

# Visual scan adaptation during repeated visual search

Christopher W. Myers

Air Force Research Laboratory, Dayton, OH, USA



Wayne D. Gray

Cognitive Science Department, Rensselaer Polytechnic Institute, Troy, NY, USA



There is no consensus as to how to characterize eye fixations during visual search. On the one hand, J. M. Wolfe, G. A. Alvarez, and T. S. Horowitz (2000) have described them as a haphazard sequence of fixations. On the other hand is research that shows systematic repetition of visual patterns when freely viewing a scene (T. Foulsham & G. Underwood, 2008; D. Noton & L. W. Stark, 1971a). Two experiments are reported that demonstrate the repetition and adaptation of visual scans during visual search, supporting an adaptive scanning hypothesis. When trials were repeated in a simple search task, visual scan similarity and search efficiency increased. These increments in similarity and efficiency demonstrate the systematic and adaptive nature of visual scans to the characteristics of the visual environment during search.

Keywords: visual scan, scanpath, adaptation, visual search, scan similarity, normalized similarity index

Citation: Myers, C. W., & Gray, W. D. (2010). Visual scan adaptation during repeated visual search. *Journal of Vision*, 10(8):4, 1–14, <http://www.journalofvision.org/content/10/8/4>, doi:10.1167/10.8.4.

## Introduction

Visual search is often considered an unsystematic, or “anarchic,” process that requires eliminating possible targets by scanning the environment (Wolfe, Alvarez, & Horowitz, 2000). However, what behavior should we expect when the same, or similar, scene is repeatedly searched? Is the sequence of fixations that leads to a target repeated, or is a new random sequence of fixations generated each time? With experience, do people somehow apprehend the entire scene in one glance and learn to make one fixation to the target location? In this paper, we present two experiments whose results imply that sequences of fixations, or *visual scans*, are acquired and adapted during repeated visual search.

Previous research has demonstrated that the search process becomes more efficient with experience, requiring less time and fewer behaviors to locate the target (Chun & Jiang, 1998; Reder, Weber, Shang, & Vanyukov, 2003; Wolfe, 1998b); however, no research has demonstrated that steps in the search process (i.e., fixations) are repeated with increased task experience. Further, no research has been reported that demonstrates a gradual refinement, or *adaptation*, of visual scans.

Since Yarbus’ (1967) seminal monograph, visual scans have been the focus of much research, some of which focused on their repetition (Chernyak & Stark, 2001; Foulsham & Underwood, 2008; Josephson & Holmes, 2002; Laeng & Teodorescu, 2002; Noton & Stark, 1971a, 1971b; Stark & Ellis, 1981; Stark et al., 1980; Zangemeister, Sherman, & Stark, 1995). Visual scan repetition has been studied under multiple contexts, and

repetition has been hypothesized to result from top–down, goal-driven processes (Noton & Stark, 1971a, 1971b) and become automatic with increased experience of the visual scene (Furst, 1971). Laeng and Teodorescu (2002) demonstrated that visual scans were stored in memory and repeated during image recognition. Additionally, if visual scans were entirely due to environmental factors, then it might be expected that visual scans produced with visual saliency models would be as similar to human scans as human scans were to other human scans. However, this is not the case (Foulsham & Underwood, 2008) and suggests that there are other processes involved in scanning that are not captured by visual saliency models.

Visual scan research and research on gains in search efficiencies with repeated distractor–target configurations both require environmental stability. For stable distractor–target configurations, Chun and Jiang (1998) have demonstrated search time savings from repeating distractor–target configurations, and Song and Jiang (2005) demonstrated that as little as 25% of the search environment (i.e., two distractor locations and the target location out of 12 total locations per trial) repeated from trial to trial resulted in search time savings. Similarly, visual scan repetition has been demonstrated using an “encode and recognize” paradigm where participants encode a stimulus and then re-encode, or recognize, the same stimulus some time later (Foulsham & Underwood, 2008; Laeng & Teodorescu, 2002). The common factor of environmental stability provides an intersection between visual scan repetition research and research using a visual search paradigm with repeating search displays. There has been no research to determine if visual scans are adapted with increased experience to repeating displays. The goals of the research

# Report Documentation Page

Form Approved  
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE <b>MAY 2009</b>		2. REPORT TYPE		3. DATES COVERED <b>00-00-2009 to 00-00-2009</b>	
4. TITLE AND SUBTITLE <b>Visual scan adaptation during repeated visual search</b>				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Air Force Research Laboratory, 2698 G Street Bldg. 190, Wright-Patterson AFB, OH, 45433</b>				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			

were to demonstrate that visual scans repeat during visual search that they adapt with regular exposure to repeating search displays.

## Visual scan adaptation

Previous research has demonstrated that visual scans repeat while freely viewing natural scenes. As with most human behavior, visual scans will repeat with some variability, and though visual scans from a repeating stimulus might not be identical they will be more similar than two random scans (Foulsham & Underwood, 2008; Noton & Stark, 1971b). Why sequences of saccades repeat is an unanswered question; however, one can speculate on possible learning mechanisms for relating one saccade/fixation to another. A step toward connecting the repeating scan phenomenon to learning mechanisms is to demonstrate that repeating scans adapt with increased task experience. When adapted visual scans result in efficiency gains, then the adaptation is considered beneficial and reflects a learned skill.

Adaptation is the process of becoming adjusted to new conditions. In order to adapt, there must be some degree of regularity in the new conditions. The adaptation of visual scans during visual search would result in a more efficient search process across repeated searches through the same stimulus. Efficiency gains would result from decreases in behaviors associated with locating the target, such as fixation durations and the number of fixations to find the target (Hidalgo-Sotelo, Oliva, & Torralba, 2005). Adaptations in behavior that lead to improved task efficiency while simultaneously maintaining a desired level of task efficacy is often associated with skill acquisition processes. Hence, visual scan adaptations that lead to more efficient searches while maintaining a high success rate would reflect an acquired visual search skill.

As a skill is acquired, it becomes automated and regularly used, which would result in similar visual scans. Card, Moran, and Newell (1983) claim that “What distinguishes cognitive skill from problem solving is the packaging of operator sequences into integrated methods...” for completing a goal (p. 367). The “packaging” of multiple saccades into a method for finding a target in a repeating search display would provide stability in the sequence of fixated items and the opportunity for reuse of the “packaged sequence.” The reuse of a packaged sequence would produce similar scans across multiple searches through the same repeating search display, thus providing a detectable hallmark of skill acquisition during visual search. Furthermore, such stability and reuse would facilitate information reduction (Haider & Frensch, 1999) in the form of fewer items fixated.

The information-reduction hypothesis of skill acquisition holds that people learn, with practice, to distinguish between task-relevant and task-irrelevant information, eventually limiting processing to task-relevant information

(Haider & Frensch, 1999). Improvements in task performance reflect increased knowledge about which information should be processed and which information can be avoided. The task-relevant knowledge is hypothesized to be available to consciousness and voluntarily used (Haider & Frensch, 1999), while task-irrelevant information is actively ignored at a perceptual, rather than conceptual, level of processing. Task relevancy has been shown to affect saccadic selectivity (Godijn & Kramer, 2006), demonstrating the relevance of the information-reduction hypothesis to visual search and saccades in general.

The processes associated with sequence packaging and information reduction would produce fixation sequences that could be regularly deployed during repeated visual search, and the number of fixations composing the sequences would be gradually reduced with increased search experience. The gradually repeated deployment of fixation sequences would produce an increase in visual scan similarity with increased task experience and provide evidence that integrated methods for visual search within repeating displays are adopted and adapted with experience, demonstrating an acquired visual search skill tailored to the environment.

We hypothesized that scans will repeat and adapt with increased task experience if processes associated with skill acquisition, such as automaticity (Furst, 1971) and memory storage and retrieval (Laeng & Teodorescu, 2002) affect scans during visual search. If visual scan repetition on the same display for a fixed location target is a skill, then this skill should express itself in two ways. First, the number of fixations composing visual scans will decrease with experience in a repeating visual search environment producing a gain in search efficiency. Second, scan pattern similarity will increase across repeated exposure to the same stimulus independent of the number of fixations composing the visual scans. Of course, if the first fixation is to a location that starts a trial and the last fixation is to the target, then it might be argued that the first and second hypothesis are not completely independent as some increase in scan similarity would be due to the reduction in the number of fixations. Hence, the third hypothesis is that when the number of fixations across repeated trials decreases, then the similarity of visual scans will increase at a faster rate than expected by chance (where chance is derived from experimental controls and Monte Carlo simulations). These three hypotheses are collectively referred to as the *adaptive scanning hypothesis*.

The adaptive scanning hypothesis is counter to two alternative hypotheses. The first alternative hypothesis is that visual scans result from an anarchic succession of saccades (Wolfe et al., 2000), and therefore scans cannot systematically change with experience. This hypothesis is henceforth referred to as the *anarchy hypothesis*. The second alternative hypothesis is the application of a general strategy that is consistently applied across all

searches (e.g., clockwise, center-out search). As this type of strategy is gradually acquired the similarity of visual scans should gradually increase. This hypothesis is henceforth referred to as the *general strategy hypothesis* and is equally applicable to deliberate (i.e., endogenous) strategies as well as to exogenously driven fixation sequences.

The anarchy and general strategy hypotheses are in opposition to the behavioral predictions associated with the adaptive scanning hypothesis. Where the adaptive scanning hypothesis predicts an increase in scan similarity and a reduction in the number of fixations to find the target, the anarchy hypothesis predicts no change in visual scan similarity or the number of fixations to find the target with increasing task experience. Further, the general strategy hypothesis predicts an increase in scan similarity with experience but no reduction in fixations to find the target. Consequently, all three of the predictions from the adaptive scanning hypothesis must not be rejected to demonstrate support that visual scans adapt during search.

Two experiments were conducted to test the adaptive scanning hypothesis. Both experiments used a visual search task that contained repeating and novel search displays.

## Methods

The same general paradigm and method for comparing the similarity of visual scans was used across both experiments. The paradigm and comparison methods are introduced, followed by each experiment.

### Paradigm

Participants searched for a target (T) among distractors (L). Target–distractor feature similarity was high to avoid salience effects, thereby making search serial and inefficient (Wolfe, 1998a). The target was the same on every trial (T) and differed only in its orientation and location. Targets could be oriented in either the 90° position where the T's top was on the right, or the 270° position where the T's top was on the left. The same orientation scheme also applied to distractors.

Each trial contained 12 stimuli—11 distractors and one target. Participants were instructed to locate the target as quickly as possible. Once the target was located, participants responded by pressing a button labeled “Right” when the target was in the 90° orientation or with a button labeled “Left” when the target was in the 270° orientation.

Participants completed multiple blocks of 24 trials. Each block was evenly divided between *repeating search displays* and *novel search displays*. Repeating displays repeated from block to block, whereas novel displays did not. To avoid a target location learning confound between

repeating and novel search displays, repeating and novel displays were crafted to control for frequency of target location. There were 24 possible target locations evenly divided between repeating and novel displays. For both display types, target locations repeated from block to block. Distractor locations in the repeating displays also repeated across blocks, but distractor locations from novel displays did not. Target and distractor locations were separated by a minimum of  $\approx 3^\circ$  of visual angle from other distractor and target locations at a viewing distance of  $\approx 56$  cm.

### Visual scan similarity

Visual scans were operationally defined as the sequence of fixated items that occur from the time a search display is presented until a participant responds to the target and included all fixations on target and distractor items. Where scans have been defined as a specific number of fixations for determining scan repetition (c.f., Foulsham & Underwood, 2008), the current operational definition provided the opportunity for the number of fixations to change with experience.

Arbitrary identifiers (i.e., letters) were assigned to display item locations. For each search display, distractor and target locations were assigned an identifier unique to the search display. Thus, repeating displays had identical letter identifiers for target and distractor locations. In novel displays, distractor identifiers were randomly assigned an identifier and the target was assigned “T.” The use of identifiers instead of display coordinates reduced the complexity of the algorithm from two comparisons for determining if two fixated items are at the same location (one for the  $x$  and one for the  $y$  display coordinates) to making one comparison (between letter identifiers). The similarity results from the novel displays served as a control for comparing results from repeating displays.

To determine the degree of similarity between two visual scans, the Levenshtein (1966) distance algorithm was used. The algorithm determines the minimum number of insertions, deletions, and replacements necessary to change one scan into another. To demonstrate how the Levenshtein algorithm works with letter identifiers, take the comparison between the two visual scans: FIREMEN and POLICEMEN, where each letter represents a fixated item, and repeated letters represent a second fixation on an item (i.e., *refixation*). To change “FIREMEN” into “POLICEMEN,” the solution would be (1) to insert a “P” to the left of the “F”; (2) to insert “O” to the left of the “F”; (3) to replace “F” with “L,” and (4) to replace “R” with “C,” resulting in minimum-edit distance of 4. Hence, similarity is based on the *sequence* of fixated items rather than the proportion of items fixated. For example, two scans that contained the same fixated items but in a

different sequence (i.e., ABCDT and DBCAT) would be as similar as two scans that contained no common fixated items other than the target (i.e., GHEJT and ABCDT).

A visual scan was defined as the first fixated item during a trial to the last fixated item at the trial response. Analyzed visual scans included fixations and refixations on distractors and targets. Target fixations were not excluded, as they too are part of the visual scan from a search trial, could be refixated during search, and there was no requirement to fixate the target before or during a response.

Minimum-edit distances are normalized to control for differences in lengths of compared visual scans (Foulsham & Underwood, 2008; Josephson & Holmes, 2002). The normalized minimum-edit distance is then subtracted from one to obtain the normalized similarity index, or *NSI*

$$NSI = 1 - \left( \frac{MED}{S_{\text{longest}}} \right), \quad (1)$$

where *MED* is the minimum-edit distance and  $S_{\text{longest}}$  is the length of the longer of the two compared scans.

There is an inverse relationship between the *NSI* metric and the length of visual scans, such that as scan lengths decrease *NSI* values increase. Monte Carlo simulations were performed to ensure that *NSI* values obtained from novel displays approximated the similarity from two random scans, providing a second control for increases in the *NSI* metric from repeating and novel displays. The Monte Carlo simulations produce a “special case” of the anarchy hypothesis. Where the anarchy hypothesis would not predict a reduction in the number of fixations to find the target, the simulations are based on human data that may demonstrate a reduction of fixations with experience. Consequently, the Monte Carlo simulation results can be considered a conservative prediction of visual scan similarity based on the anarchy hypothesis.

To obtain *NSI* values from random sequences for comparison to human *NSI* values from repeating search displays, sequences must be approximately the same length as the human visual scans. To obtain a random sequence of appropriate length, visual scan lengths from novel search displays for each epoch were first obtained from the human data. Second, two random samples of visual scan length,  $L_1$  and  $L_2$ , were sampled from the human data without replacement because the number of refixations in the human data was assumed to be very low.

After obtaining the two scan lengths, they were used in constructing the random sequences. The number of fixations for each length was reduced by one,  $L_x - 1$ , and items were randomly sampled without replacement to produce a random sequence. Once the random sequence was produced, a target identifier was added to the end of the sequence and was based on the simplifying assumption that participants fixated the target when they responded. This was done for  $L_1$  and  $L_2$ , where the first sequence,  $S_1$ , had a length of  $L_1$ , and the second sequence,  $S_2$ , had a length of  $L_2$ . These steps effectively produced

pseudorandom scans. The scans were not completely random because items were sampled without replacement. After producing the two pseudorandom sequences, their similarity was computed using the *NSI* metric.

## Experiment 1

**Experiment 1** used a 2 (search display type)  $\times$  20 (block) repeated measures design, with two within-participant levels of search display type, repeating and novel, crossed with 20 blocks. Participants were instructed to locate a target (T) among distractors (L) and respond to the target’s orientation as quickly and accurately as possible. There were three displays that constituted a trial: a *fixation control display*, a *search display*, and a *feedback display*. First, participants fixated crosshairs on a fixation control display for 600 ms. After 600 ms, a repeating or novel search display was presented. All repeating search displays were unique to individual participants. After responding to the target’s orientation, accuracy feedback was provided. After feedback, new fixation crosshairs were displayed signaling the onset of a new trial. Participants took a 10 second break after each block of 24 trials.

### Apparatus

The task environment was built in-house using ANSI common Lisp in the LispWorks development environment. The task environment ran on Apple Macintosh 10.4.4. All items were presented on a 17-inch flat-panel display at a resolution of 1280  $\times$  1024. Stimuli subtended  $\approx 2^\circ$  of visual angle at a viewing distance of  $\approx 56$  cm. All items were separated by a minimum of  $\approx 3^\circ$  of visual angle at a viewing distance of  $\approx 56$  cm from their centroids. Consequently, the minimum distance between the edges of stimuli was  $\approx 1^\circ$  of visual angle.

A binocular Eyegaze Analysis System, manufactured by LC Technologies, Inc., was used in conjunction with a chinrest to track participants’ point of gaze. The system had a 120-Hz sampling rate and was accurate to  $0.45^\circ$  of visual angle (0.38 cm) at a viewing distance of  $\approx 56$  cm. The system ran on Microsoft Windows XP and communicated with the task environment over a TCP/IP connection.

### Participants

A total of 40 students from Rensselaer Polytechnic Institute participated in the experiment. Participants in all conditions were eye-tracked and had normal or corrected-to-normal vision.

### Results

First, accuracy results are presented to demonstrate that participants are successfully completing the task and are

followed by search efficiency analyses. Finally, scan similarity results are presented.

*Trial accuracy.* A 2 (search display type)  $\times$  20 (block) repeated measures analysis of variance (ANOVA) was conducted to determine if there were systematic differences in response accuracy. The dependent variable was the proportion of correct trials, per block. Block and the Search Display Type  $\times$  Block interaction violated sphericity and the Greenhouse–Geisser correction was used.

The mean proportion of correct responses was high: 0.98. There was not a significant main effect of search display type,  $F(1, 37) < 0.001, p > 0.99, NS$ , indicating no difference in response accuracy between repeating and novel search displays. There was not a significant main effect of block,  $F(10.52, 389.29) = 1.21, p > 0.28, NS$ , nor did search display type interact with block,  $F(10.9, 403.19) = 0.69, p > 0.73, NS$ .

*Efficiency analyses.* To determine if participants' search efficiency increased, trial response times and the number of fixations and refixations to find the target were analyzed across blocks of the experiment. We hypothesized that the number of fixations would decrease with increased experience of repeating search trials. A failure to reject the hypothesis suggests that participants are increasing search efficiency, whereas a rejection of the hypothesis supports the general strategy hypothesis.

*Response times.* To determine if response times differed as a function of search display type or experience, a 2 (search display type)  $\times$  20 (block) repeated measures ANOVA was conducted. Block and the Search Display Type  $\times$  Block interaction violated sphericity and the Greenhouse–Geisser correction was used.

There was a significant main effect of block,  $F(7.31, 255.7) = 25.92, p < 0.001$ , where response times were gradually reduced from a mean of 1287.2 ms in block 1 to a mean of 943.6 ms by block 20. There was not a main effect of search display type,  $F(1, 35) = 0.12, p > 0.73, NS$ , nor was there a significant Search Display Type  $\times$  Block interaction,  $F(9.9, 346.61) = 0.57, p > 0.83, NS$ .

*Fixation count.* Fixations were determined using a sample-based fixation algorithm (see Myers & Schoelles, 2005, for a full description of the algorithm). Once a fixation was calculated, the closest display item within  $2^\circ$  of visual angle was assigned to the fixation. Concurrent fixations on the same display item were aggregated into a single fixation. If there was not a display item within  $2^\circ$  of visual angle, “middle of nowhere” was assigned to the fixation. Fixation sequences (i.e., visual scans) were determined for each trial.

To determine if the number of fixations to find the target was reduced across blocks, a 2 (search display type)  $\times$  20 (block) repeated measures ANOVA was performed on the mean number of fixations on stimulus items per block. Block and the Search Display Type  $\times$  Block interaction violated sphericity and the Greenhouse–Geisser correction was used. The average number of fixations was 2.86.

There was not a main effect of search display type,  $F(1, 35) = 0.485, p > 0.49, NS$ . Importantly, there was a main effect of block,  $F(4.39, 153.76) = 12.00, p < 0.001$ , demonstrating that the mean number of fixations to find the target was reduced with experience ( $M_{\text{Block-1}} = 3.38; M_{\text{Block-20}} = 2.38$ ). There was not a significant Search Display Type  $\times$  Block interaction,  $F(9.67, 348.9) = 0.99, p > 0.45, NS$ .

The average number of refixations was submitted to a 2 (search display type)  $\times$  20 (block) repeated measures ANOVA. There were very few refixations, less than 0.03 on average. However, there was a main effect of block,  $F(19, 342) = 2.8, p < 0.001$ , where the number of refixations was reduced across blocks ( $M_{\text{Block-1}} = 0.08; M_{\text{Block-20}} = 0.02$ ).

The results demonstrate that participants increased their search efficiency across blocks—accuracy was maintained at a high proportion of correct trials (0.98) while the amount of time to find the target decreased by 343.6 ms from the first to the last block.

*Scan similarity analyses.* The following analysis was conducted to determine if visual scans increase in similarity during visual search, reflecting an established and regularly used method for finding targets in repeating search displays. We hypothesized that visual scans would increase in similarity with increased experience of repeating search trials. A failure to reject the hypothesis suggests that participants are developing skillful search through repeating displays, whereas if visual scans do not increase in similarity, then there is support for the anarchy hypothesis.

Visual scans were aggregated into epochs of blocks, where one epoch equaled five blocks. For each participant, trials from each of the repeating search displays within each epoch (5 views of the same repeating search display) were compared against each other producing 10 NSI values for each repeating search display. The 10 NSIs were then averaged to obtain the mean epoch NSI for each of the 12 repeating search displays. Next, the mean NSI for each epoch was averaged across repeating displays to acquire the average repeating search display NSI for each epoch of trials.

In order to determine if the similarity of visual scans across repeating search displays is greater than predicted by chance, NSI values for scans from novel search displays were also determined. Novel displays were only compared with other novel displays that shared target location. This produced a chance NSI value with which to compare NSI values from repeating displays.

To determine if there were differences in NSI values between repeating and novel search displays across epochs, a 2 (search display type)  $\times$  4 (epoch) repeated measures ANOVA was performed on all mean NSI values. Epoch violated the sphericity assumption and corresponding results use the Greenhouse–Geisser correction.

There was a main effect of search display type,  $F(1, 35) = 106.08, p < 0.001$ , where visual scans from repeating

search displays ( $M_{\text{Repeating}} = 0.67$ ) were significantly more similar than novel ones ( $M_{\text{Novel}} = 0.60$ ). There was also a main effect of epoch,  $F(2.18, 69.79) = 36.13$ ,  $p < 0.001$ , demonstrating an increase in similarity across epochs. However, there was not a significant Search Display Type  $\times$  Epoch interaction,  $F(3, 99) = 0.78$ ,  $p > 0.49$ , *NS*.

The above analyses did not reveal evidence that visual scans from repeating displays increased in similarity at a faster rate than the similarity of scans from novel displays. Post hoc analyses using the Bonferroni correction revealed that there were significant differences in NSI values between repeating and novel search displays within the first epoch ( $p < 0.001$ ). Hence, it may be the case that differences in visual scans do increase faster for repeated than for novel displays, but that this increase occurs rapidly within the first epoch.

To investigate the likelihood of a rapid differential onset of similarity within the first six views of repeating displays, we derived five consecutive scan comparisons for repeated displays and for novel displays with the same target location (i.e.,  $S_1$  vs.  $S_2$ ,  $S_2$  vs.  $S_3$ ,  $S_3$  vs.  $S_4$ ,  $S_4$  vs.  $S_5$ ,  $S_5$  vs.  $S_6$ ). The NSI values from these comparisons were used in a 2 (search display type)  $\times$  5 (consecutive scan) repeated measures ANOVA. This analysis yielded a significant main effect of search display type,  $F(1, 35) = 4.69$ ,  $p = 0.037$ . However, it did not show a main effect of consecutive scan,  $F(3, 105.14) = 0.22$ ,  $p > 0.88$ , nor was there a reliable Search Display Type  $\times$  Consecutive Scan interaction,  $F(4, 140) = 0.20$ ,  $p > 0.93$ , *NS*.

**Monte Carlo simulations.** To produce pseudorandom NSI values across epochs, 10 NSI values were determined for the pseudorandom sequences, effectively reproducing the same number of NSI values from novel search displays in the experiment. This was repeated for each epoch. To mimic the reduction of human scan lengths across epochs of the experiment, scan lengths from the first epoch of Experiment 1 were only used for the first epoch of the pseudorandom sequence comparisons, scan lengths from the second epoch of Experiment 1 were only used for the second epoch of the random sequence comparisons, *et cetera*. As was the case for the human generated visual scans, the 10 NSI values were then averaged to obtain the mean epoch NSI for each of the four epochs. This process provided the same number of comparisons as that from a single human participant from the experiment. To minimize variability, the Monte Carlo simulation described above was run 200,000 times (10,000 times for each participant in the experiment, see Figure 1).

As mentioned earlier, the Monte Carlo simulation results provide a special case of the anarchy hypothesis because the number of fixations composing a scan was reduced across epochs of trials in the simulations. Consequently, the simulation results can be interpreted as a conservative estimate of scan similarities predicted by the anarchy hypothesis.

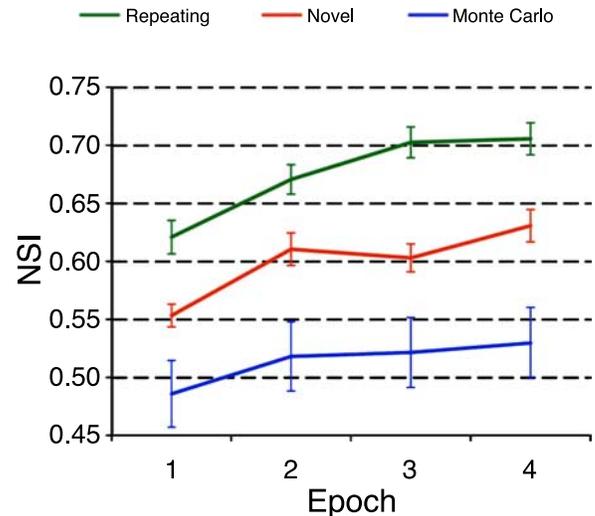


Figure 1. Normalized similarity index and Monte Carlo simulation results. Error bars from the human data represent 95% confidence intervals and represent standard error for the Monte-Carlo simulation data.

First, the results from the Monte Carlo simulations indicate that as the number of fixations to find the target are reduced with experience, NSIs gradually increase in similarity, much as visual scans from novel trials. Second, the results from the Monte Carlo simulations indicate that human visual scans are more systematic than predicted by the anarchy hypothesis, even scans from novel displays. A possible explanation is that participants may have learned the target locations associated with novel displays. An alternative explanation is that the participants are learning a more general search skill associated with non-stable search displays but that cannot be specialized as within stable search displays, as suggested by the general strategy hypothesis. Further, participants' search times decrease for novel displays (Chun & Jiang, 1998; Chun, 2000), demonstrating that participants are learning something general enough to apply to novel search displays. At the very least, NSI values obtained from the novel search displays highlights the ability of the human visual search process to adapt to the environment.

### Experiment 1 conclusion

Experiment 1 revealed that visual search becomes more efficient with experience of the search through the reduction of the fixations to find the target. This result rules out the general strategy hypothesis of scan repetition during repeated search. Furthermore, the experiment revealed that visual scans increased in similarity across searches of repeating displays, where the similarity cannot be completely attributed to chance. This result rules out the anarchy hypothesis of scan production during visual search. The results support two of the adaptive scanning

hypothesis predictions. Further, search efficiency and scan similarity increased with the small minimum distance between stimuli ( $1^\circ$  of visual angle), demonstrating that scan adaptation occurred when participants could encode multiple stimuli without gazing at a single stimulus. However, the hypothesis that visual scans from repeating displays would increase in similarity at a faster rate than chance was not supported through comparisons with the similarities between novel scans.

The consecutive scan analyses demonstrated that visual scans from repeating search displays are more similar than scans from novel displays after just two views, effectively replicating results reported in Foulsham and Underwood (2008). Foulsham and Underwood used the NSI metric to demonstrate that visual scans recorded during an encoding phase and a subsequent recognition phase were more similar than to randomly selected scans. They also demonstrated that salient areas of the presented image could not account for the visual scan repetition. The results from [Experiment 1](#) revealed similar findings to Foulsham and Underwood, demonstrating the generality of the repeating scan phenomenon as well as an extension to their results. The consecutive scan results from [Experiment 1](#) mimicked the results from Foulsham and Underwood, where the first two scans of a repeating display were more similar than two scans from novel displays. These results demonstrate the generality of repeating visual scans across visual search and encode and recognize tasks. The results from [Experiment 1](#) extend the results of Foulsham and Underwood by increasing the number of times repeating displays are viewed (or searched in the case of [Experiment 1](#)) and demonstrated an increase in scan pattern similarity with experience across both repeating and novel search displays.

There are two limitations to the results from [Experiment 1](#). First, there were too few trials in the experiment to determine if NSI values reached asymptote for visual scans. Second, the size of the target and distractors in [Experiment 1](#) were large ( $2^\circ$  of visual angle at a viewing distance of  $\approx 56$  cm) and a minimum of  $3^\circ$  separated the centroids of each item to ensure that all 12 stimuli fit on the display and did not overlap. Consequently, the edges of stimuli in the first experiment could be within  $1^\circ$  angle. This small minimum distance enabled participants to encode multiple stimuli without gazing at a single stimulus. This resulted in fewer fixations assigned to distractor and target items, reducing data for visual scan analyses. [Experiment 2](#) was designed to overcome these limitations and to uncover the roles of endogenous and exogenous influences on visual scan adaptation during search.

## Experiment 2

There were three goals for [Experiment 2](#). First was to replicate the results of [Experiment 1](#) in a search task with smaller stimuli. The second goal was to manipulate visual

scan similarity by adding a dual-task condition in an effort to demonstrate that concurrent cognitive activities disrupt visual scan adaptation. The third goal was to determine if different people similarly scan the same repeating displays.

[Experiment 2](#) differs from [Experiment 1](#) in four ways. First, the number of blocks in [Experiment 2](#) was extended from 20 to 30 (adding 240 trials). The extra trials provided 10 more views of each repeating search display compared to [Experiment 1](#) to help determine if scan pattern adaptations continue to occur past four epochs of trials.

Second, the size of distractors and targets was made smaller than those used in [Experiment 1](#). Reducing the size should make it more difficult for participants to encode multiple items with one fixation. If reducing the size increases difficulty in encoding multiple items, then search displays from [Experiment 2](#) will require more fixations, on average, to find the target compared to [Experiment 1](#).

Third, groups of participants searched through the same repeating search displays whereas in [Experiment 1](#) all repeating displays were unique to individual participants. Visual scans have been shown to be idiosyncratic (Noton & Stark, 1971a), and using the same repeating search displays across participants in [Experiment 2](#) will help to determine if different participants scan the same repeating search displays idiosyncratically. If different individuals' visual scans from repeating search displays are truly idiosyncratic, then NSI values computed by comparing visual scans from multiple participants for the same repeating displays should approximate NSI values from comparisons of visual scans from novel search displays. Consequently, participants were assigned to one of three different stimulus configuration groups, where participants within a configuration group searched through the same repeating search displays.

Fourth, cognitive load was added as a between-participant variable to dampen visual scan adaptation. The  $n$ -back task is a working memory task often used to induce increased cognitive load that produces predictable demands on prefrontal operations, such as cognitive control, planning, and goal maintenance. In the task, participants are presented with a series of items (letters or numbers) and have to decide if the current item matches the item that preceded it by  $n$  items in the presented sequence. Importantly, neuroimaging studies have confirmed that regional cerebral blood flow in a number of frontal brain regions is linearly related to the change in  $n$  (Jonides et al., 1997). Further, high cognitive load manipulated through the inclusion of the  $n$ -back task as a concurrent task negatively affected oculomotor control to abrupt onsets (Mitchell, Macrae, & Gilchrist, 2002). We used a task similar to the  $n$ -back in [Experiment 2](#) where participants classified a presented letter relative to a 1-back letter based on its occurrence within the English alphabet. If visual scan adaptation requires endogenous control, then participants performing visual search with a

concurrent task that disrupts cognitive processes hypothesized to be associated with scan adaptation (e.g., cognitive control, planning, and oculomotor control) will have less-similar visual scans than participants that only perform the visual search task.

**Experiment 2** used a 2 (load)  $\times$  3 (configuration group)  $\times$  [2 (search display type)  $\times$  30 (block)] mixed experimental design, where the factors outside the brackets were manipulated between participants and those within the brackets were manipulated within participants. The visual search task was identical to **Experiment 1**, except where noted.

Three different sets of repeating search displays were used across three groups of participants (configuration groups) to determine if scans are idiosyncratic. The same was done for target locations reserved for novel search displays. Consequently, within a configuration group, each participant searched for the same target locations.

Cognitive load was manipulated between participants using a multimodal dual-task paradigm. Participants in the single-task condition performed the visual search task. Participants in the dual-task condition performed the visual search task while simultaneously performing an auditory letter classification task.

The auditory letter classification task was a 1-back task with the additional task of deciding if the currently presented letter occurred before or after the previously presented letter in the English alphabet. The task consisted of randomly presented letters of the English alphabet with an intervening four-second interstimulus interval (ISI). Participants were to indicate whether the current letter ( $n$ ) preceded or followed the prior letter ( $n - 1$ ) in the alphabet. Participants responded by pressing one of two buttons on a response pad, labeled “BEFORE” and “AFTER.” Participants were instructed to respond quickly and accurately, and that if they failed to respond to the alphabet task within the four second ISI, then the response would be counted as incorrect. The task continued without interruption through all trials.

Participants were randomly assigned to a load condition followed by the random assignment of participants’ configuration group. Participants performed 720 visual search trials, broken into 30 blocks, where each block is a set of 24 trials. After completing the 15th block, all participants took a mandatory 5-minute break. After the break, participants completed the final 15 blocks.

### Apparatus

All stimuli in the letter classification task were presented via Apple’s speech software, using its “Victoria” voice. Responses were made using a Cedrus button box. Three letters were excluded (E, V, W) due to discrimination difficulty. The binocular eye tracking system used in **Experiment 1** was used in **Experiment 2**.

### Participants

A total of 26 students from Rensselaer Polytechnic Institute participated in the experiment. Participants in all conditions were eye-tracked and had normal or corrected-to-normal vision.

### Results

Results are presented in the same order as **Experiment 1**: accuracy analyses, efficiency analyses, and scan similarity analyses.

*Letter classification accuracy.* A 3 (configuration group)  $\times$  [2 (search display type)  $\times$  30 (block)] mixed ANOVA was conducted on participants’ accuracy of the auditory letter classification task in the dual-task condition. There were no significant effects, and accuracy was an adequate 78% correct, on average.

*Visual search accuracy.* A 2 (transfer-load)  $\times$  3 (configuration group)  $\times$  [2 (search display type)  $\times$  30 (block)] mixed ANOVA was conducted on trial accuracy. The dependent variable was the proportion of correct trials.

Accuracy for the search task was high, never falling below 91% correct across blocks of the experiment. There was a significant main effect of search display type,  $F(1, 22) = 9.344$ ,  $p = 0.006$ , where novel search displays resulted in a higher mean proportion of correct trial responses ( $M_{\text{Novel}} = 0.973$ ) than repeating displays ( $M_{\text{Repeating}} = 0.968$ ). This result demonstrates a difference in accuracy by search display type, where novel stimuli are responded to more accurately than repeating stimuli. Although there is a significant effect, the difference between the conditions (0.005) is trivial and has little bearing on the hypotheses being tested as the response accuracy for both search display types were very high and remained high throughout the experiment. There was also a Search Display Type  $\times$  Load  $\times$  Configuration Group interaction,  $F(2, 22) = 3.867$ ,  $p = 0.036$ , where one of the configuration groups led to reduced accuracy for repeating and novel search displays in the single-task condition but had little effect on repeating or novel displays in the dual-task condition. Finally, there was a Significant Block  $\times$  Load interaction,  $F(29, 638) = 2.58$ ,  $p < 0.001$ , where the dual-task group increased accuracy at a faster rate across blocks than the single-task group. These results demonstrate that participants were focused on finding the target and accurately responding and that the concurrent letter classification task negatively affected search response accuracy.

*Efficiency analyses.* To determine if participants’ search efficiency increased, trial response times and the number of fixations and refixations to find the target were analyzed across blocks of the experiment. Just as **Experiment 1**, we hypothesized that the number of fixations would decrease with increased experience of repeating search trials. A failure to reject the hypothesis suggests that participants

are increasing search efficiency, whereas a rejection of the hypothesis supports the general strategy hypothesis.

*Visual search response times.* A 2 (load)  $\times$  3 (configuration group)  $\times$  [2 (search display type)  $\times$  30 (block)] mixed ANOVA was conducted on response latency. There was a main effect of block,  $F(29, 609) = 12.84, p < 0.001$ , where response times decreased with experience. The single-task group had an average response time of 2197.13 ms in the first block and an average response time of 1657.43 ms in the 30th block. The dual-task group had an average response time of 4363.68 ms in the first block and an average response time of 1971.2 ms in the 30th block. There was also a main effect of load,  $F(1, 21) = 10.144, p = 0.004$ , where the dual-task group had longer response times than the single-task group. There was a significant Block  $\times$  Load interaction,  $F(29, 609) = 3.33, p < 0.001$ , where the dual-task group reduced trial response latencies across blocks at a faster rate than the single-task group. No other effects reached significance.

*Fixation count.* Experiment 2 used smaller stimulus items than those used in Experiment 1, and it was expected that this change would result in more fixations, on average, in Experiment 2 than in Experiment 1. The average number of fixations from Experiment 2 was 6.11 compared to 2.86 from Experiment 1.

If participants were increasing search efficiency, then the number of fixations to find a target should decrease across blocks of trials. A 2 (load)  $\times$  3 (configuration group)  $\times$  [2 (search display type)  $\times$  30 (block)] mixed ANOVA was performed on fixations on display items. There was a main effect of load, where the dual-task group had more fixations than the single-task group,  $F(1, 20) = 10.16, p = 0.005$ . There was a significant main effect of block, where the number of fixations decreased with task experience,  $F(29, 580) = 17.5, p < 0.001$ . There was also a significant Block  $\times$  Load interaction,  $F(29, 580) = 3.43, p < 0.001$ , where the number of fixations to find the target was reduced with experience and cognitive load ( $M_{\text{Dual-Block-1}} = 9.33; M_{\text{Dual-Block-30}} = 5.44; M_{\text{Single-Block-1}} = 6.43; M_{\text{Single-Block-30}} = 4.49$ ). Furthermore, there was a simple main effect for the single-task group where the number of fixations were significantly reduced across blocks,  $F(29, 319) = 1.61, p < 0.05$ . There was not a significant effect of search display type,  $F(1, 20) = 0.91, p > 0.35$ , nor was there a Significant Load  $\times$  Search Display Type interaction,  $F(1, 20) = 0.35, p > 0.56$ . No other effects reached significance.

The average number of refixations was submitted to 2 (load)  $\times$  3 (configuration group)  $\times$  [2 (search display type)  $\times$  30 (block)] mixed ANOVA. There were very few refixations, less than 0.33 on average, yet there were more than Experiment 1. There was a main effect of block,  $F(29, 609) = 13.12, p < 0.001$ , where the number of refixations was reduced across blocks. There was a main effect of load,  $F(1, 21) = 19.77, p < 0.001$ , where the dual-task group had more refixations on average per block of trials ( $M_{\text{Dual}} = 0.43$ ) than the single-task group ( $M_{\text{Single}} =$

0.24). Not surprisingly, there was a Block  $\times$  Load interaction,  $F(29, 609) = 4.39, p < 0.001$ , where the dual-task group reduced refixations across blocks at a faster rate than the single-task group.

The results demonstrate that participants increased their search efficiency across blocks—accuracy was maintained at a high proportion of correct trials (0.97) while the amount of time to find the target decreased by 539.7 ms from the first to the last block in the single-task condition and was decreased by 2392.5 ms in the dual-task group.

*Scan similarity analyses.* The similarity analysis was conducted to determine if visual scans increased in similarity during visual search, reflecting an established and regularly used method for finding targets in repeating search displays. We hypothesized that visual scans would increase in similarity with increased experience of repeating search trials. A failure to reject the hypothesis suggests that participants are repeating scans to find the target, whereas if visual scans do not increase in similarity then there is support for the anarchy hypothesis. The NSI metric used in Experiment 1 was also used in Experiment 2. Monte Carlo simulations were run to demonstrate a special case of the anarchy hypothesis and to provide a control along with the novel search displays.

To determine if there were differences in NSI values as a function of load, configuration group, search display type and epoch, a 2 (load)  $\times$  3 (configuration group)  $\times$  [2 (search display type)  $\times$  6 (epoch)] mixed ANOVA was performed on all mean NSI values. There was a main effect of search display type,  $F(1, 20) = 115.12, p < 0.001$ , where visual scans from repeating search displays ( $M_{\text{Repeating}} = 0.49$ ) were significantly more similar than visual scans from novel displays ( $M_{\text{Novel}} = 0.39$ ). There was a main effect of epoch,  $F(5, 100) = 29.29, p < 0.001$ , demonstrating an increase in similarity across epochs. Importantly, there was a reliable search Display Type  $\times$  Epoch interaction,  $F(5, 100) = 5.40, p < 0.001$ , demonstrating that visual scans from repeating search displays increased in similarity across epochs at a faster rate than visual scans from novel displays (see Figure 2). There was also a main effect of load,  $F(1, 20) = 10.55, p < 0.001$ , revealing that the single-task group's mean NSI value ( $M_{\text{Single}} = 0.47$ ) was significantly higher than the dual-task group's mean NSI value ( $M_{\text{Dual}} = 0.43$ ).

The results of the Monte Carlo simulations replicated the results from Experiment 1: The similarity of visual scans from novel and repeating search displays was greater than NSIs produced by pseudorandom sequences of comparable lengths to the participants' average scan length, again ruling out the anarchy hypothesis. These results suggest participants are learning something that leads to behavioral stability within the task environment. One possibility is that participants are learning the relatively small set of target locations used in both experiments. Another is that participants are learning a more general skill better distinguishing distractors and the target in periphery.

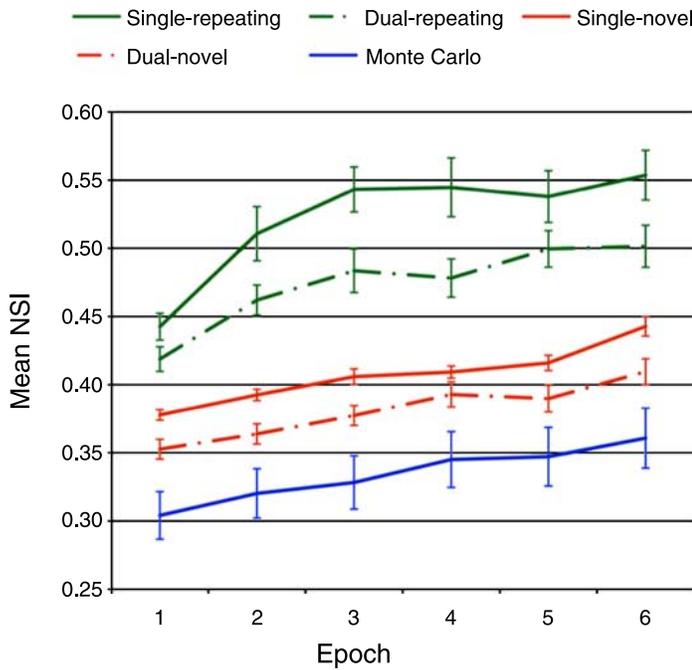


Figure 2. The epoch by search display type interaction on the normalized similarity indices from Experiment 2 with Monte Carlo Simulation results. Error bars from the human data represent 95% confidence intervals and represent standard error for the Monte Carlo simulation data.

*Underlying influences on visual scan similarity.* Increased cognitive load on cognitive resources reduced the similarity of scans, demonstrating the importance of endogenous processes (i.e., memory, attention, skill acquisition, etc.) on the repetition of scans. However, NSI values from repeating displays in the dual-task condition remained higher than NSI values from novel displays, suggesting that stable environmental information provides a means for repeating visual scans when cognitive processes are taxed. The design of Experiment 2 made it possible to determine if visual scan adaptation across epochs is solely a function of the visual stimulus. If the fixation locations composing a visual scan are determined from only external influences (e.g., the configuration of distractors and the target, visual salience, etc.) and if internal cognitive processes required for completing the task are invariant across individuals (i.e., memory, eye movement programming, etc.), then visual scans from different participants searching through the same repeating search displays, *between-participant NSI values*, should be as similar to visual scans from the same person searching through repeating displays, *within-participant NSI values*.

To determine between-participant NSI values, visual scans from participants searching through the same repeating displays were compared at each block for each of the 12 repeating trials, as well as for novel trials. For example, participant 1’s visual scan from repeating-display-A was compared to participant 2’s visual scan

from repeating-display-A and then to participant 3’s visual scan from repeating-display-A for each of the 12 repeating displays. Next, mean NSI values from each stimulus per block were averaged into epochs. Finally, all repeating stimuli were averaged together to get the mean repeating between-participant NSI at each epoch.

Although the within-participant and between-participant NSI values should not be submitted to statistical analyses because they contain the same data, Figure 3 shows that between-participant NSI values from repeating search displays are more similar than between-participant NSI values from novel search displays and that the degree of between-participant similarity is less than the within-participant similarity. Because between-participant and within-participant NSI values were not equivalent, the results indicate that visual scans were not produced solely from exogenous influences but were produced from a mix of internal cognitive processes interacting with the structure of the environment as argued by Josephson and Holmes (2002) and Foulsham and Underwood (2008). These results further rule out the anarchy hypothesis. It remains unclear which internal cognitive processes are producing the differences in visual scans across participants operating with the same goal within the same environment and provides a clear and ripe area for future research.

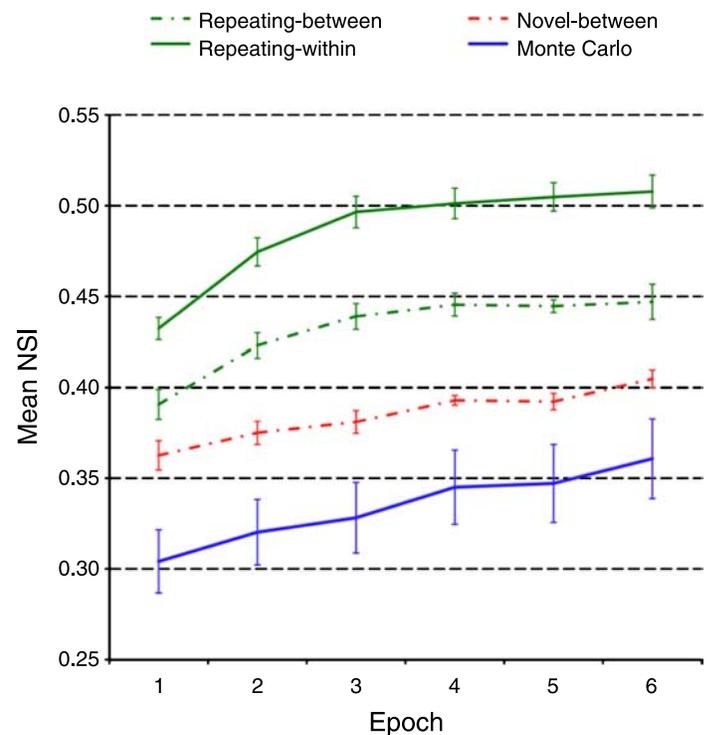


Figure 3. Between-participant and within-participant visual scan similarities from repeating and novel search displays from Experiment 2. Error bars represent standard error as no statistical test was conducted.

## Experiment 2 conclusion

There were three goals for [Experiment 2](#). The first goal was to provide further support for the adaptive scanning hypothesis by replicating [Experiment 1](#). The second goal was to manipulate visual scan similarity by adding dual-task conditions. The third goal was to determine if different people scan the same stimulus in a relatively similar manner.

The visual scan analyses corroborate the results from [Experiment 1](#), providing further support for the adaptive scanning hypothesis. First, search efficiency increased with task experience. Second, scans became more similar with increased task experience. Third, visual scan similarity increased at a rate faster than predicted by chance. Hence, results support the adaptive scanning hypothesis and reject the anarchy and general strategy hypothesis. Further, the results from [Experiment 2](#) extended the phenomena to environments that require more fixations to locate a target.

There was a reduction in visual scan similarity in the dual-task condition compared to the single task, demonstrating that endogenous processes are important to visual scan repetition. Furthermore, there is evidence that the structure of the environment (i.e., distractor and target configurations) contributed to the similarity of visual scans from repeating search displays, as evidenced with between-participant NSI values for repeating displays being greater than chance and less than within-participant NSI values ([Figure 3](#)).

## Conclusions and summary

Results from each experiment rejected the anarchy and general strategy hypotheses, providing support for the adaptive scanning hypothesis. Surprisingly, visual scans from novel search displays increased in similarity with experience at a rate similar to repeating displays in [Experiment 1](#). Further, scan adaptation occurred with very few fixations to locate the target.

[Experiment 2](#) provided further evidence of the adaptive scanning hypotheses in a search environment that required more than twice as many fixations on average to locate the target relative to [Experiment 1](#). The results of [Experiment 2](#) demonstrated that different participants searched through the same repeating search displays in a similar manner, but the similarity was not as high as that of visual scans within individual participants. This result suggests that the structure of the search display cannot completely account for visual scan similarity in repeating search displays.

The results of the experiments provide evidence that visual scans are repeated in, and adapted to, visual search

environments. In the following sections, we discuss influences on visual scans and areas for future research.

## Influences on visual scans

The presence of visual scan adaptation demonstrates that participants' adapted behavior resulted in increased search efficiency. As visual scans are composed of saccades to locations for fixation, mechanisms associated with selecting saccade locations, or saccadic selectivity, must be partly responsible for visual scan adaptation. There are typically two influences attributed to saccadic selectivity: *exogenous* and *endogenous*. Exogenous influences (i.e., bottom-up, data-driven) refer to hypothesized interactions between environmental stimuli sufficiently distinct from surrounding areas (i.e., salient) and hard-wired, invariant visual processes. Although exogenous processes clearly require some level of information processing (albeit at a relatively low level), they are typically thought of as reflexes facilitated by salient stimulus features. Salient stimuli have been shown to attract attention and affect visual scans (Franconeri, Simons, & Junge, 2004; Pomplun, Reingold, & Shen, 2003; Theeuwes, 2004; Wolfe, 1994). Exogenous influences are regularly considered non-deliberate (Everling & Fischer, 1998; Findlay, 1982, 1997; Kowler, 1990; Mitchell, Macrae, & Gilchrist, 2002) and can result in anarchic successions of individually programmed saccades (Wolfe et al., 2000).

Endogenous influences (i.e., top-down, goal-driven) refer to deliberate influences on saccades, such as the goals of making a sandwich or batting a ball. Task goals affect the distribution of dwell locations and durations (Hayhoe, 2000; Hayhoe & Ballard, 2005; Hayhoe, Shrivastava, Mruczek, & Pelz, 2003; Land, Furneaux, & Gilchrist, 2002; Land & Lee, 1994; Land & McLeod, 2000; Land & Tatler, 2001). Where endogenous processes have been shown to affect the sequence of fixations based on high-level goals, they have also been shown to influence the successions of saccades (Shen, Reingold, & Pomplun, 2000).

Scans that result from an endogenously influenced general strategy would not be adapted with increased task experience. Instead, they would be highly similar and remain approximately the same length as experience increased. Further, if participants' attempted different specific strategies throughout the experiments, then there would not be a steady rise in NSI values and a steady reduction in the number of fixations. Although the consistency of highly similar visual scans within individuals seems like a "straw man," these results are predicted by scanpath theory (Chernyak & Stark, 2001; Noton & Stark, 1971a, 1971b; Stark & Ellis, 1981; Stark et al., 1980). Visual scan similarities did not remain constant with increased task experience. Rather, visual scans

increased in similarity while their lengths were reduced with experience. Consequently, the results from both experiments do not support the general strategy hypothesis.

Visual scans resulting from only exogenous influences would not be adapted with increased experience, either. The same visual display would always influence scan patterns in nearly the same manner, leading to highly similar visual scans. Further, if exogenous influences are assumed to be relatively invariant across individuals, then visual scan repetition should be as similar across individuals as they are within individuals. [Experiment 2](#) showed that the similarity of scan patterns from different individuals across the same set of repeating stimuli did not reach the same level of similarity as scan patterns from the same stimulus within participants (see [Figure 3](#)). Results from both experiments do not support the anarchy hypothesis.

Taking a step back, it is clear that much skilled behavior entails learning where to saccade to obtain task-relevant information. For example, the difference between the best Cricket batters and the merely good ones lies in the superior ability of the best ones to execute an anticipatory saccade to where the ball will bounce (Land & McLeod, 2000). Given the fact that only the best players show these anticipatory saccades, it seems unlikely that such adaptations to environmental regularity result solely from either endogenous or exogenous influences.

The results from both experiments demonstrate that participants acquired a task-relevant skill, but that neither endogenous nor exogenous influences are the sole mechanism behind visual scan adaptation. Learning requires stability or regularity within an environment along with cognitive processes that can exploit the stability. For example, Reder et al. (2003) demonstrated that shifts of visual attention were sensitive to the probability that a given location would contain a target or a distractor. Furthermore, the authors report that their participants “were not consciously aware of any manipulation. Participants could not even report with any confidence or accuracy what locations tended to have more or fewer distractors” (p. 645). Chun and colleagues’ explanation of the contextual cuing phenomenon also implicates a non-deliberate and non-conscious adaptation to scene statistics (Chun, 2000; Chun & Jiang, 1998; Jiang & Wagner, 2004; Song & Jiang, 2005). In their paradigm, decreases in search times for repeated search displays were faster than for novel displays even in the absence of awareness that displays were repeating. Finally, the results from the dual-task condition suggest that the scan adaptation process was hampered when endogenous cognitive processes important to skill acquisition were taxed.

Each experiment demonstrated the ability of human cognition to exploit environmental stability within repeating and novel search displays. Indeed, efficiency gains associated with novel search displays in [Experiments 1](#) and [2](#) may have resulted from the adoption of a more general search skill than adapted visual scans. For example, participants may have exploited the small

number of possible target locations used across the novel search displays. Alternatively, participants may have become better at allocating attention away from distractors and toward targets. In one sense, using a small set of target locations in the novel trials is a limitation of the two experiments. If target locations from novel displays were randomly assigned making the novel displays truly unique and without repeating target locations, then the similarity of visual scans from truly unique trials should be closer to chance levels. This hypothesis was not tested in either experiment. However, using a small set of target locations in the novel displays demonstrated the sensitivity of our visual system to the statistical structure of the task environment.

The experiments presented here were just the first two on adaptive scanning in visual search and many questions remain. For example, a fruitful area for future research is to determine if visual scans are stored as whole or partial sequences of fixations that lead to a target and if so what information is stored within these “packaged sequences.” Further, is it that sequences of gross saccade trajectories are stored rather than relative object locations (e.g., cognitive map) to guide shifts of attention and saccades?

The results from the similarity analyses suggest that a strategy is increasingly used with more experience in the search environment. Another fruitful area of research is to determine how participants settle on a strategy. Previous research has indicated that individuals begin using and adapting multiple strategies until one eventually wins out and that the discovery of strategies is unconscious (Siegler & Stern, 1998). A general search strategy would lead to similar scans across repeated trials, and gradually settling on a general strategy would lead to a gradual increase in scan similarity as task experience increases. However, a general strategy would not necessarily result in a reduction in the number of fixations to find the target. However, it is possible that participants adopt a general strategy and with experience adapt it to task specifics. Uncovering the cognitive mechanisms behind the adoption and adaptation of visual search strategies that reflect skillful search is a ripe area for future research.

## Summary

The two experiments provide support for the adaptive scanning hypothesis. Interestingly, increased cognitive load was shown to disrupt scan pattern similarity. These results suggest an endogenous component in scan pattern repetition and refinement. Moreover, scan patterns from repeating stimuli viewed between participants were found to be more similar than novel stimuli, indicating an exogenous component in scan pattern repetition and refinement. Important extensions of this research could include extending the hypothesis to real-world scenes, determining the role scan patterns play during skill acquisition in complex tasks, uncovering mechanisms that

led to adaptive scanning, developing computational process models of adaptive scanning, and conducting further research on the interactions between endogenous and exogenous influences on visual scan formation.

## Acknowledgments

This research was conducted by the authors at Rensselaer Polytechnic Institute. The authors would like to thank Brett R. Fajen and members of the CogWorks Laboratory at Rensselaer Polytechnic Institute for their comments on earlier versions of this report. Support for this work was provided in part by grants to Wayne D. Gray from the Air Force Office of Scientific Research AFOSR #F49620-03-1-0143 and the Office of Naval Research ONR #N000140310046.

Commercial relationships: none.

Corresponding author: Christopher W. Myers.

Email: christopher.myers@mesa.afmc.af.mil.

Address: 2698 G Street Bldg. 190, Wright-Patterson Air Force Base, Dayton, OH 45433, USA.

## References

- Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Erlbaum.
- Chernyak, D. A., & Stark, L. W. (2001). Top-down guided eye movements. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, *31*, 514–522.
- Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, *4*, 170–178.
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71.
- Everling, S., & Fischer, B. (1998). The antisaccade: A review of basic research and clinical studies. *Neuropsychologica*, *36*, 885–899.
- Findlay, J. M. (1982). Global visual processing for saccadic eye movements. *Vision Research*, *22*, 1033–1045.
- Findlay, J. M. (1997). Saccade target selection during visual search. *Vision Research*, *37*, 617–631.
- Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, *8*(2):6, 1–17, <http://www.journalofvision.org/content/8/2/6>, doi:10.1167/8.2.6. [PubMed] [Article]
- Franconeri, S. L., Simons, D. J., & Junge, J. A. (2004). Searching for stimulus-driven shifts of attention. *Psychonomic Bulletin & Review*, *11*, 876–881.
- Furst, C. J. (1971). Automatizing visual attention. *Perception & Psychophysics*, *10*, 65–70.
- Godijn, R., & Kramer, A. F. (2006). Prosaccades and antisaccades to onsets and color singletons: Evidence that erroneous prosaccades are not reflexive. *Experimental Brain Research*, *172*, 439–448.
- Haider, H., & Frensch, P. A. (1999). Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 172–190.
- Hayhoe, M. M. (2000). Vision using routines: A functional account of vision. *Visual Cognition*, *7*, 43–64.
- Hayhoe, M. M., & Ballard, D. H. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, *9*, 187–194.
- Hayhoe, M. M., Shrivastava, A., Mruzcek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, *3*(1):6, 49–63, <http://www.journalofvision.org/content/3/1/6>, doi:10.1167/3.1.6. [PubMed] [Article]
- Hidalgo-Sotelo, B., Oliva, A., & Torralba, A. (2005). *Human learning of contextual priors for object search: Where does the time go?* Paper presented at the IEEE Computer Society Convergence on Computer Vision and Pattern Recognition (CVPR'05) Workshops, San Diego, CA.
- Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cueing—Configuration or individual location? *Perception & Psychophysics*, *66*, 454–463.
- Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., et al. (1997). Verbal working memory load affects regional brain activation as measured by PET. *Journal of Cognitive Neuroscience*, *9*, 462–475.
- Josephson, S., & Holmes, M. (2002). Attention to repeated images on the World-Wide-Web: Another look at scanpath theory. *Behavior Research Methods, Instruments, & Computers*, *24*, 539–549.
- Kowler, E. (1990). The role of visual and cognitive processes in the control of eye movement. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 1–70). New York: Elsevier.
- Laeng, B., & Teodorescu, D. S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science*, *26*, 207–231.

- Land, M. F., Furneaux, S. M., & Gilchrist, I. D. (2002). The organization of visually mediated actions in a subject without eye movements. *Neurocase*, 8, 80–87.
- Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature*, 369, 742–744.
- Land, M. F., & McLeod, P. (2000). From eye movements to actions: How batsmen hit the ball. *Nature Neuroscience*, 3, 1340–1345.
- Land, M. F., & Tatler, B. W. (2001). Steering with the head: The visual strategy of a racing driver. *Current Biology*, 11, 1215–1220.
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics, Doklady*, 10, 707–710.
- Mitchell, J. P., Macrae, C. N., & Gilchrist, I. D. (2002). Working memory and the suppression of reflexive saccades. *Journal of Cognitive Neuroscience*, 14, 95–103.
- Myers, C. W., & Schoelles, M. (2005). ProtoMatch: A tool for analyzing high-density, sequential eye gaze and cursor protocols. *Behavior Research Methods*, 37, 256–270.
- Noton, D., & Stark, L. W. (1971a). Scanpaths in eye movements during pattern perception. *Science*, 171, 308–311.
- Noton, D., & Stark, L. W. (1971b). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, 11, 929–942.
- Pomplun, M., Reingold, E. M., & Shen, J. (2003). Area activation: A computational model of saccadic selectivity in visual search. *Cognitive Science*, 27, 299–312.
- Reder, L. M., Weber, K., Shang, J., & Vanyukov, P. M. (2003). The adaptive character of the attentional system: Statistical sensitivity in a target localization task. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 631–649.
- Shen, J., Reingold, E. M., & Pomplun, M. (2000). Distractor ratio influences patterns of eye movements during visual search. *Perception*, 29, 241–250.
- Siegler, R. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology: General*, 127, 377–397.
- Song, J.-H., & Jiang, Y. (2005). Connecting the past with the present: How do humans match an incoming visual display with visual memory? *Journal of Vision*, 5(4):4, 322–330, <http://www.journalofvision.org/content/5/4/4>, doi:10.1167/5.4.4. [PubMed] [Article]
- Stark, L., & Ellis, S. (1981). Scanpaths revisited: Cognitive models direct active looking. In M. Fisher & Senders (Eds.), *Eye movements, cognitive and visual perception* (pp. 193–226). Hillsdale, New Jersey: Erlbaum.
- Stark, L., Ellis, S., Inoue, H., Freksa, C., Portnoy, Z., & Zeevi, J. (1980). Cognitive models direct scanpath eye-movements—Evidence obtained by means of computer-processing of perceptual scanpath eye-movements. *Physics in Medicine and Biology*, 25, 1000–1001.
- Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review*, 11, 65–70.
- Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, 1, 202–238.
- Wolfe, J. M. (1998a). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–73). East Sussex, UK: Psychology Press.
- Wolfe, J. M. (1998b). What can 1 million trials tell us about visual search? *Psychological Science*, 9, 33–39.
- Wolfe, J. M., Alvarez, G. A., & Horowitz, T. S. (2000). Attention is fast but volition is slow. *Nature*, 406, 691.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.
- Zangemeister, W. H., Sherman, K., & Stark, L. (1995). Evidence for a global scanpath strategy in viewing abstract compared with realistic images. *Neuropsychologia*, 33, 1009–1025.