $d = .95$ (see Table 1). We used one-tailed tests for this statistical comparison because our a priori SOS prediction, based upon the radiology research, is that detection should be better on single-target trials than on dual-target trials. Additionally, because raw trial counts were chosen to balance single-target and dual-target trials for low-salience targets, we did not focus on high-salience dual-target costs. However, inspection of that data revealed no cost on dual-target search for high-salience targets when compared with single-target high-salience searches in any of these experiments (see Appendix B).

For all three experiments, the participants had a small proportion of trials in which they did not click DONE before the 15-s time limit (Experiment 1: *M* = 1.4%, *SD* = 1.2%; Experiment 2: *M* = 1.1%, *SD* = 0.6%; and Experiment 3: *M* = 1.5%, *SD* = 1.4%). There were few false alarms, which were calculated as the percentage of all trials with one or more false positive responses (Experiment 1: *M* = 2.8%, *SD* = 2.4%; Experiment 2: *M* = 1.6%, *SD* = 1.9%; and Experiment 3: *M* = 2.3%, *SD* = 3.7%).

Discussion

The results of Experiment 1 indicated that salience differences alone are not enough to trigger SOS errors. In a dual-target search wherein one target is highly salient and another is less salient and both are equally likely to be present in any given display, detection of the high-salience target does not lead to a higher miss rate for the low-salience target than when the low-salience target is presented by itself (Experiment 1). However, when the relative frequency of the high-salience targets increased (Experiments 2 and 3), participants missed low-salience targets more often on trials in which a high-salience (and highly expected) target was detected than on trials in which the low-salience target was presented alone. The target prevalence differences appear to lead the participants to shift their biases about the type of target to expect on any given trial. Interestingly, biases about the expected number of targets also changed accuracy: single-target low-salience trial accuracy increased as the number of single-target high-salience trials increased. Thus, while salience differences alone were not enough to induce SOS (Experiment 1), it appears that changing relative frequency of target types, and subsequent expectations about those target types, can indeed generate SOS errors. We will return to this result in the General Discussion section after exploring additional possible influences on SOS in Experiments 4–10.

Effect of Time Pressure on SOS (Experiments 4–6)

Radiological studies of SOS are typically self-paced paradigms modeled after routine radiograph examinations (e.g., Berbaum et al., 1998). Even though such studies have demonstrated significant SOS effects, no study has yet to directly test the effects of a specific time pressure that might be presented in radiographic workflows. A radiologist may feel compelled to process a minimum number of radiographs in a day, a pressure that is absent in a controlled radiology experiment. Relatedly, during high passenger flow, luggage screening typically occurs on the order of 3–5 s per inspection (Schwaninger, 2005), indicating the possibility of a strict time pressure (potentially even more so than that found in radiology). Here we ask whether time pressure plays a role in the SOS effect.

In Experiments 1–3, we utilized a 15-s time limit per trial. Although search tasks such as luggage screening and radiology may indeed be faced with time pressures, we wish to establish whether the SOS effects observed in Experiments 2 and 3 were potentially driven by time pressure, instead of or in addition to frequency. In Experiments 4–6, we double the amount of time allowed to search on each trial. We again manipulate the frequency of salient targets to explore the relationship between salience, frequency, and time pressure.

Method

The search task was identical to that in Experiments 1–3. Thirty individuals participated (Experiment 4: mean age = 18.6 years, *SD* = 1.0 years; Experiment 5: mean age = 19.1 years, *SD* = 1.2 years; Experiment 6: mean age = 18.8 years, *SD* = 1.4 years). There were 10 participants per experiment with 5, 4, and 4 male participants, respectively. In Experiments 4–6, each participant was allowed 30 s to search each display, which was twice as long as allowed in Experiments 1–3. High-salience targets were either 2× (Experiment 4) or 3× (Experiment 5) more frequent than were low-salience targets, just as in Experiments 2 and 3, respectively. Because Experiment 1 revealed no evidence for SOS, we did not repeat the 1× frequency distribution here. Likewise, to increase our chances of obtaining SOS with 30 s to search, we included a new condition with a more extreme salience distribution (Experiment 6; 6×).

Results

Mean detection rates are presented in Table 1. Accuracy data from Experiments 2–5 were submitted to a repeated-measures ANOVA with the within-subject factor number of targets (1 or 2) and between-subjects factors of frequency (2× or 3×) and time limit (15 s or 30 s), which revealed a significant main effect of number of targets, $F(1, 36) = 13.78$, $p = .001$, $\eta_p^2 = .28$, but no effects of time limit or frequency. Planned, follow-up one-tailed *t* tests for low-salience targets revealed no SOS effect for Experiment 4, $t(9) = 1.07$, $p = .31$, one-tailed, $d = .11$, and a nonsignificant effect for Experiment 5, $t(9) = 2.08$, $p = .07$, one-tailed, $d = .32$. Experiment 6, with the greater frequency of high- to low-salience targets, also generated a nonsignificant SOS effect, $t(9) = 2.11$, $p = .06$, one-tailed, $d = .70$. There were very few trials in which participants did not click DONE before the 30-s time limit (Experiment 4: *M* = 0.2%, *SD* = 0.3%; Experiment 5: *M* = 0.2%, *SD* = 0.3%; and Experiment 6: *M* = 0.1%, *SD* = 0.3%). There were few false alarms (Experiment 4: *M* = 1.4%, *SD* = 1.5%; Experiment 5: *M* = 1.6%, *SD* = 2.3%; and Experiment 6: *M* = 3.1%, *SD* = 3.7%).

Discussion

Although available time to search did not interact with frequency and number of targets, the notable absence of an interaction between frequency and number of targets, as observed in Experiments 1–3, potentially indicates that time limit may play some role in exacerbating the SOS effect. Comparing the one-tailed planned comparisons and associated effect sizes of Experiments 4 through 6 reveals much weaker SOS effects when participants were allowed 30 s to search, in comparison with Experiments 1 through 3 when participants were allowed

15 s to search. One possible reason that the factor of time limit was not significant was that the time limit may have pressured searchers only on a select number of especially “difficult” trials, depending on the vagaries of the randomized background noise and the particular clustering of stimuli. Once a particular pace is set by a participant, the actual time remaining in a trial may become less of a focus and therefore not impactful on behavior. While these data do not directly indicate that time limit plays a significant role in driving SOS, the reduced effect sizes of Experiments 4–6 in relation to Experiments 1–3 indeed suggest that time pressures should be minimized in critical tasks with more than one target, such as in radiology, cytology, or airport security. We more directly explore the link between time pressure and accuracy in Experiments 7 and 8, in which we examine SOS in searches that rewarded participants for minimizing time on task.

Interaction Between Time and Reward Pressures on SOS (Experiments 7 and 8)

While Experiments 1–6 collectively demonstrated that SOS does indeed generalize beyond medical searches, the role of external pressures such as a time limit remains open. Beyond time limit pressures, laboratory-based experiments with novice searchers typically do not address the host of other pressures that a searcher may face. For example, a radiologist, a cytologist, and an airport X-ray screener all have an immeasurably higher incentive to avoid missed targets than do participants volunteering in a laboratory study. Although it is extremely difficult to replicate in the lab such pressure to perform accurately, in Experiments 7 and 8 our goal was to add pressure by motivating participants through performance-based reward and to further explore the role of time pressure by directly linking overall performance to time on task. We also wished to establish that the generalized SOS effects found in Experiments 1–6 are not simply driven by a low motivation to find all targets. In many laboratory studies, there is a concern that participants may wish to complete the experiment as quickly as possible, and in a multiple-target paradigm this is a particularly hazardous prospect. Participants may feel that they have “adequately participated” on each trial upon finding any target, thereby manifesting the exact property we wish to explore. Although arguably this prospect is inherent as well in any nonlaboratory search task, a reward manipulation addresses whether the SOS effect may be attenuated when motivation to find all possible targets is increased.

Moreover, we wanted to allow participants the ability to “budget” their time across the length of the experiment to enable them to selectively balance accuracy versus time. In many tasks, searchers are not subject to the “fixed” time limit pressure as we have utilized it so far. Instead, they have the option of economically allocating time appropriately—spending more time on the cases that require additional attention while speeding through images deemed “easier” upon initial inspection. In this manner, the time pressure is less a per-image pressure as it is a per-session pressure (e.g., to search a given number of images within a day). In Experiments 7 and 8, we implement a performance-based reward system simulating a luggage screening task to see how and when SOS would arise when participants are motivated to perform accurately and are facing a time pressure across the experiment rather than within a trial.

Method

Twenty individuals (Experiment 7: mean age = 18.7 years, *SD* = 0.9 years, 4 men; Experiment 8: mean age = 19.1 years, *SD* = 1.1 year, 8 men) were paid $10 for their participation and had the opportunity to earn an extra $10 for high performance (see below).

The search task was identical to that in the previous experiments, except that we implemented an accuracy- and time-based reward system for performance on the task rather than a fixed time limit. Participants were instructed to manage a "line of luggage," represented by a row of luggage icons at the top of the screen above the searched area, which increased and decreased in the number of icons throughout the experiment. As participants completed each trial, regardless of accuracy on the trial, one icon was eliminated from the row. Icons were added to the row as a function of time spent on each trial. In Experiment 7, this time window was set individually for each participant on the basis of their average time spent on no-target trials in the practice block. Slower participants therefore accrued "luggage icons" at a slower rate than did faster participants, and this manipulation reduced the time pressure factor overall while still demanding that participants budget accuracy versus time spent searching. In Experiment 8, this time window for luggage accrual was fixed for all participants at 9 s per icon to establish a much greater time pressure and to determine whether participants would adapt to the relatively fast rate of luggage accrual.

The luggage line was tied to a point-based reward system. Participants were told that successfully responding to any trial (either with zero, one, or two targets) would always result in a gain of points, but that these points would be greater if the luggage line at the top of the screen was kept short. Participants were warned that missed targets would result in a significant loss of points, and that even when the luggage line reached its longest length (16 icons along the top of the screen), participants would continue to gain points for each correct trial. Participants were told that accuracy was the primary goal in the task, and that line-length management was a secondary goal. The computer tracked each participant's cumulative point total offscreen, and participants were informed that the top points scorer out of every 3 participants would receive an extra $10. After each block, participants were informed of their current point total, but they were never told about the total points of any previous participant.

Participants completed 50 practice trials with feedback, followed by 200 test trials with no feedback, divided into four blocks of 50 trials each. Each display contained 30 items. Similar to Experiments 3 and 5, we utilized a high-salience to low-salience single target trial ratio of 3:1 to replicate known SOS conditions.

Results

Mean detection rates are presented in Table 1. A repeated-measures ANOVA with within-subjects factor of number of targets (1 or 2) and a between-subjects factor of time pressure (variable or fixed) revealed a significant main effect of number of targets, $F(1, 18) = 4.50$, $p = .05$, $\eta_p^2 = .20$. There was a nonsignificant interaction between number of targets and time pressure, $F(1, 18) = 2.73$, $p = .12$, $\eta_p^2 = .13$. There was a small proportion of false alarms in each experiment (Experiment 7: $M = 0.7\%$, $SD = 1.0\%$; Experiment 8: $M = 2.1\%$, $SD = 1.8\%$). Experiment 7, with the rate of luggage accrual customized for each participant as a function of their practice block speeds, mirrored the general accuracy rate of Experiment 3, but did not exhibit an SOS effect, $t(9) = .93$, $p = .38$, one-tailed, $d = .07$. Alternately, Experiment 8, with the rate of luggage accrual fixed at 9 s across all participants, produced lower accuracies overall and a significant SOS effect, $t(9) = 3.21$, $p = .01$, one-tailed, $d = .57$ (see Table 1 for data and planned *t* test).

Discussion

Experiment 7 did not itself yield an SOS effect with its individually titrated luggage accrual (i.e., lower pressure), but Experiment 8, in which luggage icons appeared in a relatively

NIH-PA Author Manuscript

NIH-PA Author Manuscript

rapid and steady stream, exhibited a strong SOS effect. Examining effect sizes across the four experiments in which frequency and salience combinations were held constant (Experiments 3, 5, 7, and 8), we found the greatest SOS effects in the two experiments in which time pressure was highest: Experiment 3 (15-s time limit, $d = .95$) and Experiment 8 (rapid luggage accrual rate, $d = .57$), in contrast to Experiment 5 (30-s time limit, $d = .32$) and Experiment 7 (individualized luggage accrual rate, $d = .07$). Collectively, these data suggest that external pressures of reward and time increase SOS-type errors. Implications for the role of external pressures are further explored in the General Discussion.

Effect of Expectation About Number of Targets on SOS (Experiments 9–10)

Experiments 1–8 explored the influences of salience, frequency, and external pressure on multiple-target accuracy. To examine these issues, each of the experiments so far has utilized two types of targets: high-salience targets and low-salience targets. This was motivated by salience-related SOS effects in medical research (Berbaum et al., 1994; Bowditch, 1996) and an attempt to map onto the saliency and frequency differences found in current airport security screening. However, we also wished to determine the extent to which SOS may arise when two present targets are identical to one another, because many searches involve looking for multiple items that may be very similar.

Recent work has demonstrated that searches for multiple categories yield decreased accuracy when compared with single-category searches (Menneer et al., 2007, 2009). This multiple-category deficit has also been discussed in radiology with respect to SOS errors; radiologists may adopt a readiness to seek out a specific pattern in the image, at the expense of detecting other target types (Berbaum et al., 2009). In such a “perceptual set” explanation, the detection of one target type, say a tumor of a particular contrast, may bias searchers to selectively search for additional targets with the same perceptual features as that tumor and to discount the features of targets from different categories (Berbaum et al., 1990). Because the first eight experiments have all utilized two classes of targets, some of the SOS effects demonstrated so far may be related to multiple-category effects.

In Experiment 9, we remove all expectation biases about both the number of targets and the target salience by making secondary targets identical to the primary target and by making dual-target trials as equally likely as single-target trials. Thus, on any given trial, it is equally likely that there may be one or two targets in the display, and if there are two targets, they are identical in salience and shape. In Experiment 10, we maintain the identical-target manipulation while reintroducing the element of frequency and expectation by altering the ratio of single-target to dual-target trials. The goal is to determine whether SOS is sensitive to expectations about the number of targets likely to be present even when all targets are identical.

Method

Twenty individuals (Experiment 9: mean age = 19.9 years, *SD* = 1.0 years, 5 men; Experiment 10: mean age = 21.3 years, *SD* = 5.0 years, 3 men) participated. To minimize the contribution of the time pressure effect to SOS implicated by the first eight experiments, we again utilized a 30-s time limit rather than the reward-based paradigm. The other difference from the previous search parameters was that all targets were presented at the low-salience values of Experiments 1–8, resulting in three trial types: no target, single low salience, and dual low salience. Because there were no salience differences between targets, dual-target accuracy was calculated differently from previous experiments: Here the accuracy was calculated by dividing the number of dual-target trials in which both were detected by the sum of those trials plus all trials in which only one target was detected.

Dual-target trials in which both targets were missed were again excluded to focus on errors that arise as a consequence of detecting one target.

Results

Mean detection rates are presented in Table 1. A repeated-measures ANOVA with a within-subject factor of number of targets (1 or 2) and a between-subjects factor of single-to-dual frequency (1× or 4×) revealed little effect of number of targets, $F(1, 18) = 2.86$, $p = .11$, $\eta_p^2 = .14$, and no significant interaction between number of targets and trial type frequency, $F(1, 18) < 1$, $p = .51$, $\eta_p^2 = .02$. Planned, follow-up one-tailed *t* tests for low-salience targets revealed no SOS effect for Experiment 9, $t(9) = 1.12$, $p = .29$, one-tailed, $d = .13$, as was expected, but in Experiment 10, when single-target trials were four times more likely to occur than were dual-target trials, participants demonstrated a significant SOS effect, $t(9) = 3.10$, $p = .01$, one-tailed, $d = .45$. There was a small proportion of trials in which participants did not click DONE before the 30-s time limit (Experiment 9: *M* = 0.5%, *SD* = 0.9%; Experiment 10: *M* = 0.2%, *SD* = 0.3%). There were few false alarms (Experiment 9: *M* = 5.6%, *SD* = 5.9%; Experiment 10: *M* = 1.5%, *SD* = 1.1%).

Discussion

The lack of an SOS effect in Experiment 9 is unsurprising, particularly considering the absence of an effect in Experiments 1 and 4 in which differing salience levels may have biased expectation about target types and yet still yielded no SOS effect. Keeping all else the same and equating salience here thus predictably generated no SOS effect. However, when the relative frequency of single- and dual-target trials was manipulated in Experiment 10, participants missed more targets in dual-target conditions than in single-target conditions. It should be noted that the between-experiments manipulation of relative frequency of target types showed no interaction with number of targets, indicating that additional work may be necessary to more fully explore the relationship. However, comparison of the a priori one-tailed *t* tests and associated effect sizes offer some preliminary evidence that increased frequency of single-target trials biased participants' expectations and led to a greater relative proportion of missed dual-target trials. Experiment 10 indicates that a "perceptual set" or "multiple-category effect" account cannot entirely drive the SOS effects observed in the previous experiments, and that SOS is indeed sensitive to expectations about number of targets.

General Discussion

The 10 experiments presented here focused on the phenomenon of "satisfaction of search" (SOS), whereby the detection of one target is hindered by the successful detection of another target. This research was conducted with the goal of establishing the commonalities between visual search as it is studied in radiology and how it is studied in cognitive psychology. Because SOS may not be a problem exclusive to radiology, we have investigated here the generality of the effect and explored the parameters to which it is sensitive. By manipulating a set of search parameters, we have revealed several contexts in which SOS errors can be observed, even in a nonmedical search task, and with novice searchers. Importantly, this general finding of SOS outside of radiology suggests that the factors that modulate accuracy in multiple- target detection may be broadly applicable to many other critical search tasks such as airport security screening. These factors include (1) the relative salience and frequency of different target types, (2) external pressures of reward and time, and (3) expectation about the number of targets present.

Summary

Experiments 1–3 established that SOS is sensitive to the interaction between salience and the expectancy about certain target events: As the frequency of easy-to-detect high-salience targets increased in relation to difficult-to-detect low-salience targets, participants missed more of those low-salience targets in a dual-target condition than when the low-salience target was presented by itself. Interestingly, salience differences alone (Experiment 1) were not sufficient to induce SOS. Rather, it is the interaction between those salience differences and biased expectation about the differing target types that leads to SOS; in Experiments 2 and 3, we shifted the frequency of targets such that participants expected high-salience targets more often than low-salience targets, and consequently dual-target low-salience performance fell.

In Experiments 4–6, we examined the role of time pressure in SOS. A potential concern with the significant SOS effects in Experiments 2 and 3 is that a 15-s time limit may not be enough time to adequately search the display and find two targets. In Experiments 4–6, we doubled the time limit to 30 s, resulting in much weaker SOS effects.

In Experiments 7 and 8, we implemented a performance-based reward system to further explore the impact of time pressure and to determine whether increased motivation can offset the SOS effect. That is, if participants are given an incentive to be both accurate and fast, will they no longer reveal SOS? Experiment 8 revealed a strong effect of the number of targets on accuracy, suggesting that SOS arises even in motivated participants when external pressure is increased (here, in the form of optimizing accuracy and time on task).

Lastly, in Experiments 9 and 10 we focused on whether expectation about number of targets reduces detection of extra targets, even in the absence of multiple categories. We wished to determine whether SOS can arise even when all targets are perceptually identical. A failure to find an effect when targets are perceptually similar would suggest that SOS arises as a function of a “perceptual set” that biases the decision to terminate a search after the “preferred target” has been detected. However, even when targets were perceptually identical, we still observed SOS, indicating that SOS represents an effect distinct from the multiple- category effect previously demonstrated (Menneer et al., 2007, 2009).

Implications for SOS in Radiology

SOS has been extensively studied in radiology, and the present experiments attempt to establish the generality of the effect to determine whether and how SOS errors arise in nonmedical domains. However, these generalized results may in turn inform the attempts to minimize SOS within radiology. We discuss here a few possible implications, but their ultimate impact on radiology remains an open question given that the current search arrays differ from radiological scans and that our novice searchers do not have the benefit of extensive training and experience.

The theory that a “perceptual set” (Berbaum et al., 2009) may act to increase SOS errors is in accordance with the data from Experiments 1 through 8, which illustrate that when participants expected a particular visual pattern for the target (in the present studies, a high-salience T shape), the result was that successfully finding such a target interfered with detection of a perceptually different target (here, a low-salience T shape). However, Experiments 9 and 10 indicate that this explanation does not fully account for SOS errors, because perceptually similar targets continued to yield SOS effects when expectations about the number of targets were biased.

Additionally, although radiologists may be well trained in efficiently optimizing the trade-off between accuracy and the need to process at least a certain number of cases in a particular session, our data linking time pressure with SOS, both within a particular case (e.g., Experiment 2 vs. Experiment 4) and across a longer session (Experiment 7 vs. Experiment 8), should be taken into consideration. The influence of external pressures on the SOS effect (here, in a performance-based paradigm rewarding efficient responding) specifically emphasizes that radiological searches should minimize pressure to process a particular number of cases in a day.

Implications for Cognitive Psychology

Visual search is a well-studied cognitive task, and this is not the first comparison of searches for one versus multiple targets (e.g., Gibson, Li, Skow, Brown, & Cooke, 2000; Körner & Gilchrist, 2008; Menneer et al., 2007, 2009; Metlay, Sokoloff, & Kaplan, 1970; Takeda, 2004). However, the influence of multiple possible targets in cognitive studies of search has typically explored the impact on search efficiency, or the added cost of extra targets on time to search, rather than on search accuracy, as we examined here. In two notable exceptions, Menneer and colleagues (2007, 2009) have demonstrated an accuracy effect for search for multiple targets. In these studies, participants searched for multiple categories of targets, but there was never more than one individual target in any given trial. A performance decline was found in this multiple-category search design, but unlike here, these effects do not reflect the cost of having detected an added target within the same search. Rather, the performance decline was a function of increasing the possible search space to include more categories. This accuracy cost maps well onto the multiple target types utilized in the present article in the form of differing salience levels, and further indicates that expectation for a single type of target can affect search performance.

The search parameters we explore in this article (relative frequency of target types, time and reward pressure, and number of targets) are not novel in visual search studies, but the current experiments offer some of the first evidence about how these factors interact in a multiple-target paradigm to modulate accuracy. These results raise possible extensions to established theories of visual search. For example, Chun and Wolfe (1996) explored how an individual ends a search when a target is not present. The most conservative strategy is to search every possible item before declaring no target is present, but searchers do not do this in practice. Instead, they make assumptions and end their search on the basis of gathered information about the nature of the targets and the distractors. Additionally, they adjust their search time on the basis of misses and hits and also guess from time to time (Chun & Wolfe, 1996). These same principles likely come into play for accuracy in a multitarget search. Once one target is detected and a second might be present, how will searchers adjust their search behavior? Future research using the current paradigm and additional variables (e.g., changes in set size) should be fruitful.

More generally, the finding of SOS in a nonmedical context and in an untrained, naïve group of participants suggests that this type of error reflects a global search heuristic. Indeed, it may be adaptive to exhibit SOS as a means to maximize efficiency of time, just as the phenomenon of “inhibition of return” (Posner & Cohen, 1984) is thought to maximize the efficiency of search and foraging behavior by deprioritizing previously attended locations. Regardless of the mechanism, that SOS is a general effect, sensitive to pressures nonspecific to radiology, suggests that any visual search incorporating multiple simultaneously present targets must take into account the interactions we have demonstrated here, adding to the previously known data on visual search including effects of item frequency (e.g., Fleck & Mitroff, 2007; Rich et al., 2008; Van Wert, Horowitz, & Wolfe, 2009; Wolfe et al., 2005), familiarity (e.g., Wang et al., 1994), set size (e.g., Carter, 1982; Wolfe, 1998), target and distractor heterogeneity (e.g., Nagy & Thomas, 2003; Palmer, Verghese, & Pavel, 2000), emotionality of stimuli (e.g., Gerritsen, Frischen, Blake, Smilek, & Eastwood, 2008), and memory (e.g., Horowitz & Wolfe, 1998; Körner & Gilchrist, 2008), to name a few.

Implications for Luggage Screening and Other Related Searches

Given the generality of SOS, these results motivate a careful analysis of search tasks to determine whether similar multiple-target effects may be contributing to overall error rates in other socially critical searches. In the particular case of airport security X-ray screening, it is obviously a vital safety issue to ensure maximal detection rates in luggage screening, yet there has been relatively little cross-talk between radiology, cognitive psychology, and transportation security, despite the many factors common to visual search tasks in all three domains (for some examples, see Fiore, Scielzo, & Jentsch, 2004; Fleck & Mitroff, 2007; Gale et al., 2000; McCarley & Carruth, 2004; McCarley, Kramer, Wickens, Vidoni, & Boot, 2004; Menneer et al., 2007; Smith, Redford, Washburn, & Taglialatela, 2005; Wiegmann, McCarley, Kramer, & Wickens, 2006; Wolfe et al., 2005).

Most of the factors explored in this article can directly inform search tasks such as airport security. Notably, the interaction between salience and expectancy is directly relevant to current security searches. As of this writing, current security protocols mandate that airport screeners search for very salient and common targets such as water bottles, hair gel, soft drinks, and toothpaste, potentially at the expense of finding additional targets that may be better concealed and less frequent, such as scissors, box cutters, or pocketknives. Although extensive training assuredly emphasizes an exhaustive search no matter the search results, the finding of salience effects in osteoradiology (Berbaum et al., 1994) and cytology (Bowditch, 1996) indicates that training may not override this basic search heuristic.

We also found that time pressure likely exacerbates the SOS effect. Critically, airport security searches are much shorter and more numerous in a session than are radiograph examinations. Although searchers have no “fixed time limit” after which they can no longer search, there is likely to be a global pressure to keep security lines moving efficiently, thus making these effects of time pressure particularly relevant. It has been suggested that screener performance is not linked to salary (Filipczak, 1996; Guzzo, Jette, & Katzell, 1985), but we did observe in the present study that reward-based pressure to perform accurately and quickly leads to increased SOS. The potentially significant factor of time pressure should therefore be an important consideration in the training of airport security screeners and in constructing the environment in which searches are conducted (i.e., obscuring screener awareness of passenger line length). Future SOS research that more closely parallels the search conditions of airport security screening might reveal optimal strategies for efficient staffing and operations to reduce both error and cost.

Conclusion

Our primary goal in this article has been to interrogate the phenomenon of SOS in a series of controlled laboratory studies to simultaneously inform radiology, cognitive psychology, and nonmedical searches such as airport baggage screening. One of the most important implications from this work is that we reveal the first evidence for SOS in a nonmedical search. By demonstrating that SOS can arise for nonexperts and in standard cognitive psychology search paradigms, we illustrate some of the factors critical to studying and understanding errors in multitarget search. SOS is a complex, generalized effect, and the only way to reduce its impact is to carefully delineate its diverse variety of underlying causes.

Acknowledgments

We thank Melissa Bulkin and Jordan Axt for help with data collection, and we thank the Duke Visual Cognition Lab for helpful conversations. We are very grateful to Ron Bowditch for extensive discussion of related issues in cytology searches. We also thank Kevin Berbaum for sharing a prepublication draft of his review chapter in press. This research was partially supported by National Institutes of Health/National Institute of Mental Health (NIH/

NIMH) grants F31 MH082582 (to Mathias S. Fleck) and R03 MH080849 (to Stephen R. Mitroff), Army Research Office Grant 54528LS (to Stephen R. Mitroff), and a grant from the Institute of Homeland Security Solutions (to Stephen R. Mitroff).

References

Ashman CJ, Yu JS, Wolfman D. Satisfaction of search in osteoradiology. American Journal of Roentgenology. 2000; 177:252–253.

Berbaum KS, El-Khoury GY, Franken EA Jr, Kuehn DM, Meis DM, Dorfman DD, et al. Missed fractures resulting from satisfaction of search effect. Emergency Radiology. 1994; 1:242–249.

Berbaum, KS.; Franken, EA., Jr; Caldwell, RT.; Schartz, KM. Satisfaction of search in traditional radiographic imaging. In: Samei, EA.; Krupinski, E., editors. The handbook of medical image perception and techniques. Cambridge, England: Cambridge University Press; 2009.

Berbaum KS, Franken EA Jr, Dorfman DD, Caldwell RT, Krupinski EA. Role of faulty decision making in the satisfaction of search effect in chest radiography. Academic Radiology. 2000; 7:1098–1106. [PubMed: 11131054]

Berbaum KS, Franken EA Jr, Dorfman DD, Miller EM, Caldwell RT, Kuehn DM, et al. Role of faulty visual search in the satisfaction of search effect in chest radiography. Academic Radiology. 1998; 5:9–19. [PubMed: 9442202]

Berbaum KS, Franken EA Jr, Dorfman DD, Miller EM, Krupinski EA, Kreinbring K, et al. The cause of satisfaction of search effects in contrast studies of the abdomen. Academic Radiology. 1996; 3:815–826. [PubMed: 8923900]

Berbaum KS, Franken EA Jr, Dorfman DD, Rooholamini SA, Kathol MH, Barloon TJ, et al. Satisfaction of search in diagnostic radiology. Investigative Radiology. 1990; 25:133–140. [PubMed: 2312249]

Bowditch R. Patterns found in false negative cervical cytology. Cytoletter. 1996; 3:22–25.

Brainard DH. The psychophysics toolbox. Spatial Vision. 1997; 10:443–446. [PubMed: 9176954]

Carter RC. Visual search with color. Journal of Experimental Psychology: Human Perception and Performance. 1982; 8:127–136. [PubMed: 6460078]

Chun MM, Wolfe JM. Just say no: How are visual search trials terminated when there is no target present? Cognitive Psychology. 1996; 30:39–78. [PubMed: 8635311]

DeMay, RM. The art and science of cytopathology. Chicago, IL: ASCP Press; 1996. p. 144-145.

DeMay RM. Common problems in Papanicolaou smear interpretation. Archives of Pathology and Laboratory Medicine. 1997; 121:229–238. [PubMed: 9111106]

Filipczak B. Can’t buy me love. Training. 1996; 33:29–35.

Fiore SM, Scielzo S, Jentsch F. Stimulus competition during perceptual learning: Training and aptitude considerations in the X-ray security screening process. International Journal of Cognitive Technology. 2004; 9:34–39.

Fleck MS, Mitroff SR. Rare targets are rarely missed in correctable search. Psychological Science. 2007; 18:943–947. [PubMed: 17958706]

Franken EA, Berbaum KS, Lu CH, Kannam S, Dorfman DD, Warnock NG, et al. Satisfaction of search in the detection of plain-film abnormalities in abdominal contrast studies. Investigative Radiology. 1994; 29:403–409. [PubMed: 8034444]

Gale AG, Mugglestone MD, Purdy KJ, McClumpha A. Is airport baggage inspection just another medical image? Proceedings of the SPIE. 2000; 3981:184–192.

Gerritsen C, Frischen A, Blake A, Smilek D, Eastwood JD. Visual search is not blind to emotion. Perception & Psychophysics. 2008; 70:1047–1059. [PubMed: 18717390]

Gibson BS, Li L, Skow E, Brown K, Cooke L. Searching for one versus two identical targets: When visual search has a memory. Psychological Science. 2000; 11:324–327. [PubMed: 11273393]

Guzzo RA, Jette RD, Katzell RA. The effects of psychologically based intervention programs on worker productivity: A meta-analysis. Personnel Psychology. 1985; 38:275–291.

Hanson CH, Hanson RD. Finding the face in the crowd: The anger superiority effect. Journal of Personality and Social Psychology. 1988; 54:917–924. [PubMed: 3397866]

Horowitz TS, Wolfe JM. Visual search has no memory. Nature. 1998; 394:575–577. [PubMed: 9707117]

Körner C, Gilchrist ID. Memory processes in multiple-target visual search. Psychological Research. 2008; 72:99–105. [PubMed: 17021837]

McCarley JS, Carruth DW. Oculomotor scanning and target recognition in luggage X-ray screening. Scanning and Recognition. 2004; 9:26–29.

McCarley JS, Kramer AF, Wickens CD, Vidoni ED, Boot WR. Visual skills in airport-security screening. Psychological Sciences. 2004; 15:302–306.

Menneer T, Barrett DJK, Phillips L, Donnelly N, Cave KR. Costs in searching for two targets: Dividing search across target types could improve airport security screening. Applied Cognitive Psychology. 2007; 21:915–932.

Menneer T, Cave KR, Donnelly N. The cost of search for multiple targets: The effects of practice and target similarity. Journal of Experimental Psychology: Applied. 2009; 15:125–139. [PubMed: 19586251]

Metlay W, Sokoloff M, Kaplan IT. Visual search for multiple targets. Journal of Experimental Psychology. 1970; 85:148–150. [PubMed: 5458320]

Nagy AL, Thomas G. Distractor heterogeneity, attention, and color in visual search tasks. Vision Research. 2003; 43:1541–1552. [PubMed: 12782068]

Palmer J, Verghese P, Pavel M. The psychophysics of visual search. Vision Research. 2000; 40:1227–1268. [PubMed: 10788638]

Posner, MI.; Cohen, Y. Components of visual orienting. In: Bouma, H.; Bouwhuis, D., editors. Attention and performance X. Hillsdale, NJ: Erlbaum; 1984. p. 531-556.

Rich AN, Kunar MA, Van Wert MJ, Hidalgo-Sotelo B, Horowitz TS, Wolfe JM. Why do we miss rare targets? Exploring the boundaries of the low prevalence effect. Journal of Vision. 2008; 8:1–17. [PubMed: 19146299]

Samuel S, Kundel HL, Nodine CF, Toto LC. Mechanisms of satisfaction of search: Eye position recordings in the reading of chest radiographs. Radiology. 1995; 194:895–902. [PubMed: 7862998]

Schwaninger A. Increasing efficiency in airport security screening. WIT Transactions on the Built Environment. 2005; 82:405–416.

Simon, H. Administrative behavior. 3rd ed.. New York, NY: Free Press; 1976.

Smith JD, Redford JS, Washburn DA, Taglialatela LA. Specific-token effects in screening tasks: Possible implications for aviation security. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2005; 31:1171–1185.

Takeda Y. Search for multiple targets: Evidence for memory-based control of attention. Psychonomic Bulletin & Review. 2004; 11:71–76. [PubMed: 15116989]

Transportation Security Administration. [Retrieved February 7, 2009] Prohibited items for travelers. 2009. from http://www.tsa.gov/travelers/airtravel/prohibited/permitted-prohibited-items.shtm

Treisman A, Gelade G. A feature-integration theory of attention. Cognitive Psychology. 1980; 12:97–136. [PubMed: 7351125]

Tuddenham WJ. Visual search, image organization, and reader error in roentgen diagnosis: Studies of the psycho-physiology of roentgen image perception. Radiology. 1962; 78:694–704. [PubMed: 13923013]

Van Wert MJ, Horowitz TS, Wolfe JM. Even in correctable search, some types of rare targets are frequently missed. Attention, Perception, & Psychophysics. 2009; 71:541–553.

Wang Q, Cavanagh P, Green M. Familiarity and pop-out in visual search. Perception and Psychophysics. 1994; 56:495–500. [PubMed: 7991347]

Wiegmann D, McCarley JS, Kramer AF, Wickens CD. Age and automation interact to influence performance of a simulated luggage screening task. Aviation, Space, and Environmental Medicine. 2006; 77:825–831.

Wolfe, JM. Visual search. In: Pashler, H., editor. Attention. London, England: UCL Press; 1998. p. 13-73.

Wolfe, JM. Guided search 4.0: Current progress with a model of visual search. In: Gray, W., editor. Integrated models of cognitive systems. New York, NY: Oxford; 2007. p. 99-119.

Wolfe JM, Horowitz TS, Kenner NM. Rare items often missed in visual searches. Nature. 2005; 435:439–440. [PubMed: 15917795]

Appendix A Response Time Data

Table A1

Response Time Data for Experiments 1–10

	Experimental parameters			Response time in seconds (with *SD*)			
Experiment	**Time limit**	**Prevalence high:low**	**Trial type counts**	**No target**	**High single**	**Low single**	**Dual target**
1	15 s	1:1	High: 50	9.02 (1.49)	9.43 (1.35)	9.51 (1.38)	8.52 (1.23)
			Low: 50				
			High-Low: 50				
			None: 100				
2	15 s	2:1	High: 100	9.65 (0.95)	9.42 (1.20)	9.58 (0.86)	8.10 (0.72)
			Low: 50				
			High-Low: 50				
			None: 50				
3	15 s	3:1	High: 120	9.93 (0.89)	9.53 (0.64)	10.01 (0.70)	8.55 (0.66)
			Low: 40				
			High-Low: 40				
			None: 50				
4	30 s	2:1	High: 100	10.45 (2.64)	9.39 (2.03)	9.83 (1.96)	8.05 (1.40)
			Low: 50				
			High-Low: 50				
			None: 50				
5	30 s	3:1	High: 120	10.97 (2.87)	10.23 (2.09)	10.72 (2.26)	8.99 (1.58)
			Low: 40				
			High-Low: 40				
			None: 50				
6	30 s	6:1	High: 150	11.05 (1.65)	10.38 (1.32)	10.88 (1.65)	8.83 (0.78)
			Low: 25				
			High-Low: 25				
			None: 50				
7	Luggage line (variable rate)	3:1	High: 96	10.15 (3.57)	9.46 (2.63)	10.04 (3.23)	7.02 (1.45)
			Low: 32				
			High-Low: 32				
			None: 40				
8	Luggage line (constant rate)	3:1	High: 96	11.77 (4.22)	10.95 (3.08)	11.12 (2.94)	8.65 (1.40)
			Low: 32				
			High-Low: 32				
			None: 40				
9	30 s	1:1 Low-single:dual	Low: 100	12.47 (3.36)	—	12.19 (3.13)	10.05 (1.73)

Experiment	Experimental parameters: Time limit	Prevalence high:low	Trial type counts	Response time in seconds (with *SD*): No target	High single	Low single	Dual target
10	30 s	4:1 Low-single:dual	Low-Low: 100	11.99 (2.06)	—	12.02 (1.35)	10.61 (0.83)
			None: 50				
			Low: 160				
			Low-Low: 40				
			None: 50				

Note. One mechanistic explanation for SOS is a "truncated search," in which participants prematurely end their search after finding a target (Samuel et al., 1995). In the current experiments, participants click on each target they find and then click a button labeled "DONE" once they have decided to terminate their search. Thus, in principle we can ask whether or not participants exhibit a truncated search by comparing the time taken to click "DONE" for high-salience single target trials in which the target was correctly detected and the no-target trials without false alarms. After successfully finding a high-salience target, are participants 'satisfied' and quicker to terminate their search? We are hesitant to make any strong conclusions from these data since they are complicated by the fact that one trial-type involves an extra mouse-click and that mouse-clicks in the fashion used here are not very sensitive. Only Experiment 4 (which did not reveal SOS) and Experiment 6 (which revealed a weak effect) had significantly faster responses for high-salience single target trials than no-target trials. Thus, these data reveal little evidence of a truncated search but likely reveal little reliable evidence in general. As a result, we focus on accuracy as our primary dependent measure.

Appendix B Accuracy Data for High-Salience Targets

Table B1

Accuracy Data for High-Salience Targets on Single- and Dual-Target Trials in Experiments 1–8 (Experiments 9 and 10 Did Not Include High-Salience Targets)

Experiment	Experimental parameters: Time limit	Prevalence high:low	Trial type counts	Accuracy (%): High single	High dual
1	15 s	1:1	High: 50	88.40 (8.42)	90.19 (7.63)
			Low: 50		
			High-Low: 50		
			None: 100		
2	15 s	2:1	High: 100	94.80 (5.57)	90.02 (9.72)
			Low: 50		
			High-Low: 50		
			None: 50		
3	15 s	3:1	High: 120	95.66 (3.42)	96.07 (3.83)
			Low: 40		
			High-Low: 40		
			None: 50		
4	30 s	2:1	High: 100	94.20 (4.10)	92.64 (6.55)
			Low: 50		
			High-Low: 50		
			None: 50		
5	30 s	3:1	High: 120	95.74 (2.35)	92.13 (8.25)
			Low: 40		

	Experimental parameters			Accuracy (%)	
Experiment	Time limit	Prevalence high:low	Trial type counts	High single	High dual
			High-Low: 40		
			None: 50		
6	30 s	6:1	High: 150	93.26 (5.77)	96.08 (5.48)
			Low: 25		
			High-Low: 25		
			None: 50		
7	Luggage line (variable rate)	3:1	High: 96	97.10 (4.60)	95.32 (6.12)
			Low: 32		
			High-Low: 32		
			None: 40		
8	Luggage line (constant rate)	3:1	High: 96	84.91 (7.81)	76.22 (16.15)
			Low: 32		
			High-Low: 32		
			None: 40		

Note. Experiments 1–8 included high-salience targets to calibrate the low-salience targets. However, since the dual-target trial counts were matched to the number of low-salience single-target trials, we did not analyze the two-target effects of salience, nor explore interactions between salience and other factors such as frequency or time limit. We provide the data here for completeness.

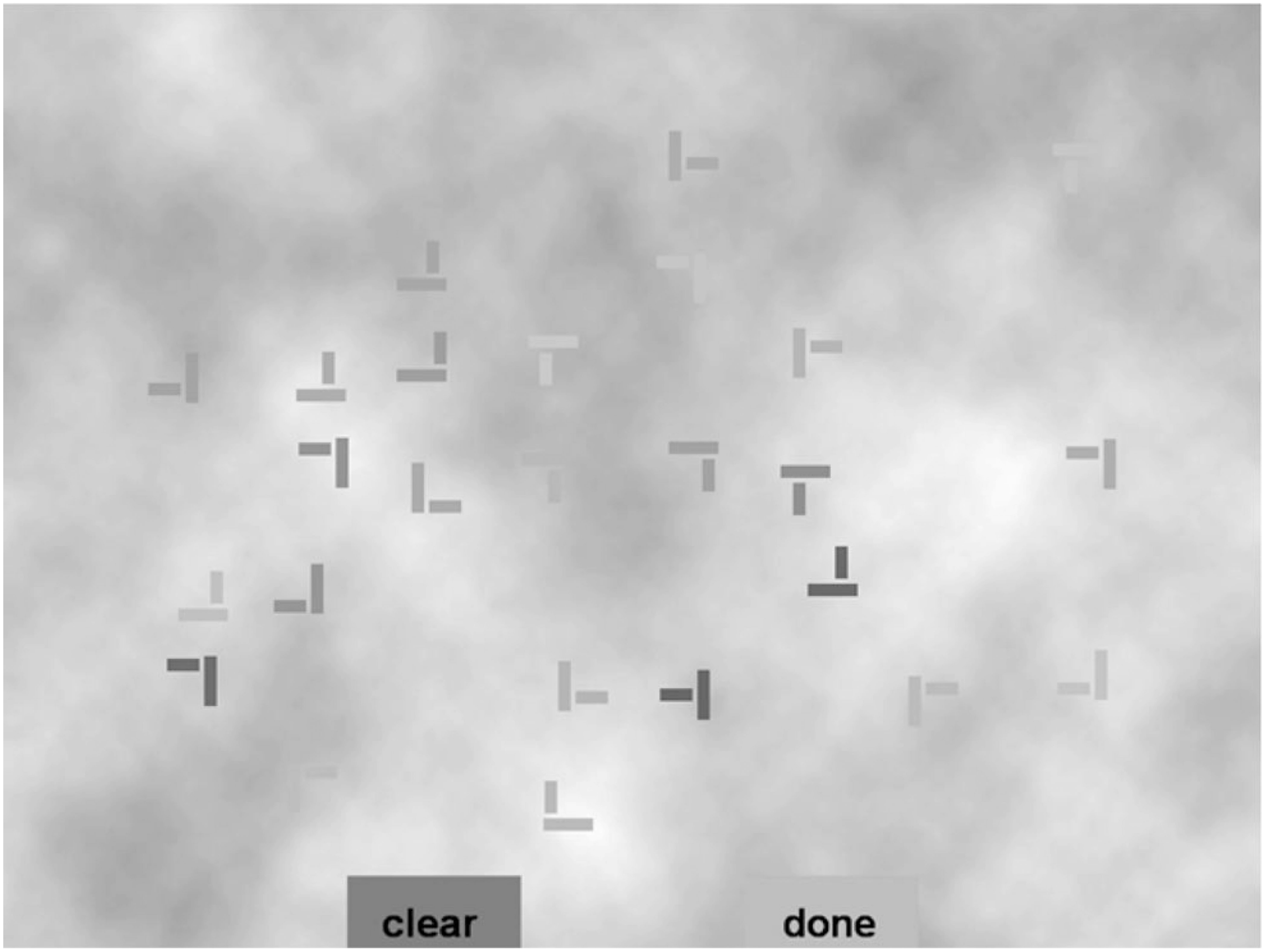

Figure 1.
Sample search display for Experiments 1–10. Each trial contained 0, 1, or 2 T-shaped targets among the L-shaped distractors. Experiments 7 and 8 additionally displayed a row of "luggage" icons across the top of the display.

Table 1

Experimental Parameters and Accuracy Data for Experiments 1–10

	Experimental parameters			Accuracy (%)			
Experiment	Time limit	Prevalence high:low	Trial type counts	Low single	Low dual	Low-single vs. Low-dual	SOS effect: SOS?
1	15 s	1:1	High: 50	60.40 (12.57)	63.09 (10.72)	$t(9) = 1.28$	✘
			Low: 50			$p = .23$	
			High-Low: 50			$d = -0.23$	
			None: 100				
2	15 s	2:1	High: 100	66.00 (16.22)	59.72 (16.62)	$t(9) = 3.18$	✔
			Low: 50			$p = .01$	
			High-Low: 50			$d = 0.38$	
			None: 50				
3	15 s	3:1	High: 120	72.50 (12.42)	60.76 (12.23)	$t(9) = 3.56$	✔
			Low: 40			$p < .01$	
			High-Low: 40			$d = 0.95$	
			None: 50				
4	30 s	2:1	High: 100	60.60 (17.26)	58.50 (19.59)	$t(9) = 1.07$	✘
			Low: 50			$p = .31$	
			High-Low: 50			$d = 0.11$	
			None: 50				
5	30 s	3:1	High: 120	60.50 (13.83)	55.64 (16.67)	$t(9) = 2.08$	~✔
			Low: 40			$p = .07$	
			High-Low: 40			$d = 0.32$	
			None: 50				
6	30 s	6:1	High: 150	68.00 (12.51)	57.90 (15.98)	$t(9) = 2.11$	~✔
			Low: 25			$p = .06$	
			High-Low: 25			$d = 0.70$	
			None: 50				
7	Luggage line (variable rate)	3:1	High: 96	74.07 (15.11)	73.08 (11.84)	$t(9) = 0.93$	✘
			Low: 32			$p = .38$	

NIH-PA Author Manuscript

NIH-PA Author Manuscript

	Experimental parameters			Accuracy (%)			
Experiment	Time limit	Prevalence high:low	Trial type counts	Low single	Low dual	Low-single vs. Low-dual	SOS effect: SOS?
			High-Low: 32			*d* = 0.07	
			None: 40				
8	Luggage line (constant rate)	3:1	High: 96	71.89 (12.77)	63.92 (15.11)	*t*(9) = 3.21	✔
			Low: 32			*p* =. 01	
			High-Low: 32			*d* = 0.57	
			None: 40				
9	30 s	1:1 Low-single:dual	Low: 100	78.17 (7.80)	76.75 (13.22)	*t*(9) = 1.12	✘
			Low-Low: 100			*p* = .29	
			None: 50			*d* = 0.13	
10	30 s	4:1 Low-single:dual	Low: 160	74.91 (8.70)	71.63 (5.43)	*t*(9) = 3.10	✔
			Low-Low: 40			*p* = .01	
			None: 50			*d* = 0.45	

Note. Accuracy reflects mean detection rates with standard deviations in parentheses; *p*-values represent 1-tailed *t*-tests of low-salience target accuracy in single vs. dual target trials.