



Subject Areas:

Behaviour, evolution

Keywords:

visual search, optimal behaviour, eye movements

Author for correspondence:

Alasdair Clarke

e-mail: a.clarke@essex.ac.uk

Stable individual differences in strategies within, but not between, visual search tasks

Alasdair D. F. Clarke¹, Jessica L Irons²,
Warren James³, Andrew B Leber² and
Amelia R Hunt³

¹Department of Psychology, University of Essex,
Colchester, UK

²Department of Psychology, The Ohio State University,
Columbus, USA

³School of Psychology, University of Aberdeen,
Aberdeen, UK

Some abstract goes here

1. Introduction

As is common in cognitive psychology, most visual search literature has focused on how the average participant performs in the task, despite it being well known that there is a great deal of variability between one subject and the next. From Treisman's work on Feature Integration Theory [22] to the latest incarnation of the Guided Search Model [24], we now have a good understanding of what makes some search targets effortless to find, while others require careful inspection. However, these theories and models have neglected the question of why some observers find visual search so much harder than others. These differences can emerge from several different sources of variation: tiredness [13], information-processing ability, speed-accuracy trade-off, motivation, visual impairments [16], and search strategies [1]. Although their existence has previously been noted [13], a rigorous examination of individual differences in visual search is a challenge that has not been taken up by many researchers, and questions about their importance and stability remain relatively under explored.

In the current study, we focus on one source of individual differences in visual search: strategy. By search strategies, we refer to a collection of behaviours that all observers can freely choose from when they search a display. Examples include choosing to adopt a systematic

© 2014 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License <http://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, provided the original author and source are credited.

A striking example of the effect of strategy is given by Boot and colleagues [1]. They asked participants to monitor a cluttered display for an object changing colour or suddenly appearing. Large individual differences were found with respect to the number of saccades participants made while monitoring the stimulus, and this was negatively correlated with detection performance. Proulx [18] demonstrated that participants differed in how they prioritised visual features (color or orientation) when searching through a set of objects for one with a particular color-orientation combination. These differences in search strategy explained some of the variation in the extent to which participants were distracted by an irrelevant object of a unique size. However, participants were largely unaware of the search strategy they employed, and the majority were unable, or unwilling, to comply with instructions to employ a particular strategy.

Eye movement strategies have also been shown to be an important source of individual differences in visual search efficiency. Nowakowski, Clarke & Hunt [15] designed a simple search paradigm to discriminate between optimal [14] and stochastic [3] search strategies. The task involved searching through arrays of line segments arranged such that those on one side of the display all had a very similar orientation ("homogeneous"), while those on the other side had higher variance ("heterogeneous", see Figure 1 for an example). This meant that targets appearing on the homogeneous side would be highly salient, and targets on the heterogeneous side would be much harder to find. With these stimuli, the optimal eye movement strategy is to only search the heterogeneous half, as targets on the homogeneous side can be detected with peripheral vision. We will refer to this paradigm here as the Split-Half Line Segment task (SHLS). While some participants initially searched the displays near optimally, others carried out strategies counter to this, failing to even match the performance of the stochastic searcher. Furthermore, the degree to which they made saccades in line with the optimal search strategy was strongly correlated with their reaction times. A related version of this paradigm has been used in research investigating eye movement strategies

in response to (simulated) hemianopia [16,17], with similar conclusions: the full spectrum of individual differences in strategy was observed, from near-optimal to the opposite of optimal. It is therefore not possible to conclude whether optimal or stochastic models better describe search without first explaining individual variability.

A similar range of strategies, from random to near-optimal, has been found by Irons & Leber with the Adaptive Choice Visual Search (ACVS) paradigm [8], designed to explore how feature-based attention is used. This paradigm involves stimuli made up of small coloured boxes (red, blue, green and a fourth colour that varies from red, through purple, to blue and back again) with numerals written inside them (See Figure 1). The target is defined as a red or blue box with a numeral between 1 and 4, and on each trial one target of each colour is present. The participant's task is to find one of either target as quickly as possible and report the numeral. On trials in which the fourth colour is red (or close to red), participants should search through the blue boxes and report the blue target, as there will be fewer distracters. As the fourth colour changes to through purple to blue, participants should update their strategy and search for the red target. The results showed that participants varied substantially along two key dimensions: how frequently they used the more effective target color to search (varying from chance performance to near optimal), and how often they changed between colours. Further work has shown that these differences are stable over time (between one and ten days) with test-retest correlations of around $r \approx 0.8$ [9].

««««< HEAD Another example of differences in search strategy comes from the foraging literature [11,12]. In this context, foraging is a visual search task in which there are multiple targets on each trial. Participants were asked to search through a set of items from four categories, with two categories classed as targets. In the conjunction condition (searching for red-horizontal and green-vertical line segments among red-vertical and green-horizontal distracters), most observers searched in runs, finding all the targets of one target category, and then switching and finding the targets in the other category. This strategy has

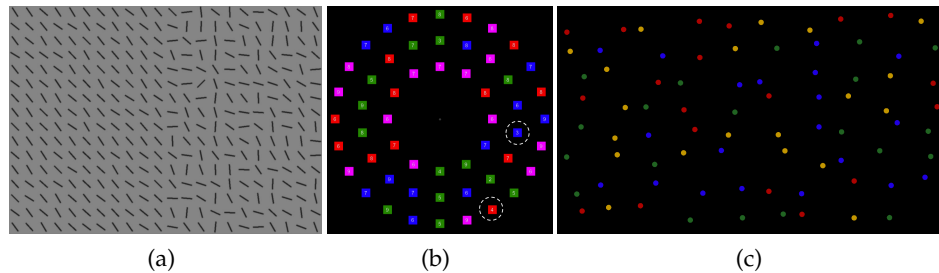


Figure 1. Example stimulus from the (a) *split-half*, (b) *adaptive choice* and (c) *foraging* paradigms.

previously been observers in animal foraging [5], and suggests holding one complex target template in mind at a time is a better strategy than switching templates. However, a subset of observers, termed ‘super-foragers’, were able to change between search target categories with no cost to performance. This study has been run on touch displays [12] (observers are asked to tap on as many targets as they can as quickly as possible) and a 3D computer game (in which participants navigate a squirrel through a virtual environment picking up targets with different feature combinations) [IN PREP], with super-foragers present in both versions. While test-retest reliability has not been measured explicitly for the foraging paradigm, the task was used as a measure to assess the effect of a six day long mindfulness retreat on cognitive performance [7]. From a re-analysis of these data, we can estimate that the test-retest reliability is $r \approx 0.7$ for the feature condition and $r \approx 0.88$ for the conjunction search. ===== Another example of differences in search strategy comes from the foraging literature [11,12]. In this context, foraging is a visual search task in which there are multiple targets on each trial. Participants were asked to search through a set of items from four categories, two of which were classed as targets. In the conjunction condition (i.e. searching for red-horizontal and green-vertical line segments among red-vertical and green-horizontal distracters), most observers searched in runs of one target category or another. This strategy has previously been observed in animal foraging literature [5]. However, a sub-set of observers, termed ‘super-foragers’ showed no switch cost. This study has been run on touch displays [12] (observers are asked to tap on as many targets as they

can as quickly as possible) and a 3D computer game (in which participants navigate a squirrel through a virtual environment picking up targets with different feature combinations) [IN PREP], with super-foragers present in both versions. While test-retest reliability has not been measured explicitly for the foraging paradigm, the task was used as a measure to assess the effect of a six day long mindfulness retreat on cognitive performance [7]. From a re-analysis of these data, we can estimate that test-retest reliability for the mean lengths of runs is $r \approx 0.88$. »»»» origin/master

Previous research has investigated the relationship between these behaviours to measures, but to date, these differences have not shown strong correlations with other attributes. In the ACVS paradigm, Irons & Leber found no evidence of a correlation between the proportion of optimal choices made by observers and measures of visual working memory or trait compulsivity [8]. Similarly, the differences in foraging behaviour are not accounted for by working memory or inhibitory control [10].

A common theme emerging from these studies is the observation that individual strategies vary in their degree of effectiveness or optimality. However, “visual search” encompasses a wide range of specific tasks, each tapping into a different aspect of behaviour (e.g. feature-based attention, information sampling). The aim of the present study is to investigate the extent to which individual differences are stable across different visual search paradigms. Are observers who use the optimal strategy in the SHLS paradigm also more optimal in the ACVS task? Does it make sense to talk about ‘super-searchers’ who show above average performance in a range of search tasks (analogous to the ‘super-recognizers’ of the

face-recognition literature [19]? Are the super-foragers consistently better or worse than more typical searchers in the other two paradigms? As a secondary question, we will measure the test-retest reliability of the differences found in the SHLS paradigm.

2. Methods

The methods and planned analysis for this study were registered on the Open Science Framework¹ before data collection started.

(a) Participants

64 students from the University of Aberdeen took part in this study². Participants were compensated for their time with either course credit or £15. All participants will sign a form giving informed consent. The study was approved by the University of Aberdeen Psychology Ethics Committee.

Sample size was determined in part by a power analysis, and in part due to constraints with counter-balancing. $n = 64$ participants means we should be able to detect correlations with $r > 0.34$ with $\alpha = 0.05$, $\beta = 0.80$ between the different visual search paradigms. The sample size is therefore of sufficient size to detect relatively small correlations.

(b) Materials and Procedures

The study consists of three paradigms from the visual search literature in which large individual differences have been found [8,12,15]. Example stimuli can be seen in Figure 1. A brief overview of each paradigm is given below, with full details in *supplementary materials*. The three tasks will be completed over two sessions, approximately one week apart. The SHLS was run in both sessions allowing us to measure test-retest reliability. The order in which participants completed the tasks was counter-balanced. There are 16 different possible orders of tasks/conditions; four participants completed each order for a total of 64.

The display was presented on a 17-inch CRT monitor with a resolutions of $1400 \times$

1050 ($n = 40$) or 1600×1200 ($n = 24$). Stimulus generation, presentation and data collection were controlled by MATLAB and the psychophysics and eyelink toolboxes [2,4] run on a Powermac.

(i) Split-half Line Segments

Stimuli consisted of arrays of black oriented line segments against a grey background. Each line segment had a length of $\approx 1.6^\circ$ of visual angle irrespective of screen resolution. The target was oriented 45° clockwise, while the distractor items had a random orientation with a mean of 45° anti-clockwise. The variance was low (18°) on one half of the display to create a homogeneous texture, and high (95°) on the other side to create a heterogeneous texture. This means that when the target is present on the homogeneous side of the stimulus, it can be easily be detected with peripheral vision, but when it is in the heterogeneous half, it is much harder to detect. This was verified in a pilot experiment reported in [15]. There were a total of 160 trials and homogeneous and heterogeneous sides of the display were randomly varied from trial to trial. The position of the dominant eye was recorded using a desktop-mounted EyeLink 1000 eye tracker (SR Research, Canada).

This paradigm was carried out twice, once in each testing session, to give us an estimate of how consistent participants are in their search strategy over time.

(ii) Adaptive Choice Visual Search

The ACVS was based on the task described in [8], Experiment 1. Each search display was composed of 54 small squares arranged in three concentric rings around fixation, with 12, 18 and 24 items in the inner, middle and outer rings respectively. The same screen was used as in the Split-half Array Search, however, due to changes of screen resolution, the size of the squares changed. Participants sat ≈ 47 cm from the screen. Of the 54 squares, 13 were red, 13 were blue, 14 were green and 14 were "variable". Variable distractors change colours from trial-to-trial according to a 24 trial cyclical pattern: the distractors would be red for five trials (called a "red plateau"), then across a period of seven trials, they would change colour from almost red to magenta (at the fourth trial in the transition) to almost blue. The

¹insert URL

²data from an additional 11 participants was discarded due to being recorded with an inappropriate screen resolution

variable distractor would then be blue for five trials (blue plateau), and then transition back from almost blue through magenta to almost red.

A white digit appeared inside each square. Two targets - a red square and a blue square each with a digit between 2 and 5 - were embedded in every search display. The two target digits were always different, to enable us to distinguish the chosen target. The remaining red, blue and variable squares all contained digits between 6-9. Green squares could contain any digit between 2-9. The location of the targets and distractor within the search display were randomized on each trial.

Participants were informed that the search displays would contain two targets on every trial, that they need only find one target on each trial and that they were always free to search for either one.

(iii) Mouse Click Foraging Task

The foraging task (MCFT) was based on [12] and [11] except that we used a mouse clicks to register responses rather than a touchpad.

In the feature foraging task, search displays contained 80 small circles, 20 red, 20 green, 20 blue and 20 yellow. Stimuli were arranged in a 10×8 grid, but the position of each item within the grid space was jittered to create a more random spatial arrangement. The location of item colours to grid locations was completely randomized. For half of the participants, targets were red and green circles, and for the other half of participants, targets were blue and yellow circles. Participants were asked to collect all of the targets within a trial by using the mouse to click on each target. Clicking on a target caused it to disappear from the display. If the participant clicked erroneously on a non-target, the trial immediately ended and a replacement trial was begun. In the conjunction foraging task, search displays were composed of both circles and squares. For half of the participants, the shapes were red and green (equal numbers of red circles, red squares, green circles and green squares), and for the remaining participants the shapes were blue and yellow. Targets were defined by conjunctions of colour and shape (e.g., red squares and green circles, with red circles and green squares as distractors). The assignment

of targets and distractors was assigned at random for each participant. The procedure was otherwise identical to the feature foraging task.

3. Results

(a) Replication of each paradigm

(i) Split-half Line Segments

Our results are consistent with the original SHLS study [15]. Only trials with a correct target absent/present response were included in the analysis. Some participants performed exceptionally poorly and were removed from further analysis³. All reaction times were \log_2 transformed. In order to characterise each participant's behaviour in this task, we will compute the proportion of the first n fixations that were on the heterogeneous (difficult) side of the stimuli, over all correct target absent trials. Previous work [15] demonstrated a strong correlation between this metric (for $n = 5$) and reaction times ($r = .93$).

The correlation between accuracy and reaction times between the two sessions is shown in Figure 2(a, b). We can clearly see that there are large differences from one participant to the next in terms of both the proportion of hard targets found, and reaction times. Furthermore, test-retest reliability appears to be reasonable, with Pearson's $r \in [0.71, 0.89]$ (95% confidence interval) for accuracy in finding targets on the hard heterogeneous half of the display. We get similar scores for the correlation in reaction times between sessions a and b for heterogeneous targets, ($r \in [0.54 - 0.81]$), homogeneous targets ($r \in [0.52 - 0.80]$) and target absent ($r \in [0.66 - 0.86]$). The reduced correlation for the homogeneous targets is likely due to the restricted range.

We can also look at the initial search strategies adopted by our participants 2(c, d). Again, we see large and stable individual differences across the two sessions (test-retest $r \in [0.63, 0.86]$ for the proportion of the first five saccades to the heterogeneous half of the display for target absent trials). More importantly, as with [15], we see that the search strategies give a good correlation with reaction

³Full details given in the *supplementary materials*

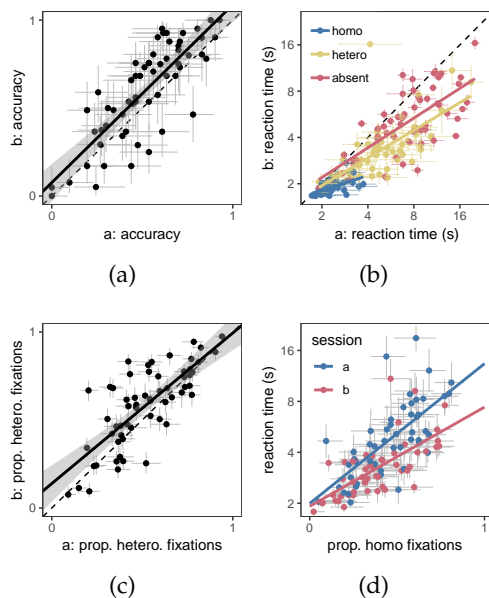


Figure 2. Each point represents a participant and the error-bars indicate 95% confidence intervals. Correlation between the two sessions of the SHLS paradigm for (a) accuracy (TP-heterogeneous trials only); (b) reaction times and (c) search strategy (TA trials only). (d) Initial search strategy correlates with reaction times in both sessions.

times in both session a, $r \in [0.52, 0.82]$ and session b, $r \in [0.50, 0.80]$.

(ii) Adaptive Choice Visual Search

Participants with accuracy more than 3 SD below the group mean were excluded from any ACVS analyses. This resulted in the exclusion of two participants, leaving 61 remaining. For RT analyses, trials with RTs less than 300ms or more than 3 SD about the participant's mean were excluded.

Two measures of individual strategy use were used: 1) Optimal choices, defined as percent of plateau trials in which the individual chose the optimal target (i.e., the target with the fewest distractors. When the variable distractor was red, the optimal choice was blue, and vice versa), and 2) Switch rate, the percent of trials in which the individual switched target colour (i.e., the colour chosen on trial N was different to the colour chosen on trial $N - 1$).

The results for the ACVS were consistent with previous findings [8,9]. We can clearly see from figure 3 that there are individual

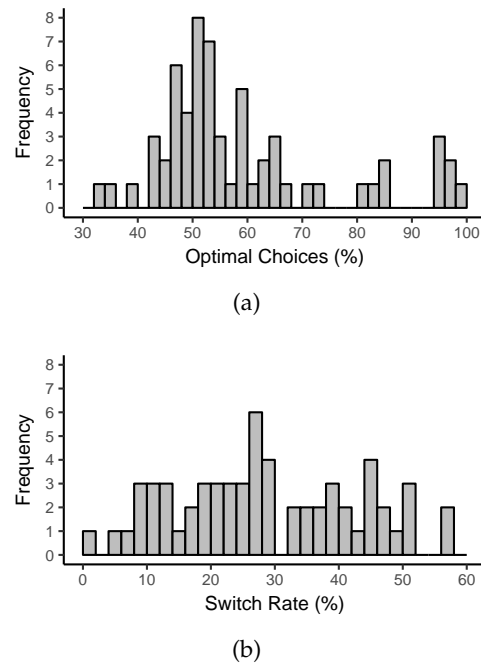


Figure 3. Histograms showing the range of individual differences found in the ACVS task.

differences in both the proportion of optimal targets reported (range 33.62% - 100.00%, $M = 59.15$, $SD = 16.54$) and switch rate (range = 0% - 56.68%, $M = 28.50$, $SD = 14.13$).

(iii) Mouse Click Foraging Task

Only completed, accurate trials were analysed. RTs were defined across the entire trial (i.e., from the start of the trial until the final target was collected). The main measure of interest was average run length per trial. A run was defined as a succession of one or more of the same target type, which was followed and preceded by the other target or no target. The average run length was the average number of target selections in a run. [Five participants did not have feature foraging data for some reason?]

The multiple-target foraging results were also in line with previous findings [11,12]. As expected, run lengths were shorter for feature foraging ($M = 3.16$, $SD = 3.14$) than conjunction foraging ($M = 11.73$, $SD = 7.09$), $t(55) = 8.74$, $p < .001$ (PREFER DIFFERENT STATS HERE? CI [-9.97,-6.26]) suggesting more frequent foraging for multiple targets

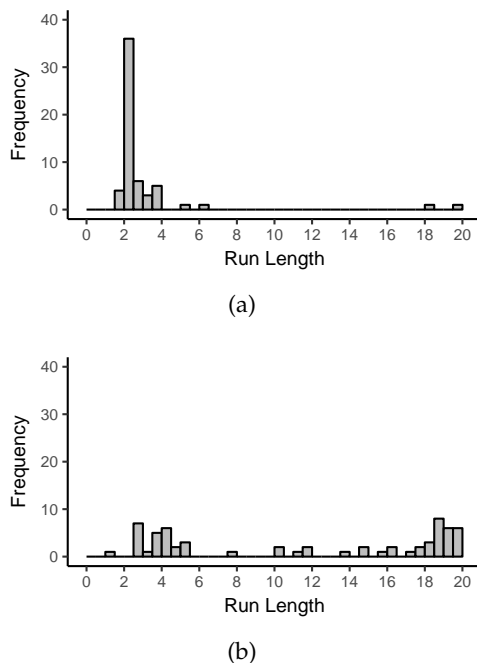


Figure 4. Histogram of mean run lengths for (a) feature and (b) conjunction foraging tasks.

concurrently when those targets were defined by features than by conjunctions. However, both showed significant individual variation: Feature run length varied between 1.87 and 20, and conjunction run length varied between 1.00 and 20. Figure 4 depicts run length across the trials separately for each individual. Twenty “super-searchers” were revealed, identified by short run lengths for both feature and conjunction targets.

(b) Correlations Between Paradigms

The results above demonstrate that we have successfully replicated the previous findings around individual differences in visual search strategy in each of the three paradigms. Furthermore, the split-half paradigm has a test-retest reliability of $r \in [0.65, 0.86]$, similar to that of the ACVS task ($r \in [0.72, 0.9]$), suggesting that the strategies different individuals chose to use in each of these tasks are somewhat stable. Given this, we now investigate the extent to which an individual’s performance in one of the tasks tells us about how well they will do in the other two.

The feature search condition from the MCFT will be discarded from this analysis, as the individual differences of interest relate to how people search in the conjunction condition. Measures for the SHLS task will be taken from

The results presented in Figure 5 convincingly demonstrate that there are no strong correlations between any of the strategies chosen by observers in SHLS, ACVS and MCFT tasks⁴. While there is no indication of a relationship between strategies across the different paradigms, perhaps there is still a correlation between the reaction times due to some participants having higher levels of motivation than others, or stable speed-accuracy trade-offs between the different tasks. Differences relating to their information-processing speed and their ability to ignore distractors would also play in to such a correlation. However, as with search strategy, we again find all the correlations to be weak, typically $0.27 < r < 0.30$ but positive (Figure 5). None of these correlations are statistically significant ($p > 0.05$). Even if we optimistically take all the data together as suggesting a robust correlation in reaction times from paradigm to paradigm, the upper bound of $r = 0.30$ means that this correlation only accounts for at best 10%⁵ an individual’s performance.

4. Discussion

While we have successfully replicated the range of individual differences found in each of these studies, with a larger sample size than the original papers, the between-paradigm correlations give at most $R^2 = 0.1$, and this is a generous interpretation that would fail to pass the usual criteria for null hypothesis significance testing. Knowing how one person will behave in one of these paradigms apparently tells us very little about how they will perform in the others. This can be contrasted with the relatively high test-retest correlations of all three of the tasks individually, which range from $[x \text{ to } x]$.

I think he sees it as a positive thing that they tasks don’t correlate, because it suggests they capture unique variation in behaviour glass half full approach

⁴More comparisons are presented in the *supplementary materials*.

⁵i.e., $R^2 = 0.3^2 = 0.09$

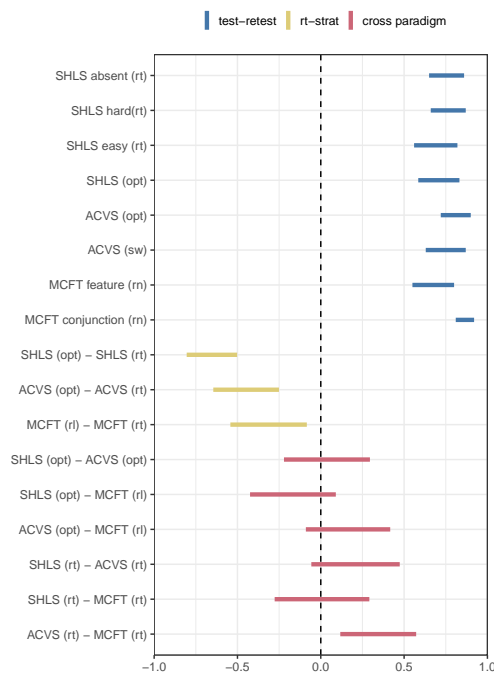


Figure 5. Lots of correlations!

Mention [23]. From [18] - Vogel and colleagues have found that an individual's ability to remember a greater number of items using working memory is related to a filtering capacity in visual search that suppresses attentional capture by distracting visual information [10,11,12]. Behavioural work in visual search has extended this research to demonstrate that working memory correlates only with top-down visual search performance where task relevance is crucial, but not with bottom-up visual search tasks where salience in the environment guides attention [20]

Mention [21].

References

- Walter R Boot, Arthur F Kramer, Ensar Becic, Douglas A Wiegmann, and Tate Kubose. Detecting transient changes in dynamic displays: The more you look, the less you see. *Human Factors*, 48(4):759–773, 2006.
- David H Brainard. The psychophysics toolbox. *Spatial vision*, 10:433–436, 1997.
- Alasdair DF Clarke, Patrick Green, Mike J Chantler, and Amelia R Hunt. Human search for a target on a textured background is consistent with a stochastic model. *Journal of vision*, 16(7):4–4, 2016.
- Frans W Cornelissen, Enno M Peters, and John Palmer. The eyelink toolbox: eye tracking with matlab and the psychophysics toolbox. *Behavior Research Methods, Instruments, & Computers*, 34(4):613–617, 2002.
- Marian Dawkins. Shifts of 'attention' in chicks during feeding. *Animal Behaviour*, 19(3):575 – 582, 1971.
- Iain D Gilchrist and Monika Harvey. Evidence for a systematic component within scan paths in visual search. *Visual Cognition*, 14(4-8):704–715, 2006.
- Mirjam Hartkamp and Ian M Thornton. Meditation, cognitive flexibility and well-being. *Journal of Cognitive Enhancement*, 1(2):182–196, 2017.
- Jessica L Irons and Andrew B Leber. Choosing attentional control settings in a dynamically changing environment. *Attention, Perception, & Psychophysics*, 2016.
- Jessica L Irons and Andrew B Leber. Characterizing individual variation in the strategic use of attentional control. *Journal of Experimental Psychology: Human Perception and Performance*, pages 1–18, 2018.
- Ómar I. Jóhannesson, Árni Kristjánsson, and Ian M. Thornton. Are foraging patterns in humans related to working memory and inhibitory control? *Japanese Psychological Research*, 59(2):152–166.
- Ómar I Jóhannesson, Ian M Thornton, Irene J Smith, Andrey Chetverikov, and Árni Kristjánsson. Visual foraging with fingers and eye gaze. *i-Perception*, 7(2):2041669516637279, 2016.
- Árni Kristjánsson, Ómar I Jóhannesson, and Ian M Thornton. Common attentional constraints in visual foraging. *PloS one*, 9(6):e100752, 2014.
- Norman H Mackworth. The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, 1(1):6–21, 1948.
- Jiri Najemnik and Wilson S Geisler. Eye movement statistics in humans are consistent with an optimal search strategy. *Journal of Vision*, 8(3):4–4, 2008.
- Anna Nowakowska, Alasdair D F Clarke, and Amelia R Hunt.

Human visual search behaviour is far from ideal.

Proceedings of the Royal Society B: Biological Sciences, 2017.

16. Anna Nowakowska, Alasdair DF Clarke, Arash Sahraie, and Amelia R Hunt.
Inefficient search strategies in simulated hemianopia.
Journal of Experimental Psychology: Human Perception and Performance, 42(11):1858, 2016.
17. Anna Nowakowska, Alasdair D.F. Clarke, Arash Sahraie, and Amelia R. Hunt.
Practice-related changes in eye movement strategy in healthy adults with simulated hemianopia.
Neuropsychologia, 2018.
18. Michael J Proulx.
Individual differences and metacognitive knowledge of visual search strategy.
PLoS One, 6(10):e27043, 2011.
19. Richard Russell, Brad Duchaine, and Ken Nakayama.
Super-recognizers: People with extraordinary face recognition ability.
Psychonomic bulletin & review, 16(2):252–257, 2009.
20. Kenith V Sobel, Matthew P Gerrie, Bradley J Poole, and Michael J Kane.
Individual differences in working memory capacity and visual search: The roles of top-down and bottom-up processing.
Psychonomic Bulletin & Review, 14(5):840–845, 2007.
21. Gijsbert Stoet.
Sex differences in search and gathering skills.
Evolution and Human Behavior, 32(6):416–422, 2011.
22. Anne M Treisman and Garry Gelade.
A feature-integration theory of attention.
Cognitive psychology, 12(1):97–136, 1980.
23. Edward K Vogel and Edward Awh.
How to exploit diversity for scientific gain: Using individual differences to constrain cognitive theory.
Current Directions in Psychological Science, 17(2):171–176, 2008.
24. Jeremy Wolfe, Matthew Cain, Krista Ehinger, and Trafton Drew.
Guided search 5.0: Meeting the challenge of hybrid search and multiple-target foraging.
Journal of vision, 15(12):1106–1106, 2015.