Human visual search behaviour is far from ideal

Anna Nowakowska

Alasdair D.F. Clarke

Amelia R. Hunt

University of Aberdeen

Corresponding Author:

Anna Nowakowska

Address and email for reader correspondence: Room T32, William Guild Building, King's College, University of Aberdeen, a.nowakowska@abdn.ac.uk

Running Head: Human visual search behaviour is far from ideal

Author note:

Anna Nowakowska, Department of Psychology, University of Aberdeen

Alasdair D.F. Clarke, Department of Psychology, University of Aberdeen and the Department of Psychology, University of Essex

Amelia R. Hunt, Department of Psychology, University of Aberdeen

Abstract

Evolutionary pressures have made foraging behaviors highly efficient in many species. Eye movements during search present a useful instance of foraging behavior in humans. We tested the efficiency of eye movements during search using homogeneous and heterogeneous arrays of line segments. The search target is visible in the periphery on the homogeneous array, but requires central vision to be detected on the heterogeneous array. For a compound search array that is heterogeneous on one side and homogeneous on the other, eye movements should be directed only to the heterogeneous side. Instead, participants made many fixations on the homogeneous side. By comparing search of compound arrays to an estimate of search performance based on uniform arrays, we isolate two contributions to search inefficiency. First, participants make superfluous fixations, sacrificing speed for a perceived (but not actual) gain in response certainty. Second, participants fixate the homogeneous side even more frequently than predicted by inefficient search of uniform arrays, suggesting they also fail to direct fixations to locations that yield the most new information.

Keywords: Visual Search, Optimal Behaviour, Eye Movements

1. Introduction

Imagine that you are searching for a red pen, and you know it could be on either of two desks. The top of one desk is clean, while the other desk is cluttered with papers, other pens, books and coffee cups. What is the most effective way to find the red pen? Common sense suggests that a glance at the empty desk should be enough to detect the target if it is present, and the observer should spend the rest, or all, of their time searching the cluttered desk. An efficient visual system would not waste any time on the clean desk.

Several models of efficient foraging behaviour (e.g. 1, 2) have been developed, against which actual foraging behaviour can be measured. In humans, optimal models of search sample information efficiently by directing eye movements to locations that yield the maximum possible information or reward (3-5). In their influential model of visual search, Najemnik and Geisler (6, 7) demonstrated that eye movements are well-described by an optimal strategy, in which each saccade during search is directed to the location that will maximise the probability of detecting a target. A few recent studies, however, contradict key assumptions of the optimal search model. Notably, observers appear to be unable to adapt their fixation strategies on a trial-by-trial basis to changes in target frequency (8), or to changes in the expected difficulty of detecting the target in the periphery (9-11).

Alternatives to optimal foraging have been proposed: for instance, selection of eye movements during search have also been shown to be well-described by a stochastic process (12). In the stochastic model, each eye movement during search is randomly selected from the population of eye movement vectors that tend to be executed from the region of the search array that is currently fixated. The apparent contradiction with an optimal process can be resolved by the possibility that a combination of experience and evolution has shaped the population of eye movement vectors to produce relatively efficient search, without the need for complex calculations that must take into account information that can be difficult to estimate under most circumstances, such as expected target visibility across the retina. Eye movements can thereby *appear* optimal, even though the underlying process driving them is a far simpler heuristic. Consistent with stochastic processes driving selection of eye movements, there is some evidence that eye movements in reading follow a random walk (13), at least partially (14). However, models with a degree of guidance in reading tend to be favoured (for a review see 15, 16), with an emphasis on the orthographic and phonetic features that contribute to fixation selection processes.

In summary, the optimal and stochastic search models present two very different, but similarly effective, ways of explaining eye movements during search. To discriminate between these two models, here we test a straightforward prediction of an optimal search model: eye movements should be directed to locations that yield the most information. When faced with the search array depicted in Figure 1, and instructed to search for a line oriented 45° to the right, optimal observers should only make fixations to the more heterogeneous half of the array. If the target were on the more homogeneous side of the array, it would be easily detected using peripheral vision, making fixations to that side superfluous (details of a pilot experiment checking the suitability of our stimuli are given in the Supplementary Materials). If search is optimal, therefore, the proportion of fixations directed to the heterogeneous side on any given trial should be 1, because inspection of the homogeneous side will provide no additional information about the target location.

In the first experiment, we find that most participants over-fixate the homogeneous half of the display at the cost of increased reaction times. There are two possible (non-conflicting) explanations for this search inefficiency. First, it could reflect a failure to direct fixations in a manner that maximizes information gain, which would present a direct challenge to the optimal search model of Najemnik and Geisler (6). Second, participants may make unnecessary confirmatory fixations on both sides of the display. To separate these two plausible contributions to search inefficiency, we ran a second experiment using a mix of uniform homogeneous, uniform heterogeneous, and compound arrays. Search in the compound display may simply reflect an additive combination of how (in)efficiently participants search uniform displays. To the extent that search in compound displays is slower than predicted based on performance in uniform displays, we can conclude a failure to distribute fixations optimally across the two types of search arrays also contributes to search inefficiency.

2. Methods

a) Participants. Each experiment had 14 participants (28 total, with females=17; age range =20-62; mean age=25.3). Previous seminal experiments on this topic had a very small numbers of participants (e.g. N=2 in (6); N=4 in (9)) but report results from individuals separately rather than averaging them. Our sample is larger, but we maintain the approach of reporting individual differences (as in (10)).

b) Apparatus. The display was presented on a 17inch CRT monitor with a resolution of 1024x768. Stimulus generation, presentation and data collection were controlled by Matlab and psychophysics toolbox (17, 18) run on a Powermac. The position of the dominant eye was recorded using a desktop-mounted EyeLink 1000 eye tracker (SR Research, Canada) sampling eye position at 1000Hz.

c) Stimuli. The line segments were aligned in 22 columns and 16 rows on a uniform grey background. The target line was always tilted 45 degrees to the right. The mean distractor angle was perpendicular to the target angle. Search difficulty was manipulated by sampling from either a narrow 30° range of distractor line orientations ("homogeneous") or a wide 106°

range ("heterogeneous"). In a pilot study reported in full in the Supplementary material, we show that, when viewed while fixating screen centre, accuracy to detect the target was close to ceiling for homogeneous distractors (96 ± 5% for target present, 89 ± 13% for target absent) and close to chance for heterogeneous distractors (61 ± 13% for target present, 57 ± 17% for target absent). In Experiment 1, one half of each search array consisted of line segments with a homogeneous orientation, while the other half was heterogeneous (see Figure 1 for an example). Which side was heterogeneous was random on each trial. There were 160 trials in total, half of which contained a target. The side of the target relative to the search difficulty was counterbalanced. The target could be located in any of the possible locations apart from the middle four vertical columns.

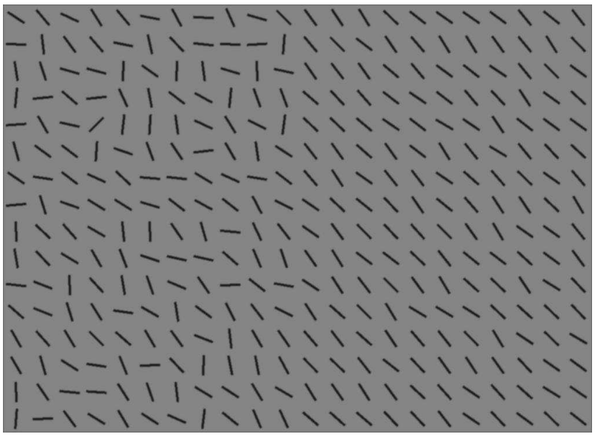

Figure 1. Example of a compound search array. The target is a line oriented 45° to the right. The target is present on the heterogeneous side in this example. The heterogeneous half of the array is shown on the left side.

In Experiment 2, the stimuli consisted of 80 homogeneous arrays, 80 heterogeneous arrays and 80 compound arrays. There were 240 trials in total, half of which contained a target. All the stimuli were displayed until the participant made a response (or timed out after 60 seconds).

d) Procedure. On arrival at the laboratory each participant was asked to read and sign a consent form and was seated alone in a low-lit room. Participants were told they would see line segments on the screen, and their task was to determine whether a line tilted 45° to the right was present among other lines. Participants were asked to respond as quickly and accurately as possible. Each trial consisted of a black fixation point (letter x) subtending 1.5x2.5cm (1.9°x3.1°), presented at the centre of the computer screen. On the press of a space bar, the stimulus was displayed until the participant made a response (or timed out after 60 seconds). Participants had to press either the left (present) or right (absent) arrow key. Auditory feedback in the form of a beep immediately followed incorrect key presses. Before the start of the experiment participants underwent a nine-point calibration sequence and a block of 10 practice trials.

3. Results

a) Experiment 1: search efficiency in compound arrays.

Reaction times (RT) for targets on the homogeneous side of the search array were faster than for targets on the heterogeneous side (mean RT and SD for homogeneous (1.75 ± .13), heterogeneous (3.94 ± 2.19) and absent (7.0 ± 4.5) conditions). Mean accuracy for target absent trials was ≈100%. For target present, participants were more accurate when the target was located on the homogeneous side of the display (98.4%), than the heterogeneous side (72.8%, ($t(13)=6.7$, $p<0.001$).

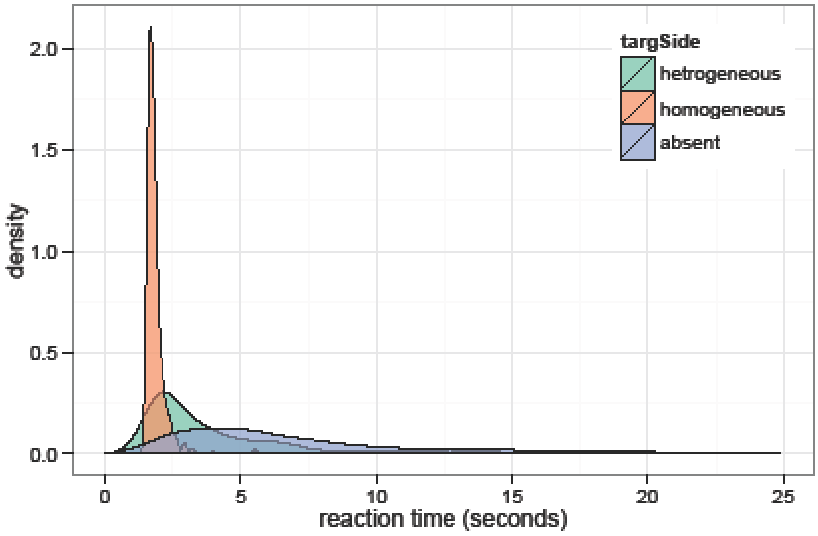

Figure 2. Distribution of reaction times across conditions.

Figure 3 shows the proportion of fixations each observer made on the heterogeneous side of the display on target absent trials only. The strictest criteria of optimal strategy in this experiment is not to look to the homogeneous side at all. (The pilot study in the Supplementary material demonstrates it can be easily ascertained whether the target is present on this side or not from the central fixation point.) Fixations on this side will provide no new information on the target's location, so participants should direct all fixations to the heterogeneous side. As we can see in Figure 3, only Participant 11 is close to executing the optimal strategy. In aggregate, our participants spend more time fixating the heterogeneous than the homogeneous side (Figure 4), but for the majority of participants a large proportion of fixations are made to the homogenous side.

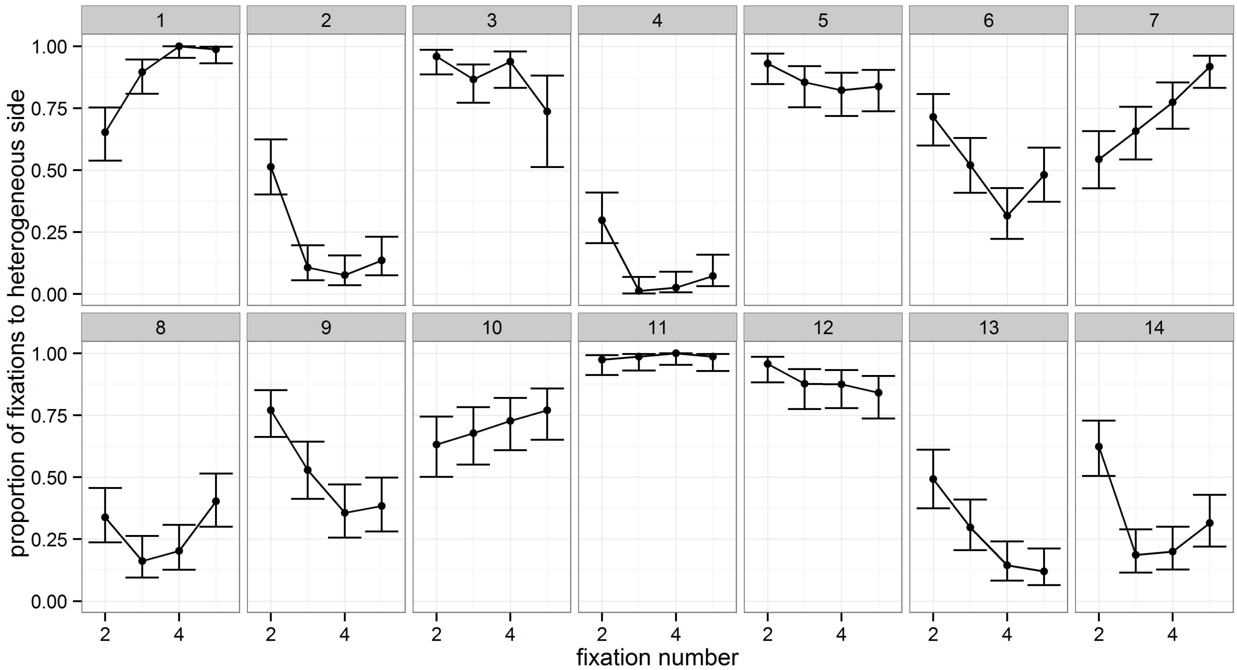

Figure 3. Proportion of the first five fixations on the homogeneous side for each observer. Only target absent trials are shown here. Fixations in the central region (1 degree to the left and right of the centre of the screen) have been excluded.

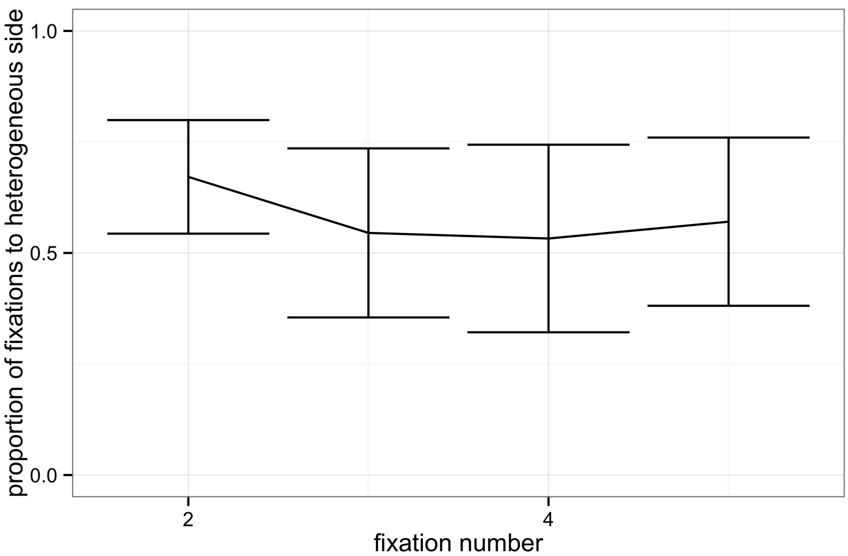

Figure 4. Mean proportion of saccades directed towards the heterogeneous side of the search array on target absent trials. Only fixations that are further than 1° to the left or right of the center of the display have been included in this analysis.

Next we measured the effect of this fixation inefficiency on the search performance of each participant. Inefficiency was defined as the proportion of the first five fixations made during target absent trials that were directed to the homogeneous side of the display. This measure was significantly correlated (see Figure 5A) with the median reaction time on target present trials, both when the target was located on the heterogeneous half of the display ($r=0.93$, $p<0.001$) and on the homogeneous side ($r=0.81$, $p=0.002$). These correlations are also significant when taking the proportion of the first 10 fixations (heterogeneous $r=0.89$, $p<0.001$; homogeneous $r=0.71$, $p=0.01$).

We also quantified the effect of fixation inefficiency on search time using a linear mixed-effect model (using the **lme4** (19)) package for **R** (20) with random intercepts and slopes. We were specifically interested in the effect of the number of homogeneous fixations on any given trial on the reaction time to find the target (including participant as a random factor). For target absent trials, we find an additional 357ms (bootstrapped 95% confidence interval: 196-516ms) in reaction times for every fixation made to the homogeneous side of the array (see Figure 5B). When the target is present on the heterogeneous side, each fixation on the homogeneous side slows reaction time by 547ms. Homogenous fixations even slow reaction time to find the target when it is present on the homogeneous side (by 159ms), consistent with the conclusions from our pilot study (see Supplementary Information) that these fixations are not necessary to find the target.

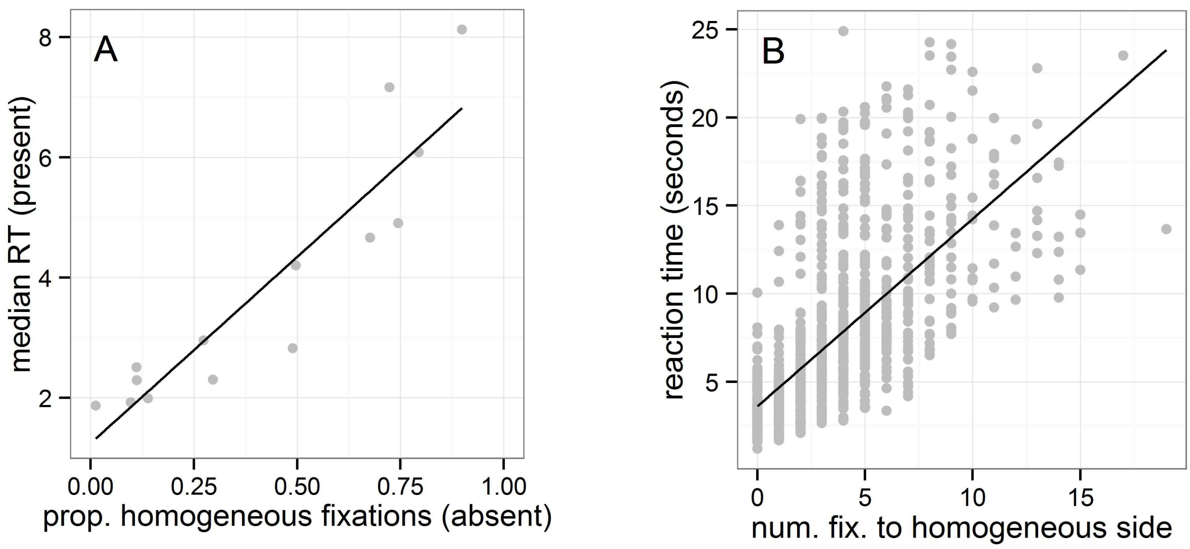

Figure 5. A. Mean reaction time on trials where the target was present on the heterogeneous side for each observer is highly correlated with the mean proportion of the first five fixations directed to the homogenous side of the display on target absent trials. B. Reaction time on each target absent trial as a function of how many fixations were made on the homogeneous side of the display. For every homogeneous-side fixation, reaction time increases by 360ms.

b) Predicting search performance on compound arrays from performance on uniform arrays.

In this experiment, participants searched uniform homogenous, uniform heterogeneous and compound search arrays. Summary of participants' reaction times and accuracy across all the conditions can be seen in Table 1.

Search array	Target Condition	Reaction Time (±SD)	Accuracy (±SD)
Homogeneous	Present	1.77 (.13)	97.32 (3.32)
	Absent	2.84 (.73)	97.86 (3.91)
Heterogeneous	Present	3.28 (2.23)	55.00 (20.55)
	Absent	6.94 (4.55)	93.39 (6.09)
Compound	Homogeneous side	1.84 (.17)	97.42 (4.60)
	Heterogeneous side	3.42 (2.42)	48.10 (23.99)
	Absent	6.03 (3.28)	95 (6.36)

Table 1. Mean of the median Reaction Times (s) and mean Accuracy (%) across conditions.

When the target was absent, participants made, on average, seven eye movements in the uniformly homogenous display before making a response. Each of these fixations can be considered unnecessary, given that participants in the pilot experiment were close to 100% correct with no eye movements at all. If search on compound trials is simply an (optimal) combination of suboptimal search behaviour on the two types of uniform trials, then RT on the compound trials should equal the average of RT on the uniform homogeneous and uniform heterogeneous trials. If equal, this would suggest our participants simply sacrifice efficiency to satisfy an overly conservative certainty criterion. To the extent that search is slower on compound trials compared to the average of the two types of uniform trials, an inflated certainty criterion alone does not explain poor search behaviour.

Figure 6 shows predicted and actual RT for each participant on the target absent trials. All participants lie above the red line (although three are very close). This indicates that participants are taking longer than predicted from the uniform trials. To quantify the size of the difference, we calculated the ratio of split versus predicted RT for each participant. If participants' behaviour on the compound trials matches an average of the behaviour they exhibit on the uniform trials, the ratio should be around 1. The mean ratio was 1.21(± .15), significantly higher than 1 ($t(13)=30.95$, $p<.001$). This additional slowing of reaction time in the compound trials can be attributed specifically to an inefficiency in allocating fixations to locations that yield the most information.

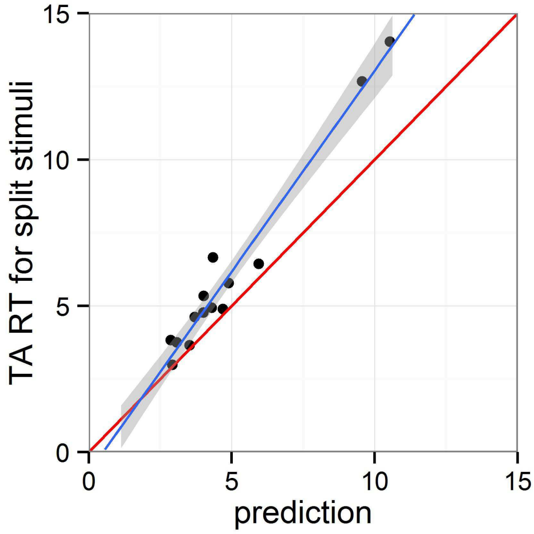

Figure 6. The red line represents predicted RT on target absent trials (mean homogeneous and mean heterogeneous RT averaged together). The blue line represents the actual RT on the split screen trials. Most points are above the line, suggesting participants take longer on the split screen trials than predicted from their behaviour on the full screen trials.

4. General Discussion

Our participants consistently failed to adopt an optimal strategy when searching a compound array with easy search on one side and difficult search on the other. In the first experiment, a large number of saccades were directed to the easy side of the display, even though the target would be clearly visible from the central fixation point if it were present on this side. Each one of these unnecessary fixations slows search substantially. In the second experiment, we demonstrated that participants also search uniform displays inefficiently, generally making many more fixations than is necessary to find the target. Although we demonstrated in the pilot experiment that the peripheral information is sufficient to decide the target is present or absent, observers may be driven to verify their peripheral estimate based on the clearer, higher-resolution visual information that can be obtained by bringing that image onto the fovea, even though this verification comes at great cost to speed. Indeed, previous results suggest participants tend to make saccades even when they are not necessary (21, 22). Importantly, the inefficiency of search in the compound display reflects more than an additive combination of how inefficiently participants search the two types of uniform displays. The additional inefficiency associated with the more complex array can be attributed to a failure to direct saccades to locations that can easily be estimated to provide the most information.

Taken together, these experiments clearly demonstrate that a large proportion of fixations made during visual search are not guided by the principles behind the optimal search model (6, 7). Not only do observers demonstrate a preference for making far more fixations than is required – presumably to increase their perceived certainty – but even taking these sub-optimal fixations into account, fixations in the split-screen array are not directed to locations that yield the most information. Participants were instructed to respond as quickly as possible, and responses on target absent trials were slowed by 360ms for every fixation they made on the homogenous side of the array. Nonetheless, it is possible that participants are capable of searching more efficiently but, for reasons of motivation or distraction, fail to implement an efficient strategy. Further research would be needed to determine the extent to which reward or greater pressure speed (for example by using response deadlines) would increase efficiency. It is important to note, however, that our results demonstrate an efficient strategy is not the dominant or default mechanism for fixation selection.

What is the mechanism for fixation selection? A viable alternative to the optimal search model, recently been proposed by Clarke et al (12), is that a scan-paths during visual search can be modelled using a random walk. This model is consistent with the mean performance of our participants, which is around 50% to each side. A largely stochastic model would predict this pattern. That said, this average performance masks a large range of individual differences. Indeed, one of our participants does follow the predictions of the ideal search model, and two others come quite close. Similarly, the stochastic model can explain some, but not all, of our individual participants. It therefore seems likely that different models will be required to fit different observers. An intriguing question is the extent to which search and foraging strategies are stable in individuals over time and across different contexts, shedding light on the nature of the efficient foraging, as well as the constraints on fixation selection mechanisms and how these are imposed. It would also be interesting to text the extent to which an individual's (in)efficient foraging decisions generalise to other kinds of decisions. For example, we have recently reported profound inefficiencies in decisions about how to allocate resources over multiple possible goals (10). The wide range of individual differences observed in both that study and the current one presents an intriguing parallel.

Individual differences aside, on the whole we can conclude that eye movements are not driven preferentially to locations that produce the most information. These results demonstrate that the processes underlying fixation selection during visual search may be more random and less efficient than current popular models suggest.

5. References

1. Charnov, E. L. Optimal foraging, the marginal value theorem. Theor Popul Biol. 1976; 2: 129-136.

2. McNamara, J. Optimal patch use in a stochastic environment. Thoer Popul Biol. 1982; 21: 269-288.

3. Cain, S. M., Vul, E., Clark, K., & Mitroff, S. R. A Bayesian optimal foraging model of human visual search. Psychol Sci. 2012; 23: 1047 – 1054.

4. Ma, W.J., Navalpakkam, V., Beck, J.M., Van den Berg, R. & Pouget, A. Behaviour and neural basis of near-optimal visual search. Nat Neurosci. 2011; 14: 783–790.

5. Navalpakkam, V., Koch,C., Rangel, A., & Perona,P. Optimal reward harvesting in complex perceptual environments. Proc Nati Acad Sci U. S. A. 2010; 107: 5232-5237.

6. Najemnik, J. & Geisler, W. S. Optimal eye movement strategies in visual search. Nat. 2005; 434: 387-391.

7. Najemnik, J., & Geisler, W.S. Eye movement statistics in humans are consistent with an optimal search strategy. J Vis. 2008; 8:4.

8. Verghese, P. Active search for multiple targets is inefficient. Vision Res. 2012; 74: 61-71

9. Morvan, C. & Maloney, L.T. Human visual search does not maximise the post-saccadic probability of identifying targets. *PLoS* Comput Biol. 2012; 8: e1002342.

10. Clarke, A.D.F. & Hunt, A.R. Failure of intuition when choosing whether to invest in a single goal or split resources between two goals. Psychol Sci. 2016; 27: 64-74.

11. Nowakowska, A., Clarke, A.D.F., Sahraie, A., & Hunt, A. R. Inefficient search strategies in simulated hemianopia. J Exp Psychol Hum Percept Perform. 2016; DOI: 10.1037/XHP0000250

12. Clarke, A.D.F., Green, P., Chantler, M.J., & Hunt, A. R. Human search for a target on a textured background is consistent with a stochastic model. J Vis. 2016; 16: 4.

13. Fend, G. Eye-movements as time-series random variables: A stochastic model of eye movement control in reading. Cog Syst Res, 2006; 7: 70-95.

14. Engbert, R. & Kliegl, R. Mathematical models of eye movements in reading: a possible role for autonomous saccades. Biol Cybernetics. 2001; 85: 77-87.

15. Rayner, K, & McConkie, G.W. What guides a reader's eye movements? Vis Res. 1976; 16: 829-837.

16. Rayner, K. Eye guidance in reading: Fixation locations within words. Percept. 1979; 8: 21-30

17. Brainard, D.H. The psychophysics toolbox. Spat Vis. 1997; 10: 433-436.

18. Pelli, D.G. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spat Vis. 1997; 10: 437-442.

19. Bates, D., Maechler, M., Bolker, B., & Walker, S. Package ‘lme4’. J Stat Softw. 2015; 67: 1-48, doi:10.18637/jss.v067.i01.

20. R Core Team. R Foundation for Statistical Computing, Vienna, Austria. 2015; URL https://www.R-project.org/.

21. Watson, M.R., Brennan, A.A., Kingstone, & A. Enns J.T. Looking versus seeing: strategies alter eye movements during visual search. Psychonomic Bull Rev. 2010; 17: 543-549.

22. Klein, R. & Farrell, M. Search performance without eye movements. Percept Psychophys. 1989; 46: 476-482.

6. Supplementary Materials - Pilot Experiment

a) Method

Participants. Ten participants (females=8; age range=21-30; mean age=25) with normal or corrected to normal vision completed the experiment.

Apparatus. The display was presented on a 17inch CRT monitor with a resolution of 1024x768. Stimulus generation, presentation and data collection were controlled by Matlab and psychophysics toolbox (Brainard, 1997; Pelli, 1997) run on a Powermac. The position of the dominant eye was recorded using a desktop-mounted EyeLink 1000 eye tracker (SR Research, Canada) sampling eye position at 1000Hz.

Stimuli and procedure. Search arrays of line segments are illustrated in Figure 1. The line segments were aligned in 22 columns and 16 rows on a uniform grey background. The target line was always tilted 45 degrees to the right. The mean distractor angle was perpendicular to the target angle. Search difficulty was manipulated by sampling from either a narrow 30°

range of distractor line orientations ("homogeneous") or a wide 106˚ range ("heterogeneous"). The side of the target was counterbalanced. The target could be located in any of the possible locations apart from the middle four vertical columns. There were 80 trials, half of which contained a target.

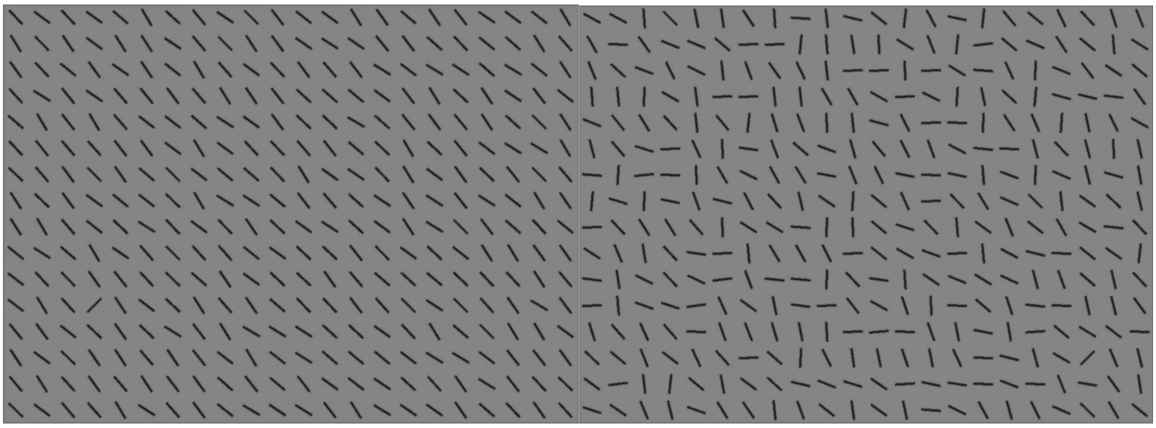

Figure 1: Example stimuli.

On arrival at the laboratory each participant read and signed a consent form and was seated alone in a low-lit room. Participants were told they would see line segments on the screen for a very short time, and their task was to determine whether a line tilted 45° to the right was present among other lines. Participants were asked to respond as accurately as possible.

Each trial consisted of a black fixation point (letter x) subtending 1.5x2.5cm (1.9°x3.1°), presented at the centre of the computer screen. On the press of a space bar, the stimulus was displayed for 200ms follow by a blank screen. Participants had to press either the left (present) or right (absent) arrow key. Auditory feedback in the form of a beep immediately followed incorrect key presses. Before the start of the experiment participants underwent a five-point calibration sequence and a block of 10 practice trials.

b) Results

Mean accuracy for the homogeneous stimuli was close to 100% (96 ± 5 % for target present, 89 ± 13% for target absent), while accuracy for the heterogeneous line segments was close to chance (61 ± 13% for target present, 57 ± 17% for target absent). When viewed from a central point, our observers were close to 100% correct to detect the target in the homogeneous array and close to chance in the heterogeneous array.