



# Redundancy gain in visual search of simulated X-ray images

Claudia R. Hebert<sup>1</sup> · Li Z. Sha<sup>1</sup> · Roger W. Remington<sup>1,2</sup> · Yuhong V. Jiang<sup>1</sup>

Published online: 6 January 2020  
© The Psychonomic Society, Inc. 2020

## Abstract

Cancer diagnosis frequently relies on the interpretation of medical images such as chest X-rays and mammography. This process is error prone; misdiagnoses can reach a rate of 15% or higher. Of particular interest are false negatives—tumors that are present but missed. Previous research has identified several perceptual and attentional problems underlying inaccurate perception of these images. But how might these problems be reduced? The psychological literature has shown that presenting multiple, duplicate images can improve performance. Here we explored whether redundant image presentation can improve target detection in simulated X-ray images, by presenting four identical or similar images concurrently. Displays with redundant images, including duplicates of the same image, showed reduced false-negative rates, compared with displays with a single image. This effect held both when the target's prevalence rate was high and when it was low. Eye tracking showed that fixating on two or more images in the redundant condition speeded target detection and prolonged search, and that the latter effect was the key to reducing false negatives. The redundancy gain may result from both perceptual enhancement and an increase in the search quitting threshold.

**Keywords** Visual search · Selective attention · Eye movements · Visual attention

Many routine activities involve visual search, such as finding your car in a busy parking lot or searching for a product on a website. Visual search tasks can have far-reaching consequences in the real world. Professionals engage in visual search during important tasks such as screening for weapons in airport luggage or detecting abnormalities in medical images. Although models developed from simpler laboratory displays are useful (Drew, Evans, Vö, Jacobson, & Wolfe, 2013), applied tasks pose unique challenges. Of particular interest here are medical images, which are typically complex, involving the detection of subtle tumor signals not well segmented from the background. In some situations, such as routine cancer screening, the occurrence rates of tumors are low, creating a “low-prevalence” scenario that increases decision errors (Evans, Tambouret, Evered, Wilbur, & Wolfe, 2011). Image ambiguity, low target prevalence, inattention, and other factors contribute to errors in medical image perception, as

well as in laboratory studies using visual search (Wolfe, Evans, Drew, Aizenman, & Josephs, 2016). In routine cancer screening, the errors are largely false negatives, in which a tumor is present but undetected. Effective methods for decreasing the prevalence of false negatives would have a significant positive impact on cancer detection.

Several approaches have been used to reduce errors in medical image perception. Training of medical professionals involves conceptual development, such as learning key features of skin cancer through a checklist (Xu, Rourke, Robinson, & Tanaka, 2016). Perceptual learning also plays an important role, especially when the training regimen adapts to the observers' accuracy and speed (Kellman, 2013; Kellman & Garrigan, 2009). In some types of images, such as detecting bone fractures, novices may rapidly gain skills through perceptual learning (Chen, HolcDorf, McCusker, Gaillard, & Howe, 2017). Learning takes longer in most cases, however (Eng et al., 2000), with performance remaining imperfect even for radiology specialists. Another approach to reducing errors is with computer-assisted diagnostic (CAD) tools. These tools highlight potential abnormalities and can complement clinicians' initial reading. Although useful, CAD is associated with high false positive rates (Doi, 2007; Drew, Cunningham, & Wolfe, 2012). More recently, artificial intelligence, including deep-learning algorithms, has been successfully applied to cancer image diagnosis (Akselrod-Ballin et al., 2019; Topol,

---

Claudia R. Hebert and Li Z. Sha contributed equally to this work.

✉ Claudia R. Hebert  
heber145@umn.edu

<sup>1</sup> Department of Psychology, University of Minnesota, Minneapolis, MN, USA

<sup>2</sup> School of Psychology, University of Queensland, Brisbane, Australia

2019), although challenges remain (Bi et al., 2019; Hosny, Parmar, Quackenbush, Schwartz, & Aerts, 2018).

Search errors also affect other tasks, such as airport baggage screening. These tasks represent a low-prevalence scenario in which the target rarely appears. As the prevalence of targets is reduced, false-negative rates increase (Wolfe, Horowitz, & Kenner, 2005). At a prevalence rate of 1%, the target may be missed 30% of the time (Wolfe et al., 2005). A few methods have been employed to reduce false negatives in low-prevalence conditions. False-negative rates are reduced when participants have a second chance to change their response (Fleck & Mitroff, 2007). Increasing the number of response options from a two- to a four-alternative task also reduces false negative rates (Rich et al., 2008), suggesting that the search errors reflect, in part, decision errors that are correctable under some circumstances. Nonetheless, false negatives associated with rare targets are stubbornly resistant to many forms of correction, including reward (Wolfe et al., 2007).

The psychological literature on visual search has revealed that perception and search performance can be enhanced by adding redundant information. In visual search, response times (RTs) are faster when the display contains two targets rather than just one (Fischer & Miller, 2008). In addition, both short-term and long-term memory for objects are significantly better when participants view four duplicates of the same object rather than a single image at a time (Jiang, Kwon, Shim, & Won, 2010). Gender and expression judgments of faces are also significantly faster when participants are presented with multiple copies of the same face rather than a single face (Won & Jiang, 2013). Functional imaging studies have revealed higher activity in V1 and V2, as well as high-level category-selective brain regions, in reaction to the simultaneous presentation of four identical images, relative to either a single image or four different images (Shim, Jiang, & Kanwisher, 2013). Electroencephalograph recordings have shown differences as early as 90 ms between redundant and single-image conditions (Murray, Foxe, Higgins, Javitt, & Schroeder, 2001). In the present study, we explored the utility of *redundancy gain* to improve the detection of targets in simulated X-ray images similar to those used in the diagnosis of tumors.

The mechanism underlying the redundancy gain remains unclear, though several factors may contribute to it. First, presenting multiple targets can increase the effective signal-to-noise ratio. Because different neurons code different regions of visual space, images presented in multiple spatial regions will be processed in parallel by different populations of neurons. When targets or background differ, each of these neural populations will process images with a different efficiency, and the image processed most quickly may “win the race” (Morey, Thomas, & McCarley, 2018; Ulrich, Miller, & Schröter, 2007). This factor is particularly important if the multiple images are not identical, increasing the chance that

one image will yield faster or more accurate processing than the others. Performance is then dictated by the “best” image. In contrast, in the single-image condition, some trials will receive fast processing, and other trials will receive slower processing. Performance in this condition will reflect the average processing speed and accuracy across all images, rather than that of the fastest/most accurately processed image. Second, even when the images are identical, projecting redundant information to distinct neural populations may enhance perception, by drawing on more neural resources, especially those from both hemispheres (Miller, 2004). Essentially, decision processes can pool target information from multiple sources in parallel. This type of global image processing may increase the signal-to-noise ratio, enhancing the perception of target information (Alvarez & Oliva, 2009). Finally, redundant presentation may convey other advantages to visual search, such as a change in quitting threshold (Chun & Wolfe, 1996). With multiple images to search from, the observer may set a higher threshold for quitting, perhaps by inspecting multiple images, therefore lowering false-negative rates.

Here we tested whether presenting multiple copies of an image, or presenting multiple similar images concurrently, may reduce errors in the visual search of simulated X-ray images. Experiment 1 demonstrated that redundant images substantially reduced false-negative rates without an increase in false alarms. Experiment 2 extended the advantage to situations in which the target’s prevalence rate was low. In Experiment 3 we used eye tracking to characterize overt search behavior when searching multiple images. Together, these experiments provide a proof of concept that redundant presentation may hold potential in facilitating search for tumor-like stimuli. Analyses of RTs and eye fixations also shed light on the mechanisms of redundancy gain in visual search of complex images.

## Experiment 1

Medical images, such as mammography, contain subtle “targets” (e.g., tumors) hidden among background noise characterized by  $1/f^3$  noise (Burgess, Jacobson, & Judy, 2001). Following previous work (Drew et al., 2012; Sha, Remington, & Jiang, 2018), we simulated cancer image perception by asking participants to make a present/absent judgment of a low-contrast “T” target embedded among  $1/f^3$  noise. Although the “T” is not identical to a tumor, its presentation shares the important property of being difficult to separate from the background. Indeed, Sha et al. found significant differences in search procedure between this condition and one in which the search array was easily segmented from the background. The target was present on 50% of the trials, representing a high-prevalence situation.

In the single-image condition, the search display was presented in a randomly selected quadrant. In the multiple-image conditions, the search display contained four images, one per quadrant. If the target was present on one image, it was present on all four images. Conversely, if it was absent on one image, it was absent on all four images. Thus, the four images presented redundant information. The key question we addressed was whether redundant images facilitate search performance, compared with the single-image condition. Furthermore, to identify the generalizability of the redundancy gain across many types of redundant images, we included four types of redundant images that differed in whether the background noise was identical in the four quadrants, and whether the target's relative location in the quadrants was identical. The "same-background–same-location" condition presented four identical copies of an image. The "same-background–different-location" condition presented the same background noise in the four quadrants, but the target, when present, was in different relative locations of the quadrant. In the other two conditions, the background noise was not identical across the four quadrants. The inclusion of these conditions allowed us to test whether redundancy gain occurred primarily when the four images contained different information, or whether it also happened when the four images were identical. Figure 1 illustrates the single-image condition and one redundant-image condition.

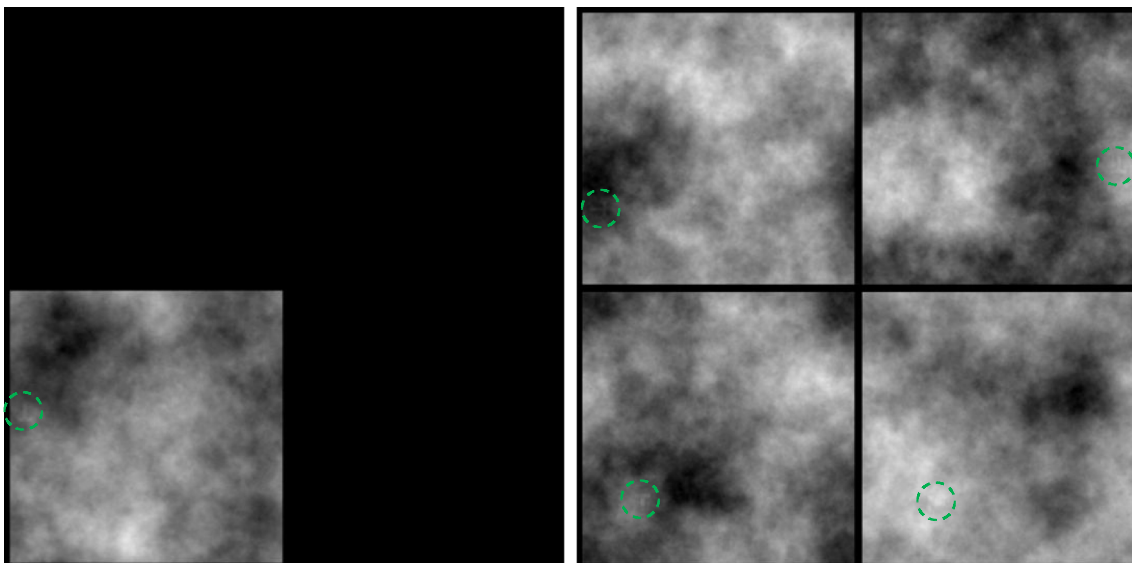
If redundancy gain arises exclusively from the "horse race" model, with the fastest image winning the race (Ulrich et al., 2007), then this effect should occur when the four images differ (i.e., when either the background noise differed or when the target's relative location differed). When the images differ, it is more likely that the target in at least one image will

happen to be in a salient spot relative to background noise and will yield above-average processing speed. In this case, the "winning" image will facilitate performance in the trial. Nonetheless, redundancy gain may also occur when an identical image is duplicated in the four quadrants, given that different populations of neurons process the images in parallel. In fact, mechanisms based on pooled neural resources (Miller, 2004), ensemble encoding (Alvarez & Oliva, 2009), or a change in search-quitting threshold (Chun & Wolfe, 1996) would predict redundancy gain in all of the multiple-image conditions. A comparison of the target-absent and target-present responses further informs us whether redundant presentation affects detection sensitivity or response criterion.

## Method

**Participants** College students between 18–25 years completed this study for extra course credit. In Experiment 1 we tested 11 females and five males (mean age 20.3 years). The sample size was predetermined on the basis of previous work examining redundancy gain (Jiang et al., 2010; Won & Jiang, 2013). The estimated power was greater than .90 with a sample size of 16.

**Equipment** Participants were tested individually in a room with normal interior lighting. The experiment was coded using Psychtoolbox (Kleiner, Brainard, & Pelli, 2007), implemented in MATLAB. The stimuli were projected on a 19-in. CRT monitor (spatial resolution  $1,024 \times 768$  pixels), with a vertical refresh rate of 75 Hz. The viewing distance was approximately 48 cm.



**Fig. 1** Example images from the study. (Left) The single-image condition. (Right) A redundant-image condition in which the target appears in different relative locations against different background

noise. The green dotted circles marking out the location of the target are shown here for illustrative purposes. They did not appear in the actual experiment.

**Stimuli** The search target was a letter T rotated 90° clockwise or counterclockwise, randomly determined on target-present trials, and subtended 0.7°×0.7°. The pixel value of the T was determined using the BlendFunction in Psychtoolbox, which combines the pixel values of the background noise and the target T using an opacity parameter. This function aimed to equate local intensity differences between the T and the background, regardless of the pixel value of the background. In our study, the opacity of the T was 0.92—a value determined on the basis of a previous study using similar stimuli (Sha et al., 2018). For example, if the background noise had a pixel value of 255 (white), the pixel value of the T would be 245; if the background noise had a pixel value of 0 (black), the pixel value of the T would be 10. The differences in pixel values between the T and the background were comparable, making the T difficult to see regardless of whether it was surrounded by darker or lighter noise.

When present, the target's location was randomly chosen from an invisible 10×10 matrix (12.1° × 12.1° per quadrant). In two of the redundant-image conditions, the target's relative location differed across the four quadrants. The background consisted of noise with a power spectrum of  $1/f^3$ . The noise was chosen to resemble the power spectrum of mammograms (Burgess et al., 2001). It changed from trial to trial, and in two of the redundant-image conditions it differed across the four quadrants.

**Procedure and design** Following eight trials of practice, participants completed ten blocks of 40 trials each. Participants clicked on a white fixation square in order to initiate each trial. The search display then appeared and remained until participants had made a target-present (“Y” key) or target-absent (“N” key) response. Participants were informed that when four images were presented, either the T was absent altogether or it would appear on all four images. Participants were instructed to respond as quickly and as accurately as possible. They were given a maximum of 10 s to make the response. Following the response, the display disappeared, and a beep indicated whether the response was correct (high beep) or incorrect (low buzz). The average accuracy of a block was displayed at the end of that block.

The experiment consisted of a 2 (target present/absent) × 5 (image condition) factorial design. The images contained a target 50% of the time. The images with the target could be presented in one of five formats: single image, four identical images (same background and same relative target locations), four images with the same background noise but different relative target locations, four images with different background noise but the same relative target locations, and four images with different background noise and different relative target locations. A blank space of 0.3° separated images from adjacent quadrants. Target-absent trials were matched with target-present trials, except that the target was removed from

all images. When four images were presented, a target was either absent from all four images or present in all of them. The 400 experimental trials were randomly and evenly divided into these ten trial conditions. Figure 1 shows example images.

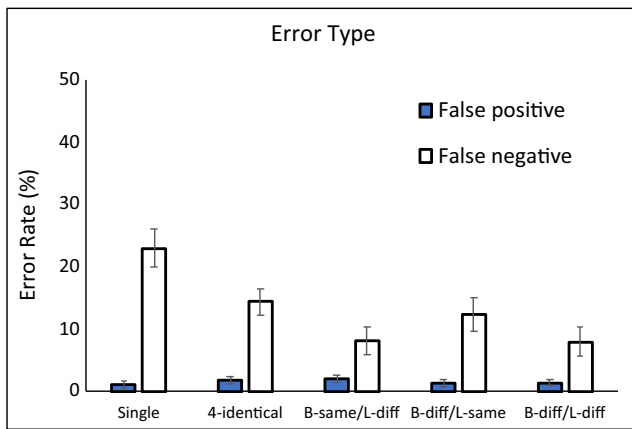
## Results

**Accuracy** Errors were of two types: False positives (false alarms) occurred when participants mistakenly reported a target when there was none, and false negatives (misses) occurred when participants missed a target. As is shown in Fig. 2, false positives were uniformly low across conditions, but false negatives were considerable and varied across conditions. In particular, the single-image condition was associated with a high false-negative rate, and this declined in the redundant-image conditions. These observations were confirmed in an analysis of variance (ANOVA) of error types and image conditions, showing significant results for the main effect of error type,  $F(1, 15) = 31.92, p < .001, \eta_p^2 = .68$ ; the main effect of image condition,  $F(4, 60) = 20.13, p < .001, \eta_p^2 = .57$ ; and their interaction,  $F(4, 60) = 14.80, p < .001, \eta_p^2 = .50$ . Follow-up tests showed no effects of image condition on false positives,  $F < 1$ , but significant effects on false negatives,  $F(4, 60) = 19.06, p < .001, \eta_p^2 = .56$ .<sup>1</sup>

The preceding analysis found a redundancy gain in all of the redundant-image conditions, even when the four redundant images were identical. Nonetheless, as is predicted by the horse race model, introducing variability across the redundant images further increased the gain. False negatives declined when the four images varied as compared with when they were identical. An ANOVA of background noise (same vs. different) and the target's relative location in a quadrant (same vs. different) showed a significant main effect of the target's location on false-negative rates,  $F(1, 15) = 12.35, p = .003, \eta_p^2 = .45$ . Background similarity did not produce statistically significant effects,  $F < 1$ , nor was there an interaction between target location and background similarity,  $F < 1$ .

When accuracy was analyzed using signal detection theory (Macmillan & Creelman, 2005), we found that image condition affected both detection sensitivity  $d'$ ,  $F(4, 60) = 8.67, p < .001, \eta_p^2 = .37$ , and response criterion  $c$ ,  $F(4, 60) = 8.91, p < .001, \eta_p^2 = .37$ . Because these differences were driven mainly by the false-negative rates, the statistical results for these two indices mirrored the pattern reported for false-negative rates and are omitted here. Relative to the redundant-image

<sup>1</sup> When comparing the single-image condition with each of the four redundant-image conditions, the paired  $t$  test statistics were: against four identical images,  $t(15) = 3.53, p = .003$ ; same background–different target locations,  $t(15) = 6.49, p < .001$ ; different backgrounds–same target location,  $t(15) = 5.13, p < .001$ ; different backgrounds–different target locations,  $t(15) = 8.22, p < .001$ .



**Fig. 2** Accuracy data from Experiment 1. Condition labels: *Single*, single-image condition; *4-identical*, four identical images with the same background noise and the same relative target location; *B-same/L-diff*, four images with the same background noise but different relative target locations; *B-diff/L-same*, four images with different background noise but the same relative target location; *B-diff/L-diff*, four images with different background noise and different relative target locations. Error bars show  $\pm 1$  SEs of the means. Some error bars may be too small to see.

conditions, the single-image condition was associated with a lower  $d'$  and a more conservative response criterion.

Did the redundancy gain in false negatives change across the ten blocks of experimental trials, as might be expected if participants adjusted their search strategy over time? Our data suggest that it did not. In this analysis, we combined data from all four redundant-image conditions and contrasted them with data from the single-image condition (Table 1). An ANOVA of block and condition showed that false negatives declined significantly across the ten blocks,  $F(9, 135) = 2.06, p = .037, \eta_p^2 = .12$ . In addition, the redundant conditions yielded lower false negatives than did the single-image condition,  $F(15) = 44.40, p < .001, \eta_p^2 = .75$ . These two factors did not interact,  $F(9, 135) = 1.14, p = .34$ , suggesting that redundancy gain was robust across the entire experimental session. Subsequent experiments also did not reveal an interaction between redundant conditions and block, and these data are summarized in Table 1.

**Response times** An analysis of RTs yielded insights into the source of the redundancy gain observed above. This analysis included correct trials only. As is shown in Fig. 3 (left), RTs

were longer on target-absent than on target-present trials,  $F(1, 15) = 91.60, p < .001, \eta_p^2 = .86$ , consistent with previous visual search findings (Wolfe, 1998). When the target was present, RTs were comparable across all conditions,  $F(4, 60) = 1.43, p > .23$ . However, when the target was absent, RTs differed across conditions,  $F(4, 60) = 40.79, p < .001, \eta_p^2 = .73$ : The single-image condition was faster than all the other conditions ( $ps < .001$  with Bonferroni correction), which did not differ from each other,  $F(3, 45) = 1.48, p = .23, \eta_p^2 = .09$ . The RT data on target-absent trials suggest that with redundant images, participants were less likely to give up search when they did not find the target.

If the redundancy gain came mainly from greater persistence in search, this should have led to an overall longer target-present RT on redundant than on single-image displays. Yet the mean RTs on target-present trials were comparable across conditions. However, the global average did not fully characterize the data, as the conditions differed in their proportions of fast and slow responses. Figure 3 (right) plots the RTs for single-image and redundant-image trials, with RTs averaged into ten percentile bins. For the fastest six percentiles of each condition (10th through 60th percentiles), the redundant conditions were associated with faster RTs than the single-image condition. The reverse was true for the slowest four percentiles (70th through 100th percentiles), with slower RTs in the redundant than in the single-image condition. This crossover pattern in RT distributions yielded a significant Condition  $\times$  Percentile interaction,  $F(9, 135) = 13.43, p < .001, \eta_p^2 = .47$ .

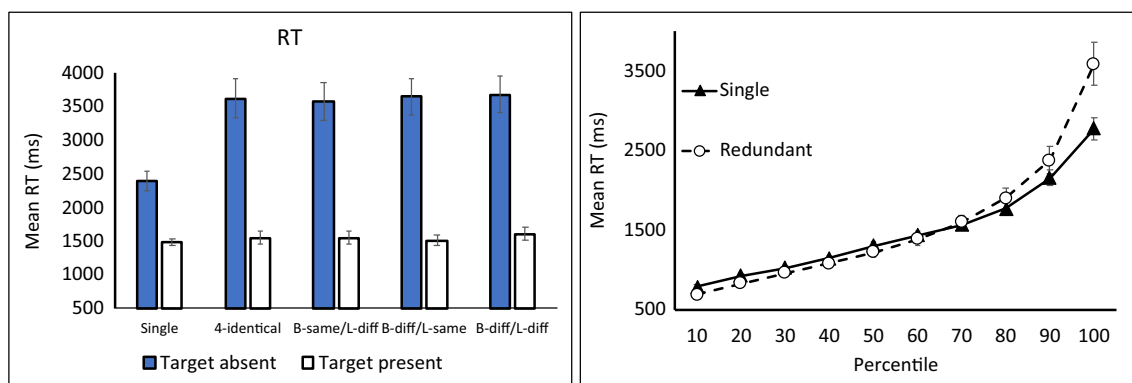
**Discussion**

In all conditions, false positives were uniformly low. However, Experiment 1 demonstrated a striking reduction in false negatives (misses) when four images were presented concurrently. Moreover, the benefits of multiple-image presentation were retained across all redundant display conditions. False negatives were lowest in the two conditions in which target locations differed, but the effect size was small compared with the overall difference from the single-image condition.

**Table 1** Mean false-negative rates across experimental blocks

Exp.	Cond.	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10
1	Single	.30	.34	.22	.25	.28	.16	.20	.17	.16	.22
	Redundant	.16	.10	.09	.12	.15	.07	.11	.11	.11	.05
2	Single	.31	.50	.50	.66	.47	.47	.34	.56		
	Redundant	.27	.31	.29	.24	.33	.39	.30	.39		
3	Single	.33	.30	.34	.28	.19	.27	.23	.30	.22	.19
	Redundant	.14	.12	.13	.09	.09	.14	.09	.10	.09	.08

The data were averaged across the four redundant conditions.



**Fig. 3** (Left) Mean response time (RT) data from Experiment 1. (Right) Target-present RT distribution in the redundant-image conditions and the single-image condition. RTs are binned into percentiles (e.g., the fastest

10% of each condition). Error bars show  $\pm 1$  SEs of the means. Some error bars may be too small to see.

These data suggest that redundant images changed performance in two ways. First, redundancy allowed people to more rapidly detect the target when it was present. The fastest 60% of RTs of each distribution showed slower RTs for single as compared with multiple images. This suggests that some form of ensemble encoding facilitates rapid detection of the target with multiple images. Second, when the target was not found initially, redundant displays altered the quitting threshold, such that participants searched for a longer time. This change in quitting threshold led to a net increase in RTs on target-absent trials and to an increase in the slowest percentile responses on target-present trials in the multiple as compared with the single-image conditions. Enhanced perception and prolonged search contributed to reduced misses in the redundant-image conditions.

Redundancy gain was robust even when the four images were identical, suggesting that the effect reflected, in part, a pooling of signals from different populations of neurons even as they coded the same stimuli. Introducing variability in the four images added an additional benefit, as was predicted by the horse race model. This effect was small and reached significance only when the T was in different relative locations. The lack of a consistent variability effect is likely due to our use of BlendFunction, which equated local contrast between the T and the background regardless of where the T was. This minimized the variability in target salience across the four images, reducing the difference between the “best” image and the average. The next two experiments will further assess the role of image variability in modulating redundancy gain.

## Experiment 2

The first experiment showed that redundant displays enhanced detection sensitivity mainly by reducing false negatives. This finding raises the possibility that redundant presentation may be an effective approach to addressing a stubborn problem in visual search: high miss rates for low-prevalence targets.

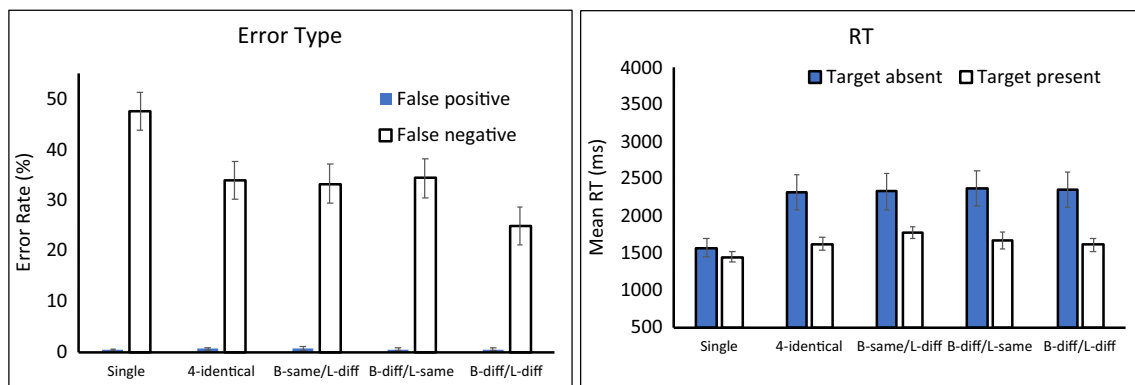
Targets that occur rarely, such as weapons in airport baggage or tumors in routine cancer screenings, are sometimes missed. In Evans et al. (2011), radiologists missed tumors about 22% of the time when the cancer prevalence rate in their stack of images was 50%. But when the prevalence rate declined to 2%–5%, the miss rate rose to 35%. The increase in the miss rate was accompanied by a reduction in false positives, suggesting that low target prevalence primarily affected decision criterion (Wolfe et al., 2007).

The low-prevalence effect is pernicious in situations in which false negatives are costly, such as when radiologists miss tumors or when a radar screener misses a dangerous signal. However, this effect is difficult to ameliorate experimentally. Task instructions or changes in the payoff matrix have not consistently reduced the low-prevalence effect (Rich et al., 2008; Wolfe et al., 2007). Experiment 1’s finding raises the possibility that redundant display presentation may be an effective method to reduce false negatives. However, that experiment was conducted with a high target prevalence of 50%. Therefore, in Experiment 2 we tested whether redundant presentation might also reduce false negatives under low-target-prevalence conditions. Consistent with previous studies, we selected a low target prevalence of 10% in Experiment 2 (Peltier & Becker, 2017; Wolfe et al., 2005).

## Method

**Participants** Sixteen new participants (nine females and seven males; mean age = 19.9 years) completed Experiment 2.

**Design and procedure** This experiment was identical to Experiment 1, except that the target prevalence rate was reduced to 10%. The experiment was extended to eight blocks of 100 trials each. The target was present 10% of the time and absent 90% of the time, and this manipulation was implemented for all five display conditions. Participants were tested using the same procedure and design as in Experiment 1, except that target-present trials occurred only 10% of the time.



**Fig. 4** Results from Experiment 2. (Left) Error types. (Right) Response times (RTs). Error bars show  $\pm 1$  SEs of the means. The condition labels are the same as in Fig. 2.

## Results

**Accuracy** When the target's prevalence rate was just 10%, false negatives reached nearly 50% in the single-image condition (Fig. 4). This doubled the rate of false negatives in Experiment 1. False negatives also rose in the other conditions. Importantly, the benefit of redundant image presentation was observed even when the target's prevalence was low.

As in Experiment 1, false-positive rates were uniformly low,  $F < 1$  for the main effect of image condition. False-negative rates varied across conditions,  $F(4, 60) = 10.12$ ,  $p < .001$ ,  $\eta_p^2 = .40$ . All four redundant display conditions produced lower false negatives than in the single-image condition.<sup>2</sup>

As in Experiment 1, redundancy gain was robust in all four redundant-display conditions, including the condition in which the four images were identical. Introducing variability to the redundant images had a modest effect. Owing to the small number of target-present trials in Experiment 2 (there were just 16 target-present trials per condition), the effects of image variability failed to reach statistical significance. An ANOVA of background condition (same vs. different) and the target's relative locations (same vs. different) did not reveal significant effects of the target's relative location,  $F(1, 15) = 3.92$ ,  $p = .066$ ,  $\eta_p^2 = .21$ ; of background similarity,  $F(1, 15) = 2.42$ ,  $p = .14$ ,  $\eta_p^2 = .14$ ; or of their interaction,  $F(1, 15) = 2.95$ ,  $p = .11$ ,  $\eta_p^2 = .16$ .

How did redundant presentation affect the low-prevalence effect? To address this question, we conducted a cross-experiment comparison on the false-negative rates, using prevalence rate as a between-subjects factor and image condition as a within-subjects factor. This analysis showed a higher false-negative rate in Experiment 2 than in Experiment 1,  $F(1,$

$30) = 33.74$ ,  $p < .001$ ,  $\eta_p^2 = .53$ , demonstrating the low-prevalence effect. The main effect of image condition was also significant, showing a redundancy gain,  $F(4, 120) = 23.01$ ,  $p < .001$ ,  $\eta_p^2 = .43$ . However, the two factors did not interact,  $F(4, 120) = 1.37$ ,  $p = .25$ . Thus, although redundant presentation was effective in reducing false negatives, it did not modulate the target's prevalence effect.

**Response times** Response time varied across image conditions and target-present/-absent status. When the target was present, RTs showed a small but significant variation across conditions,  $F(4, 60) = 3.19$ ,  $p < .019$ ,  $\eta_p^2 = .18$ , with the fastest RTs in the single-image condition. When the target was absent, RTs differed significantly across image conditions,  $F(4, 60) = 34.18$ ,  $p < .001$ ,  $\eta_p^2 = .70$ . The single-image condition was significantly faster than all the other conditions ( $ps < .001$  with Bonferroni correction), which did not differ from each other. Thus, as in Experiment 2, redundant displays were associated with more persistent search when the target was not found.<sup>3</sup>

RTs were comparable on target-present trials across the two experiments,  $F < 1$ , but target-absent RTs were considerably faster in Experiment 2 than in Experiment 1,  $F(1, 30) = 13.73$ ,  $p < .001$ , consistent with the idea that participants reduced their search times under low target-prevalence conditions.

## Discussion

Experiment 2 replicated and extended the redundancy gain observed in the first experiment, this time to a task that required finding a target that appeared infrequently. Participants now missed the target often, accompanied by faster RTs on target-absent trials. Importantly, image redundancy remained effective in reducing false negatives. Thus, the redundancy gain occurred under both high- and low-target-prevalence conditions.

<sup>2</sup> When comparing the single-image condition with each of the four redundant-image conditions, the paired  $t$  test statistics were: against four identical images,  $t(15) = 4.10$ ,  $p = .001$ ; same background–different target locations,  $t(15) = 3.49$ ,  $p < .003$ ; different backgrounds–same target location,  $t(15) = 3.89$ ,  $p < .001$ ; different backgrounds–different target locations,  $t(15) = 6.00$ ,  $p < .001$ .

<sup>3</sup> Owing to the small number of target-present trials, we could not perform the same percentile analysis as we had done in Experiment 1.

The lack of an interaction between target prevalence and image redundancy suggests that these factors make independent contributions to search performance. Whereas high target prevalence raises both target detection and false-alarm rates, redundant images raise target detection rates without also increasing false alarms. By providing additional signals, redundant image presentation increases detection sensitivity, in addition to altering response criterion.

Image variability did not significantly modulate the redundancy gain in Experiment 2. Owing to the low target prevalence, only a small number of target-present trials were observed in each condition, limiting the power of this analysis. Experiment 3 further explored the effect of image variability.

### Experiment 3

Experiment 1 had revealed evidence that the benefits of multiple-image trials may be due to prolonged search. In Experiment 3 we used eye tracking to gain further insight into the nature of overt search behavior among redundant images. We focused on the amount of time that participants fixated on the four images. One possibility is that participants spent most of the search time on one of the four images, relying on the other images only occasionally. If this were the case, then fixations should be predominantly spent on one of the four images, with a steep drop-off for the other images. A second possibility is that participants sample all images equally often, trading time per image for the number of images covered. To what degree did participants sample multiple rather than a single image to render their decisions?

### Method

**Participants** Sixteen new participants (15 females and one male, with a mean age of 19.1 years) completed Experiment 3. Five additional participants were tested, but the data were excluded due to computer problems.

**Eye tracking** An EyeLink 1000 eye tracker (SR Research Ltd., Mississauga, ON, Canada) tracked the left eye at a sampling rate of 2000 Hz. Participants rested their head on a chinrest throughout the experiment. The viewing distance was 75 cm. Eye position was calibrated before the experiment and verified prior to each trial, with recalibration done as needed. We focused on fixation patterns across the four quadrants. We also defined interest areas centered on the target, to examine fixation behavior. At a viewing distance of 75 cm, the T subtended  $0.67^\circ \times 0.67^\circ$ . The interest area was a square region around the target that subtended  $1.41^\circ \times 1.41^\circ$ .

**Design and procedure** This experiment was identical to Experiment 1, except for the addition of eye tracking and a change in viewing distance.

### Results

**Accuracy and response time** As is shown in Fig. 5, the accuracy and RT data replicated those from Experiment 1: Redundant images lowered false-negative rates,  $F(4, 60) = 18.82, p < .001, \eta_p^2 = .56$ , without altering false-positive rates,  $F < 1$ . On target-absent trials, RTs were significantly longer in the redundant-image conditions than in the single-image condition,  $F(4, 60) = 20.03, p < .001, \eta_p^2 = .57$ . As in Experiment 1, although mean RTs were comparable between the single- and redundant-image conditions on target-present trials,  $F(4, 60) = 1.03, p = .40$ , RT distributions differed between the single and redundant conditions. The redundant conditions were associated with faster RTs than in the single-image condition for the fastest 80 percentiles, but with slower RTs for the slowest 20 percentiles. The interaction between condition and percentile was significant,  $F(9, 135) = 2.80, p < .005, \eta_p^2 = .16$ . Owing to the similarity of the results to those from Experiment 1, detailed statistical reports are omitted.

As in Experiment 1, a redundancy gain was observed even when all four images were identical, but image variability further increased the effect. The exact patterns of statistical results, however, differed across the two experiments. In Experiment 3, an ANOVA of background similarity and target location revealed a significant main effect of background similarity on false negatives,  $F(1, 15) = 17.00, p < .001, \eta_p^2 = .53$ . The main effect of the target's relative location did not reach significance,  $F(1, 15) = 1.20, p = .29, \eta_p^2 = .07$ , and the two factors did not interact,  $F < 1$ . Thus, although in both experiments false negatives were numerically lower when either the background or the target's location varied, the exact patterns of statistical results differed. As we will discuss later, this difference may be attributed to the randomization of background and target location, introducing stimulus noise to otherwise identical experiments.

**Eye movement data** Similar to visual search tasks involving well-segmented items (Harris & Remington, 2017), in our study the number of saccades was strongly correlated with the RT, Pearson's  $r = .94$ . As in the RTs, on target-absent trials the number of saccades was greater in the redundant-image conditions ( $M = 9.0, SEM = 1.0$ ) than the single-image condition ( $M = 7.0, SEM = 0.7$ ),  $t(15) = 4.95, p < .001$ . On target-present trials, the mean numbers of saccades made before finding the target did not differ between the single-image ( $M = 4.8, SEM = 0.3$ ) and redundant-image ( $M = 4.6, SEM = 0.3$ ) conditions,  $t(15) = 0.95, p > .35$ . The distribution of saccades followed RT data to show a crossover pattern. Specifically, when trials were divided into ten bins from the fastest to the



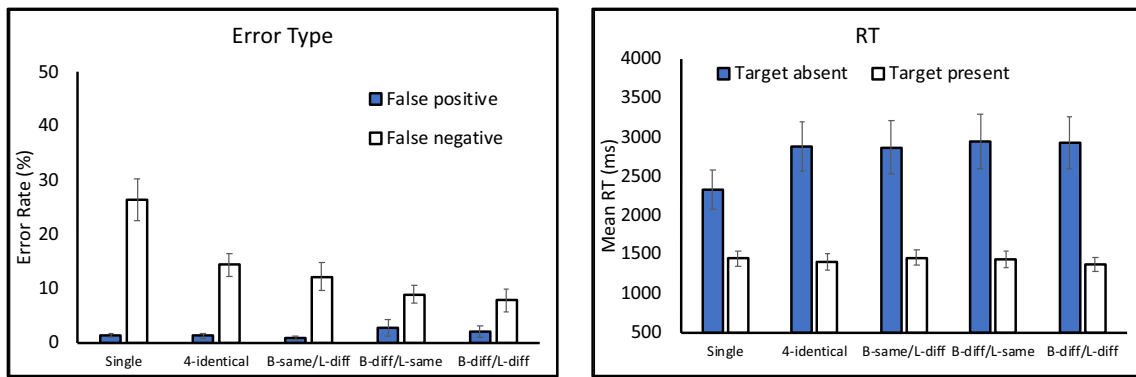


Fig. 5 Accuracy and RTs from Experiment 3. Error bars show  $\pm 1$  SEs of the means. The condition labels are the same as in Fig. 2.

slowest, the fastest 40th percentile of the redundant condition was associated with fewer saccades than the fastest 40th percentile of the single-image condition, and this pattern reversed for the slowest 10th percentile.

Whereas saccades were restricted to the single image in the single-image condition, they were distributed across space in the redundant-image conditions. Figure 6, left panel, shows the mean numbers of images fixated. On redundant-image displays, the average number of images fixated was 2.2 when the target was present, and 3.7 when the target was absent. That is, when the target was in the display, participants typically found it upon visiting two of the four images. When it was absent, participants visited nearly all four images. There were no statistical differences among the four redundant-image conditions ( $F < 1$ ).

The right panel of Fig. 6 shows how much time was spent examining each quadrant in the redundant-image conditions. The data from the four redundant-image conditions were combined, because they produced highly similar results. The duration of each fixation in each quadrant was summed to yield the total fixation time per quadrant, and then was rank-ordered. When the target was present, fixations were

disproportionately allocated to one, most visited quadrant. The mean fixation time on that quadrant was 824 ms, with the remaining 436-ms fixation time allocated to the other quadrants. The time spent on the most visited quadrant was shorter than the fixation time in the single-image condition (1,391 ms). Thus, participants spent about 65% of the fixation time in the most visited quadrant, but they also searched other quadrants. Target-absent trials showed a more even distribution of fixation times. The most visited quadrant was fixated for a total of 956 ms, which was about 40% of the total fixation time.

We also examined which quadrant was the most visited and if it bore any relationship to the quadrant that participants searched first. At the level of individual participants, each participant showed some idiosyncratic preference for some quadrants over others. For example, across all trials, one participant fixated primarily on the lower right (about 50% of the time). However, across all participants there was no systematic preference,  $F < 1$ . The proportions of time spent fixating on the upper left, upper right, lower left, and lower right quadrants were 25%, 24%, 24%, and 26%, respectively. Similarly, across all participants the first fixation was equally likely to

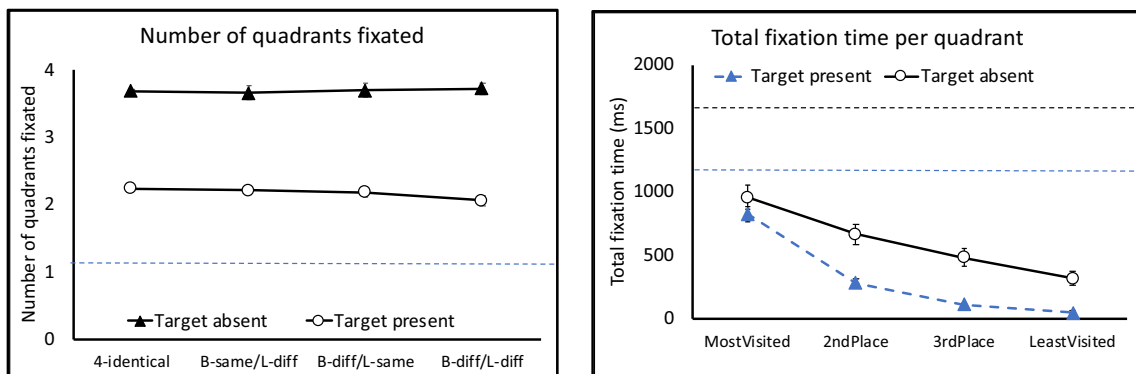


Fig. 6 Eye movement data from Experiment 3. (Left) Numbers of quadrants fixated in the redundant-image conditions. The dotted line shows the data from the single-image condition. The condition labels are the same as in Fig. 2. (Right) Total fixation times per quadrant, ranked from the most visited to the least visited quadrant. The dotted lines show data from the single-image condition on target-absent (upper dotted line) and target-

present (lower dotted line) trials. Condition labels: *MostVisited*, the quadrant that participants fixated for the longest time; *2ndPlace*, the quadrant fixated for the second longest; *3rdPlace*, the quadrant that participants fixated for the third longest; *LeastVisited*, the quadrant fixated for the shortest time. Error bars show  $\pm 1$  SEs of the means. Some error bars may be too small to see

land on any of the four quadrants (29%, 24%, 25%, and 22% for the upper left, upper right, lower left, and lower right quadrants, respectively),  $F < 1$ . There was a significant correspondence, however, between the first fixated quadrant and the most visited quadrant. On target-present trials, the first-fixated quadrant was also the most visited quadrant 66% of the time, which was significantly higher than chance (25%,  $p < .001$ ). On target-absent trials, the frequency that the first-fixated quadrant was also the most visited quadrant declined to 44%, though this was also higher than chance ( $p < .001$ ). Thus, although there was no systematic preference for a specific visual quadrant, participants spent much of their search time on the quadrant they happened to search first.

Participants differed in the degree to which they relied primarily on one dominant quadrant. On target-present trials, the proportion of time that participants spent fixating the most visited quadrant varied from 52% to 81% across individuals. This variability was not associated with a reduction in false-negative rates in individual participants, Pearson's  $r = -.16$ . This might in part be attributed to the fact that all participants favored one quadrant, limiting individual variability. Future studies using larger samples will be needed to further explore this issue.

Previous studies using radiological images have examined the time it took participants to first fixate a target. These studies showed that the time to fixate was sensitive to expertise, with expert observers fixating the target more quickly than novices (Donovan & Litchfield, 2013; Manning, Ethell, Donovan, & Crawford, 2006). Did the presence of multiple images influence the time to first fixate a target? To address this question, we examined fixations that landed on the interest area surrounding a target and computed the latency before participants first fixated the target interest area.<sup>4</sup> Note that participants sometimes responded without ever fixating the target directly, so this analysis was based on trials in which a target interest area was fixated (54% of trials in the single-image condition, and 57% of trials in the redundant-image conditions). This index did not reveal significant differences: The mean times to first fixate a target were 909 ms in the single-image condition and 962 ms in the redundant-image conditions,  $t(15) = 1.09$ ,  $p = .29$ . The lack of a difference was consistent with the findings from RTs, since mean RTs were also comparable between these conditions. Binning the data into percentiles from the fastest to the slowest time to fixate revealed the same crossover pattern as in RTs. For example, in the fastest 10th percentile of the data, the time to first fixate the target was 282 ms in the single-image condition, which was slower than the 266 ms in the redundant-image conditions. This trend continued through the 40th percentile of the data. The slowest 20th percentile showed the reverse—slower times to first fixate in the redundant (2,105 ms) than in

the single-image (1,937 ms) conditions. Thus, the time to first fixate a target showed results comparable to those for RTs.

## Discussion

Experiment 3 showed that when confronted with redundant images, participants inspected about two images before making a target-present decision, and nearly all four images before making a target-absent decision. Fixation times were distributed across images. On target-present trials, about 65% of the fixation time went to one most visited image. This proportion dropped to 40% on target-absent trials.

The four types of redundant images produced similar fixation patterns, and largely similar accuracy and RTs, as well. Nonetheless, variability in the four redundant images further increased the redundancy gain. The two factors that we examined, background noise and target location, should both introduce random variability in target salience across the four images. Because the stimuli were generated randomly, we did not expect uniform or additive effects between the background and target location factors. Across all three experiments, variability in background or target location yielded numerically similar results—target detection was numerically better if the background varied or if the target's location varied. This pattern was anticipated by the “horse race” model. When the four images differed somewhat (in either background or target location), the target would be displayed against a different noise patch. Even though the local intensity differences were comparable across regions, differences in the background's texture, density, and uniformity would introduce variability in the target's detectability across the four quadrants. Assuming that the “best” image drives performance, both background difference and target location difference were expected to enhance detection. Neither effect was particularly strong, however, given that local intensity differences were equated. In addition, the background was generated randomly, and so was the target's location. Hence, the degree of benefit from the different images varied randomly across the stimuli, even in experiments that were otherwise identical.

## General discussion

Medical images, such as chest X-rays and mammography, are important for cancer diagnosis. However, perception of these images is plagued by unacceptable false negatives, especially when targets are rare (low prevalence). Attempts to improve false-negative rates often yield high false-alarm rates. Here we provided a proof of concept that a simple change in image presentation can significantly lower false-negative rates on simulated X-ray images without raising false-positive rates. When presented with multiple identical or similar images, all of which yielded the same target-present/-absent response,

<sup>4</sup> We thank D. Litchfield for suggesting this analysis.

participants reduced their false-negative rates by about half, relative to viewing a single image. Redundant displays did not increase false-positive rates, yielding a net improvement in detection sensitivity. Eye tracking showed that participants sampled information from two (on target-present trials) to four (on target-absent trials) of the images.

Redundancy gain affected false negatives both when the target's prevalence rate was high (50%) and when it was low (10%), suggesting that this gain is applicable to low-prevalence situations such as routine cancer screening. This effect did not interact with the low-prevalence effect, suggesting that redundancy and target prevalence may influence performance independently. Thus, although redundant displays are not a remedy for the low-prevalence effect, they can be used as a general approach to reducing false negatives.

Although the mean RTs for detecting the target were comparable between single-image and redundant-image presentations, the distributions of RTs differed. In Experiment 1, the fastest 60% of the trials were faster in the redundant than in the single-image condition, suggesting that redundant presentation sped up target detection on a majority of the trials. At the same time, the slowest 40% of the trials were slower in the redundant than in the single-image condition. This cross-over pattern was observed again in Experiment 3. The fact that the fastest trials were faster in the redundant condition suggests that redundant displays did not just influence search strategy (e.g., by prolonging search). Instead, participants could more quickly discern target signals. This finding suggests that the redundancy gain reflects, in part, enhanced perception. In addition, participants also demonstrated greater persistence in searching among the redundant displays. In fact, the main cost of redundant presentation is a lengthening of search times on target-absent trials. On these trials, instead of quitting search after approximately 2.5 s, search persisted for another second or so.

The increase in search time may be an important practical factor to consider, given the high workload that radiologists already face. Nonetheless, our study showed that the gain in accuracy was substantial, amounting to a reduction of 15%–25% in false-negative rates. In addition, the majority of target-present trials received faster RTs in the redundant than in the single-image condition. Overall, the large reduction in false negatives may counter the cost of slower responses, especially in situations in which accurate diagnosis, rather than time on task, is more essential.

The redundancy gain was substantial even when the four images were identical. This finding reaffirms previous findings in visual search, face perception, and visual short-term and long-term memory tests (Fischer & Miller, 2008; Jiang et al., 2010; Miller, 2004; Shim et al., 2013; Won & Jiang, 2013). In addition, the present study involved visual search, a component that may have led to the large redundancy gain observed here. It's likely that in the present paradigm, the

redundant presentation both enhanced perception and raised the quitting threshold for target-absent responses.

In our study, we only examined situations in which the four images were truly redundant—they all yielded the same target-present/absent responses. The results would likely differ if the responses conflicted among the four images. This could be the case, for example, if a subset of the images contained a target but the others did not. In that scenario, presenting multiple images would amount to an increase in the search set size, likely reducing performance compared with single-image presentations. In fact, when multiple images or regions of the image are not truly redundant, search is characterized by another type of error—premature termination of search upon finding one target. This effect, known as “satisfaction of search” (Berbaum et al., 1994) or “subsequent-search misses” (Adamo, Cain, & Mitroff, 2013), manifests as a twofold reduction in finding additional abnormalities after the first abnormality has been identified. How redundancy gain interacts with satisfaction of search remains to be tested in the future.

Our study is also limited in its examination of similar images and the presentation of these images concurrently. Future research will be needed to extend these findings to other situations, such as when the redundant images are completely different (e.g., brain images from different slice prescriptions) or when the images are presented sequentially. Other researchers have found that toggling between similar, but slightly misaligned mammograms could enhance detection (Drew et al., 2015).<sup>5</sup> Future studies should examine whether the redundancy gain observed here reflects the same mechanism as the advantage offered by toggling through similar images.

Our study also differs from radiological search in our use of a well-defined target. The target T was the same shape across trials and in the four quadrants. This enabled observers to adopt a fixed search template. Radiological search, on the other hand, often involves less well-defined targets, such as the general category of tumor. The mechanisms that contributed to the redundancy gain observed in the present study, such as the pooling of neural resources across visual fields and a gain from the horse race model, should still apply when searching for an ill-defined target. However, search strategies may change, and decisions may become more complex when the target lacks a clear template. Future studies should extend the present finding to search tasks that involve less well-defined targets, such as search for a category of objects (Hout & Goldinger, 2015; Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004) or a highly diverse set (e.g., “incidental findings” in medical image perception). The use of real cancer images and the testing of medical experts are also important directions to take in the future.

<sup>5</sup> We thank A. Carrigan for raising this point.

## Summary and conclusion

Using a visual search task with simulated X-ray images, this study showed that displaying multiple identical or similar images yields significantly lower false-negative rates. The redundancy gain may reflect a combination of enhanced perception, an alteration in search procedure, and a change in the threshold for when to quit search. Redundant presentation may be a simple but effective method to reduce false-negative rates in practical tasks, such as routine cancer screening or airport baggage screening.

**Acknowledgments** This study was supported in part by a seed grant from OFAA–Social Sciences at the University of Minnesota. We thank Y. N. Toh for help with data collection.

**Open Practices Statement** A privacy clause in the consent form precludes us from making the raw data available in a data repository. Deidentified data in an aggregated format are available upon reasonable request from the authors.

## 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>
- Akselrod-Ballin, A., Chorev, M., Shoshan, Y., Spiro, A., Hazan, A., Melamed, R., ... Guindy, M. (2019). Predicting breast cancer by applying deep learning to linked health records and mammograms. *Radiology*, *292*, 331–342. <https://doi.org/10.1148/radiol.2019182622>
- Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences*, *106*, 7345–7350. <https://doi.org/10.1073/pnas.0808981106>
- Berbaum, K. S., El-Khoury, G. Y., Franken, E. A., Kuehn, D. M., Meis, D. M., Dorfman, D. D., ... Kathol, M. H. (1994). Missed fractures resulting from satisfaction of search effect. *Emergency Radiology*, *1*, 242–249. <https://doi.org/10.1007/BF02614935>
- Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., Mehrtash, A., ... Aerts, H. J. W. L. (2019). Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA: A Cancer Journal for Clinicians*, *69*, 127–157. <https://doi.org/10.3322/caac.21552>
- Burgess, A. E., Jacobson, F. L., & Judy, P. F. (2001). Human observer detection experiments with mammograms and power-law noise. *Medical Physics*, *28*, 419–437. <https://doi.org/10.1118/1.1355308>
- Chen, W., HolcDorf, D., McCusker, M. W., Gaillard, F., & Howe, P. D. L. (2017). Perceptual training to improve hip fracture identification in conventional radiographs. *PLoS ONE*, *12*, e0189192. <https://doi.org/10.1371/journal.pone.0189192>
- 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. <https://doi.org/10.1006/cogp.1996.0002>
- Doi, K. (2007). Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, *31*, 198–211. <https://doi.org/10.1016/j.compmedimag.2007.02.002>
- Donovan, T., & Litchfield, D. (2013). Looking for cancer: Expertise related differences in searching and decision making. *Applied Cognitive Psychology*, *27*, 43–49. <https://doi.org/10.1002/acp.2869>
- Drew, T., Aizenman, A. M., Thompson, M. B., Kovacs, M. D., Trambert, M., Reicher, M. A., & Wolfe, J. M. (2015). Image toggling saves time in mammography. *Journal of Medical Imaging*, *3*, 011003. <https://doi.org/10.1117/1.JMI.3.1.011003>
- Drew, T., Cunningham, C., & Wolfe, J. M. (2012). When and why might a computer-aided detection (CAD) system interfere with visual search? An eye-tracking study. *Academic Radiology*, *19*, 1260–1267. <https://doi.org/10.1016/j.acra.2012.05.013>
- Drew, T., Evans, K., Vö, M. L.-H., Jacobson, F. L., & Wolfe, J. M. (2013). Informatics in radiology: What can you see in a single glance and how might this guide visual search in medical images? *RadioGraphics*, *33*, 263–274. <https://doi.org/10.1148/rgr.331125023>
- Eng, J., Mysko, W. K., Weller, G. E. R., Renard, R., Gitlin, J. N., Bluemke, D. A., ... Scott, W. W. (2000). Interpretation of emergency department radiographs. *American Journal of Roentgenology*, *175*, 1233–1238. <https://doi.org/10.2214/ajr.175.5.1751233>
- Evans, K. K., Tambouret, R. H., Evered, A., Wilbur, D. C., & Wolfe, J. M. (2011). Prevalence of abnormalities influences cytologists' error rates in screening for cervical cancer. *Archives of Pathology and Laboratory Medicine*, *135*, 1557–1560. <https://doi.org/10.5858/arpa.2010-0739-OA>
- Fischer, R., & Miller, J. (2008). Differential redundancy gain in onset detection versus offset detection. *Perception & Psychophysics*, *70*, 431–436. <https://doi.org/10.3758/PP.70.3.431>
- Fleck, M. S., & Mitroff, S. R. (2007). Rare targets are rarely missed in correctable search. *Psychological Science*, *18*, 943–947. <https://doi.org/10.1111/j.1467-9280.2007.02006.x>
- Harris, A. M., & Remington, R. W. (2017). Contextual cueing improves attentional guidance, even when guidance is supposedly optimal. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 926–940. <https://doi.org/10.1037/xhp0000394>
- Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. L. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer*, *18*, 500–510. <https://doi.org/10.1038/s41568-018-0016-5>
- Hout, M. C., & Goldinger, S. D. (2015). Target templates: The precision of mental representations affects attentional guidance and decision-making in visual search. *Attention, Perception, & Psychophysics*, *77*, 129–149. <https://doi.org/10.3758/s13414-014-0764-6>
- Jiang, Y. V., Kwon, M., Shim, W. M., & Won, B.-Y. (2010). Redundancy effects in the perception and memory of visual objects. *Visual Cognition*, *18*, 1233–1252. <https://doi.org/10.1080/13506281003791074>
- Kellman, P. J. (2013). Adaptive and perceptual learning technologies in medical education and training. *Military Medicine*, *178*(Suppl. 10), 98–106. <https://doi.org/10.7205/MILMED-D-13-00218>
- Kellman, P. J., & Garrigan, P. (2009). Perceptual learning and human expertise. *Physics of Life Reviews*, *6*, 53–84. <https://doi.org/10.1016/j.plrev.2008.12.001>
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(ECVP Abstract Suppl.), 14. <https://journals.sagepub.com/doi/pdf/10.1177/03010066070360S101>
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
- Manning, D., Ethell, S., Donovan, T., & Crawford, T. (2006). How do radiologists do it? The influence of experience and training on searching for chest nodules. *Radiography*, *12*, 134–142. <https://doi.org/10.1016/j.radi.2005.02.003>
- Miller, J. (2004). Exaggerated redundancy gain in the split brain: A hemispheric coactivation account. *Cognitive Psychology*, *49*, 118–154. <https://doi.org/10.1016/j.cogpsych.2003.12.003>
- Morey, S. A., Thomas, N. A., & McCarley, J. S. (2018). Redundant-target processing is robust against changes to task load. *Cognitive*

- Research: Principles and Implications*, 3, 4. <https://doi.org/10.1186/s41235-017-0088-x>
- Murray, M. M., Foxe, J. J., Higgins, B. A., Javitt, D. C., & Schroeder, C. E. (2001). Visuo-spatial neural response interactions in early cortical processing during a simple reaction time task: A high-density electrical mapping study. *Neuropsychologia*, 39, 828–844. [https://doi.org/10.1016/S0028-3932\(01\)00004-5](https://doi.org/10.1016/S0028-3932(01)00004-5)
- Peltier, C., & Becker, M. W. (2017). Eye movement feedback fails to improve visual search performance. *Cognitive Research: Principles and Implications*, 2, 47. <https://doi.org/10.1186/s41235-017-0083-2>
- Rich, A. N., Kunar, M. A., Van Wert, M. J., Hidalgo-Sotelo, B., Horowitz, T. S., & Wolfe, J. M. (2008). Why do we miss rare targets? Exploring the boundaries of the low prevalence effect. *Journal of Vision*, 8(15), 15.1–17. <https://doi.org/10.1167/8.15.15>
- Sha, L. Z., Remington, R. W., & Jiang, Y. V. (2018). Statistical learning of anomalous regions in complex faux X-ray images does not transfer between detection and discrimination. *Cognitive Research: Principles and Implications*, 3, 48. <https://doi.org/10.1186/s41235-018-0144-1>
- Shim, W. M., Jiang, Y. V., & Kanwisher, N. (2013). Redundancy gains in retinotopic cortex. *Journal of Neurophysiology*, 110, 2227–2235. <https://doi.org/10.1152/jn.00175.2013>
- Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25, 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
- Ulrich, R., Miller, J., & Schröter, H. (2007). Testing the race model inequality: An algorithm and computer programs. *Behavior Research Methods*, 39, 291–302. <https://doi.org/10.3758/BF03193160>
- Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, 9, 33–39. <https://doi.org/10.1111/1467-9280.00006>
- Wolfe, J. M., Evans, K. K., Drew, T., Aizenman, A., & Josephs, E. (2016). How do radiologists use the human search engine? *Radiation Protection Dosimetry*, 169, 24–31. <https://doi.org/10.1093/rpd/ncv501>
- Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Cognitive psychology: Rare items often missed in visual searches. *Nature*, 435, 439–440. <https://doi.org/10.1038/435439a>
- Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*, 44, 1411–1426. <https://doi.org/10.1016/j.visres.2003.11.024>
- 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>
- Won, B.-Y., & Jiang, Y. V. (2013). Redundancy effects in the processing of emotional faces. *Vision Research*, 78, 6–13. <https://doi.org/10.1016/j.visres.2012.11.013>
- Xu, B., Rourke, L., Robinson, J. K., & Tanaka, J. W. (2016). Training melanoma detection in photographs using the perceptual expertise training approach. *Applied Cognitive Psychology*, 30, 750–756. <https://doi.org/10.1002/acp.3250>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.