View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Northumbria Research Link

Northumbria Research Link

Citation: Rajsic, Jason, Wilson, Daryl E. and Pratt, Jay (2018) The price of information: Increased inspection costs reduce the confirmation bias in visual search. The Quarterly Journal of Experimental Psychology, 71 (4). pp. 832-849. ISSN 1747-0218

Published by: SAGE

URL: http://dx.doi.org/10.1080/17470218.2016.1278249 <http://dx.doi.org/10.1080/17470218.2016.1278249>

This version was downloaded from Northumbria Research Link: http://nrl.northumbria.ac.uk/41671/

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: http://nrl.northumbria.ac.uk/policies.html

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

CUSTOMER SERVICE EXCELLENCE® CSE UniversityLibrary

Northumbria
University
NEWCASTLE

The Price of Information: Increased Inspection Costs Reduce the Confirmation Bias in Visual Search

Jason Rajsic[1], Daryl E. Wilson[2] & Jay Pratt[1]

University of Toronto[1]

Queen's Unversity[2]

The final publication is available via Taylor and Francis, in the Quarterly Journal of Experimental Psychology, DOI: 10.1080/17470218.2016.1278249

Word Count: 11576

Abstract

In visual search, there is a confirmation bias such that attention is biased towards stimuli that match a target template, which has been attributed to covert costs of updating the templates that guide search (Rajsic, Wilson, & Pratt, 2015). In order to provide direct evidence for this speculation, the present study increased the cost of inspections in search by using gaze- and mouse-contingent searches, which restrict the manner in which information in search displays can be accrued, and incur additional motor costs (in the case of mouse-contingent searches). In a fourth experiment, we rhythmically mask elements in the search display to induce temporal inspection costs. Our results indicated that confirmation bias is indeed attenuated when inspection costs are increased. We conclude that confirmation bias results from the low-cost strategy of matching information to a single, concrete visual template, and that more sophisticated guidance strategies will be used when sufficiently beneficial. This demonstrates that search guidance itself comes at a cost, and that the form of guidance adopted in a given search depends on a comparison between guidance costs and the expected benefits of their implementation.

The Price of Information: Increased Search Costs Reduce the Confirmation Bias in Visual Search

In many situations, visual perception feels rapid and effortless, with decisions about how to resolve perceptual ambiguities and prioritize information taken care of by automated processes (Gregory, 1997). Often, however, we require visual information that pertains to one particular proposition (e.g., whether there are unread e-mails in my inbox). In these cases, we engage in a visual search to find target stimuli (e.g., unread email icons), and visual information processing becomes guided by top-down control (Wolfe, Cave, & Franzel, 1989). This guidance steers the inspection of stimuli towards those that are visually similar to the target. One consequence of this guidance is that the information that could be provided by visually dissimilar stimuli will be less likely to reach awareness. In a recent study, Rajsic, Wilson, & Pratt (2015) have shown that this guidance can indeed lead to a confirmation bias (Klayman, 1995; Nickerson, 1998), where observers perseverate in searching for a template-matching target when more efficient strategies are available. In this paper, we investigate a possible cause of this perseveration: the relative costs and benefits of conducting a visual search versus the planning of a visual search. First, we review how it is that a confirmation bias might occur in visual search.

Confirmation bias is a broadly used term that describes biases in both the selection and evaluation of information (Nickerson, 1998; Mackenzie, 2004). While on the surface confirmation bias is problematic, a tendency to seek positive information (positive testing) has been shown to be a reasonable approach to hypothesis testing under a range of conditions thought to characterize real-world situations (Klayman & Ha, 1987; Oaksford & Chater, 1994). This is because the number of “positive” claims made by a hypothesis (i.e., the set of events that it claims should occur) is usually smaller than the number of “negative” claims made by a hypothesis. For instance, the hypothesis “if it is a cat, then it meows” can be evaluated more efficiently by inspecting cats to see if they meow than by inspecting things that don’t meow to see if they aren’t cats. Both types of information searches can falsify the hypothesis, but one – the former, positive testing approach – will likely entail fewer tests (as there are probably fewer things that are cats than there are things that don’t meow). Indeed, positive testing does not necessarily lead to a confirmation bias, but typically does when combined with neglect of potentially unique falsifying information that negative tests provide (e.g., if it were the case that all animals meow, one could not arrive at this correct hypothesis through positive tests alone). However, it should be noted that there is no single explanation of confirmation bias, and biases are likely to occur due to the combination of several factors (Klayman, 1995; Mackenzie, 2004). In this article, we focus on the biased selection of information that occurs when individuals focus on one of several possible hypotheses, where confirmation biases manifest as a temporal bias towards confirmation (that is, faster confirmation than disconfirmation). To study the impact of focal hypotheses on information selection, we use an instruction-based framing manipulation that renders one possible percept more salient. More specifically stated, this study uses visual search to study attention to stimuli during visual hypothesis testing.

Following theorists in the decision-making and memory literatures (Mynatt, Doherty, & Dragan, 1993; Thomas, Dougherty, & Buttaccio, 2014), we have claimed that the confirmation bias in visual search stems from limitations in top-down guidance of attention (Rajsic et al., 2015). Guided visual searches can be considered a series of visual hypothesis tests, and we consider a visual template to be a sort of visual hypothesis that can be confirmed or falsified. When a template is used to guide visual attention, stimuli that match this template are prioritized for inspection. The prioritization of template matching stimuli leads to a confirmatory search, because these sorts of searches will terminate earlier when the hypothesis is true (i.e., when the display contains a target that matches the template). Template-based guidance is a feature of many models of visual attention, such as Guided Search (Wolfe, Cave, & Franzel, 1989; Wolfe, 2007), Theory of Visual Attention (Bundsen, 1990; Bundesen, 1998), the Target Acquisition Model (Zelinsky, 2008), and the Biased Competition Model (Desimone & Duncan, 1995), each of which describes mechanisms by which a template can shape search. Importantly, we do not believe that template-driven guidance is the only source of prioritization in search but rather that such prioritization coexists with other sources of guidance, such as physical salience, selection history, reward (Awh, Belopolsky, & Theeuwes, 2012), global scene properties (such as the category and spatial structure of a scene, and feature statistics; Wolfe, Võ, Evans, & Greene, 2011), and guidance from long-term memory (Fan & Turk-Browne, 2015). Further, we hypothesize that searches will be biased when cognitive limitations prevent multiple hypotheses from being tested in parallel. As a starting point, we have shown that for unfamiliar targets and search contexts, only one template will be used to guide search at one time (Rajsic et al., 2015). This fits with similar claims for the capacity of top-down guidance in search (Olivers, Peters, Houtkamp, & Roelfsema, 2011) as well as for the capacity for evaluation of hypotheses (Mynatt, Doherty, & Dragan, 1993). Indeed, Buttaccio, Lange, Thomas and Dougherty, (2015) have suggested that search is guided by the first visual hypothesis (i.e., template) that is generated from memory. We note, however, that the issue of the capacity of guidance is contentious (see Beck, Hollingworth, & Luck, 2012; Stroud, Menneer, Cave, & Donnelly, 2012; Barrett & Zokay, 2014) and remains unresolved.

To measure the presence of a confirmation bias in visual search, we developed a search task that isolated the tendency to preferentially attend to stimuli because of their confirmatory properties (Rajsic et al., 2015). In typical visual search tasks that use target-present and target-absent trials, the former should be confirmatory because search can be terminated early upon the detection of a present target while the later should be exhaustive. In our task, targets are always present, but on different trials, they may or may not match a positive target template, as set out in search instructions. Hence, in this paradigm, it is useful to distinguish between *targets* – stimuli that possess the response-defining features – and *templates*, which are features, or collections of features, that are used to guide search towards a particular target, or type of target. Importantly, when multiple varieties of targets can occur in a search, a template might specify one particular target, and not another. Critically, in our task, an observer’s decision to adopt a particular template can be attributed solely to the task-framing set out in the instructions, and not to performance-based incentives (i.e., valid cues to the target’s identity or location).

In the task that we have used (Rajsic et al., 2015), one target is always present in a display, and it may be either the Template Matching target, or a Template Mismatching target. Templates for search are elicited using search instructions that ask participants to execute one type of response when a particular target is present, and execute another response if that particular target is not present. For example, as depicted in Figure 1, a participant might be instructed to respond with a left key-press if the target P is green, and respond with a right key-press if the target P is not green. By phrasing the instructions in this way, we establish green P’s as Template Matching targets, and red P’s as Template Mismatching targets. For each subsequent search, overall set size is constant, but Template Matching Subset Size varies. In the example shown in Figure 1, the Matching Subset and Mismatching Subset are of equal size: four stimuli each. Varying the subset size allows us to track the relative prioritization of each stimulus type, based on the logic that search times are proportional to the attended subset (Bacon & Egeth, 1994; Sobel & Cave, 2002).

Our search task has revealed that, indeed, search response times monotonically increased as a function of the Template Matching Subset Size, indicating that participants possessed a confirmation bias of searching the Template Matching colour (Rajsic et al., 2015). Further experiments ruled out explanations attributing the confirmation bias to the need to maintain a template across trials, the need to switch templates between blocks, and a failure to grasp the more economical strategy of searching the smaller subset. Instead, the bias towards stimuli that would confirm the goal proposition was attributed to a preference to search by matching visual input to target template and to avoid the covert cognitive costs of updating templates on a given trial (for evidence that participants prefer to avoid cognitively costly operations, see Kool, McGuire, Rosen, & Botvinick, 2011). Previous estimates of the time required to update a template suggest that updating takes at least 200ms (Vickery, King, & Jiang, 2005; Dombrowe, Donk, & Olivers, 2011), by which time at least one item could have been overtly inspected, and possibly more could have been covertly inspected (Liversedge & Findlay, 2000). Further time would be required to process the colour statistics of the display to determine the appropriate template. Rajsic et al., however, did not directly test the cost-benefit account of confirmatory searching.

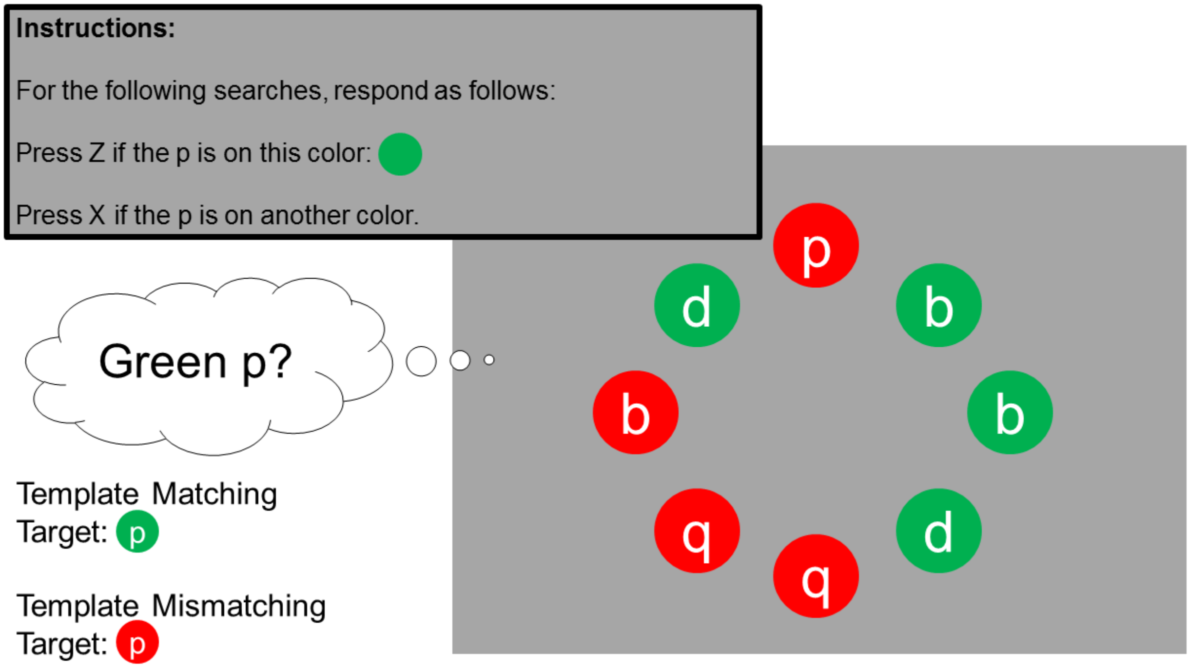

Figure 1. A schematic of the search instructions and displays used in Experiment 1. The instructions before each search block, pictured in the upper left, specified the stimulus-response mapping for a block of 24 trials. The supposed Target Template, expressed as a proposition that may be answered in the affirmative or negative, is pictured in the thought bubble in the middle-left portion. Template Matching and Template Mismatching Targets are pictured in the bottom-left, for this set of exemplified instructions. On the right is a sample search display, with a Matching Subset Size of 4, and a Template Mismatching Target.

In the present paper, we directly examined the cost-benefit account of confirmatory searching by reducing the relative costs of template updating (or, of switching to a strategy of falsification, in hypothesis testing terms). Although it is not possible to reduce the cognitive costs associated with trial-to-trial template decisions, it is possible to add costs to search so that cognitive costs are relatively lessened. To reduce the relative costs of template updating, we measured participants' search behaviour in a task where the costs associated with inspecting stimuli in search are higher than standard visual searches. In a typical search, individual search stimuli (i.e., targets and distractors) are inspected by some combination of overt and covert shifts of attention, and so the inspection costs in such searches would be the corresponding costs of these shifts. In the present study, we measured searches in three experiments that varied the dynamics of inspections used to scan search displays. Experiment 1 replicated the confirmation bias finding with the stimulus modifications necessary for subsequent experiments (including eye tracking), showing again that searches are biased towards stimuli matching a template, and uncovering the oculomotor correlates of this effect. Experiment 2 used a gaze-contingent search task, eliminating the contribution of covert shifts of attention to search, arguably the quickest and cheapest method of visual data acquisition. In Experiment 3, we used a mouse-contingent search task, where inspections required limb movements by having the presence of target-defining features on a given stimulus be contingent on mouse cursor position. Such movements require a host of additional costs, including the recruitment of larger muscle groups, increased degrees of freedom during movement, longer efferent delays, and muscle contraction times. This experiment further increased the costs of acquiring visual information. We predicted that, as inspection costs increased from Experiments 1 to 3, we would observe a complimentary reduction in the confirmation bias in visual search. In Experiment 4, we address a possible alternative explanation for changes in search strategy due to the additional inspection times associated with the manipulations in the first three experiments.

Experiment 1

Our goal for Experiment 1 was to replicate the design of Rajsic et al. (2015) with the addition of eye-tracking, and with the slightly modified stimuli that were to be used in Experiment 2's gaze-contingent searches. On each trial, participants reported whether a given letter was on a given coloured disc, or not. Trials where the letter was on the given coloured disc are referred to as Template Matching Target trials, and trials where the letter is on a disc of the other colour used in a block are referred to as Template Mismatching Target trials. Trials also varied in the number of each coloured disc that were present. All trials contained eight search stimuli (coloured discs with superimposed letters), but any given trial could have two, four, or six Template Matching stimuli, with respect to their colour. The design of this experiment was identical to that of Experiment 1 in Rajsic et al. (2015), with the exception that search stimuli were letters on coloured discs, instead of the letters themselves being coloured. In terms of search time, we expected to replicate our previous finding of an increasing, monotonic relationship between the Template Matching Subset size and search time, paired with an overall cost to search time when the target appeared in the Template Mismatching colour. In terms of oculomotor performance, we expected to find that more saccades would be made to Template Matching stimuli, especially early in search.

Methods

Participants. Twelve undergraduate students from the University of Toronto participated in this study for course credit. All participants provided informed consent prior to participation.

Stimuli. The stimuli and procedure from this experiment were very similar to those reported in Rajsic et al. (2015). All stimuli were generated using Matlab by Mathworks and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli, 2007). Stimuli for each trial consisted of circularly arranged stimulus arrays. These search arrays were drawn with a white fixation mark, 0.8° visual angle, in the centre of the screen. Search stimuli were coloured circles, 2° visual angle in diameter, positioned 8° visual angle from the fixation cross, at eight positions on the circumference of an imaginary circle, separated by 45° of arc. On each stimulus, a letter – one of p, q, d, or b, in lowercase – was drawn in white. The particular circle colours varied by condition, described in the procedure. The set of colours used was purple, yellow, green, orange, pink, blue, and red (RGB values: 200, 0, 255; 200, 200, 0; 0, 255, 0; 255, 128, 0; 255, 128, 255; 50, 50, 255; 255, 50, 50).

Procedure. Each experimental session consisted of 288 trials, broken into 12 blocks of 24. At the outset of each block, participants were presented with an instruction that defined the Target Template for that block. Two stimulus colours were randomly selected from the total set of colours, and of those two colours, one was randomly selected as the Template Colour for that trial. The Template was defined by wording the instructions as can be seen in Figure 1. In the example provided, the Target would be a p, and the Template Colour would be green. The keys (Z and X) corresponding to each response type (detection of a Template Matching Target, and detection of a Template Mismatching Target), were randomly assigned in each block.

Trials within each block belonged to one of six conditions, with presentation randomized at the trial-to-trial level. These six conditions were given by a 3 x 2 factorial design, with the factors of proportion Template Matching Stimuli (referred to for brevity simply as Matching Subset Size) with the levels of 2, 4, and 6; and Target Colour, with the levels of Template

Matching Colour and Template Mismatching Colour. Search displays remained onscreen until a response was given, at which point the search stimuli were removed from the screen, and response feedback was given, in the form of the word “Correct” or word “Incorrect” printed in the centre of the screen. The next trial began following a drift check, where correspondence between the predicted and actual values from the eye tracker were confirmed with a key press, initiated by the participant.

While participants completed the search tasks, eye positions were recorded using the S-R Eyelink 1000 desktop eyetracker. Before each experiment, participants were calibrated using a 9-point calibration routine, and drift-checks were performed before every trial. If the trial could not be initiated, due to poor correspondence between actual and predicted values in the drift check, the experimenter performed another 9-point calibration routine to recalibrate.

At the end of the experimental session, we assessed participants’ self-reported selection strategies using a brief questionnaire. Participants were first asked which colour, if any, they searched first in an open-ended manner. The next question included a hypothetical template instruction (“Press X if the P is on a blue circle, Press Z if the P is on a yellow circle”), and participants were shown a sample display with a Mismatching Subset Size of 2. Participants were asked to indicate the circle they would inspect first. The final two questions asked whether participants used the strategy they had described above for the entire session, or whether they had developed it, and – if they had switched strategies – what their initial strategy was. Responses to these questionnaires were used to classify search strategies as confirmatory search or minimal search using the answer to the second question.

Results and Discussion

Overall, the results of Experiment 1 show that both search RT and number of fixations increased with Template-Matching Subset Size, showing confirmatory search. Three additional findings also emerged. First, despite having an overall bias towards fixating Template-Matching stimuli, this bias decreased with Template-Matching Subset Size. Second, first-fixation durations towards Template Matching stimuli tended to actually be longer than towards Template Mismatching stimuli. Third, searches were more often terminated without fixating the target when targets were Template-Mismatching, suggesting that searchers indeed tend to preferentially search Template Matching subsets.

We first analysed median correct search times to assess whether search exhibited a confirmation bias. These search times are depicted in Figure 2a. Search RT overall increased with Template Matching Subset Size, linear contrast: $F(1, 11) = 18.93$, $MSE = 1.66$, $p = .001$, $\eta^2 = 0.15$[1], although the increase was not entirely linear, as suggested by a marginal quadratic trend, $F(1, 11) = 4.04$, $MSE = 0.08$, $p = .07$, $\eta^2 = 0.01$. Overall, searches were also faster when the Target Colour matched the template than when it did not, $F(1, 11) = 39.66$, $MSE = 1.68$, $p < .001$, $\eta^2 = 0.15$. In addition, the overall accuracy was high, $M = 93.1\%$, $SE = 1.4\%$, and did not differ by condition, *F*s < 1.47, *p*s > .25. These data, then, replicated the results of Experiments 1-4 in Rajsic et al. (2015) in showing a confirmation bias in visual search.

---

[1] Here, as elsewhere in the paper, effect sizes are reported as η^2 values, rather than the partial η^2 values typically reported in repeated measures designs.

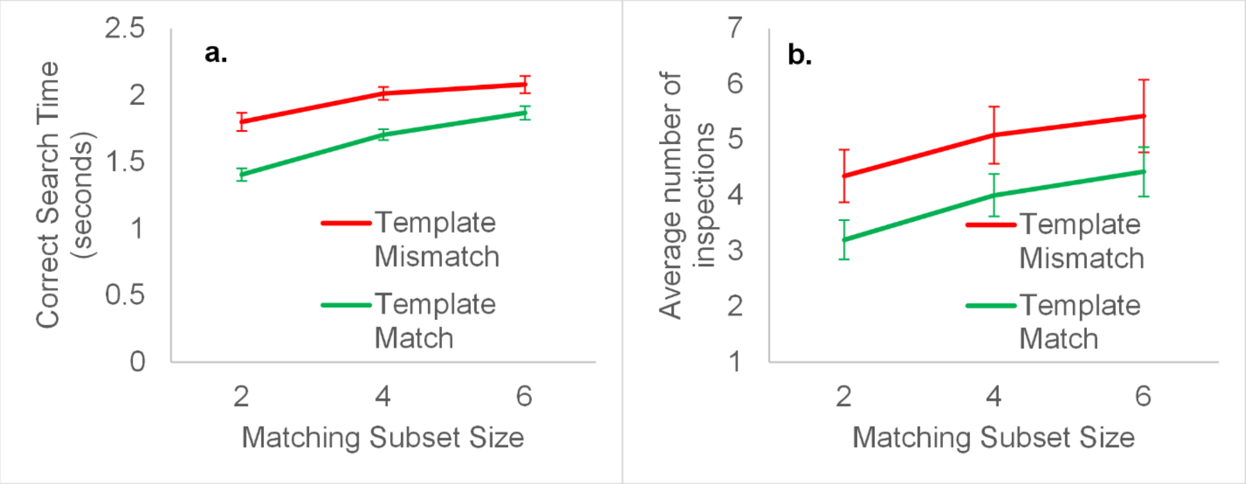

Figure 2. Panel A depicts median correct search times in Experiment 1, and Panel B depicts the average number of fixations per search. Error bars depict one within-subjects standard error.

Given that we collected eye movement data in Experiment 1, we took this opportunity to measure the oculomotor basis of confirmatory search through three analyses; a simple analysis of the number of inspections used in each of our six conditions, as well two other analyses: of how biased inspections were towards Template Matching items, and of how often participants used inference (i.e., reporting the target's colour without inspecting it) in their searches. We first analysed the total number of stimulus inspections in each condition. An inspection was defined as any fixation, or set of fixations, occurring within 2.5 degrees of the centre of a search stimulus before a fixation occurred on either another stimulus, or no stimulus. For one participant, gaze data recorded from the eyetracker was lost, and so the following analyses are of the remaining 11 participants' data. The number of fixations per condition are depicted in Figure 2b. As can be seen, the number of fixations per search increased monotonically with Template Matching Subset Size, $F(2, 11) = 37.72$, $MSE = 19.16$, $p < .001$, $\eta^2 = 0.06$. Both linear, $F(1, 10) = 10.08$, $MSE = 14.63$, $p = .01$, $\eta^2 = 0.06$, and quadratic, $F(1, 10) = 3.87$, $MSE = 0.55$, $p = .08$, $\eta^2 = 0.002$, trends were present, and so the effect of Matching Subset Size on number of fixations was decelerating. Fewer fixations were necessary with the Target Colour matched the template, mirror search RT, $F(1, 10) = 9.52$, $MSE =$ $p = .001$, $\eta^2 = 0.08$. This result shows that overt searching was most efficient when the target's presence could be confirmed, and very closely mirrored the search RT data, suggesting that suggesting that confirmatory searching does affect the number of inspections used during search.

We next sought to determine whether selectivity of stimuli may have changed during the search when confirmatory searching was inefficient. To accomplish this, for each Matching Subset Size and Target Colour, the proportion of first stimulus inspections that went to a Template Matching stimulus was determined and compared to the proportion of all other inspections that went to Template Matching stimuli. So that we assessed a bias towards confirmatory stimuli, we first corrected these measured proportions in both Search Epochs (first inspection, and all subsequent inspections) by accounting for the proportion of stimulus inspections that would be expected by chance given the display. Thus, we used a guessing correction of $p(Bias) = \frac{p(Obs)-p(Chance)}{1-p(Chance)}$, where p(Obs) was the measured probability of inspecting the Template Matching colour and p(Chance) was 0.25, 0.5, and 0.75 for the Matching Subset Sizes 2, 4, and 6, respectively. Importantly, when p(Obs) was below p(Chance), p(Chance) was adjusted to the proportion of Template Mismatching colours in the display. The resulting stimulus inspection tendencies are plotted in Figure 3.

A repeated measures ANOVA on the resulting proportions showed a main effect of Matching Subset Size, $F(2, 20) = 6.47$, $MSE = 0.52$, $p = .007$, $\eta^2 = 0.08$, such that the bias towards Template Matching Stimuli decreased linearly as more Template Matching Stimuli were in a search display, $F(1, 10) = 6.96$, $MSE = 1.01$, $p = .025$, $\eta^2 = 0.08$. That the bias was larger when fewer Template Matching stimuli were present, and smaller when more Template

Matching stimuli were present, is consistent with a contribution of either bottom-up salience, or strategic searching, to stimulus inspections. A main effect of Target Colour was also observed, *F*s(1, 10) = 16.36, *MSE* = 0.85, *p* = .002, η^2 = .07, but was qualified by an interaction with Search Epoch, *F*(1, 10) = 12.58, *MSE* = 0.39, *p* = .005, η^2 = 0.03. Separating analyses by Target Colour revealed that the likelihood of inspecting a Template Matching stimulus only changed between the first inspection and subsequent inspections when the target was in the Template Mismatching Colour, *t*(10) = 3.46, *p* = .006, reflecting the fact that participants – on these trials – likely tended to continue to search until the target had been inspected, thus altering the proportion of fixations to Template Matching stimuli, as the target was itself Template Mismatching in these trials. No difference in stimulus selectivity was present between the first and subsequent stimulus inspections when the target was in the Template Matching colour, *t*(10) = 1.09, *p* = .30.

To complement the selectivity analysis, we also analysed the duration of first inspection on trials where the target was not the first fixated item. This allowed us to obtain a measure of the initial duration of item processing, without contamination from search termination-related processing. A three-way ANOVA including Target Colour, Template Matching Subset Size, and Stimulus Type (Template Matching or Template Mismatching) revealed only a main effect of Stimulus Type, *F*(1, 11) = 8.72, *MSE* = 10247 *p* = .014, η^2 = .01, such that Template Matching Stimuli were inspected for more time, *M* = 221 ms, *SE* = 7 ms, than Template Mismatching Stimuli, *M* = 203 ms, *SE* = 8 ms. All other factors and interactions did not reliably affect first inspection durations, *F*s < 1.87, *p*s < .18, η^2s < .004.

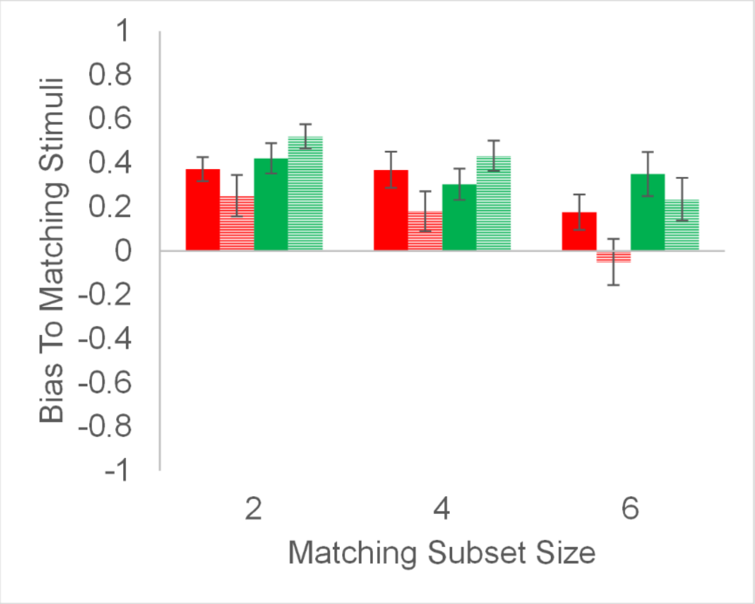

Figure 3. Bias towards Template Matching Stimuli, above (or below) chance, plotted for each Template Matching Subset Size, for Mismatching Colour Targets (red bars) and Matching Colour Targets (green bars) in Experiment 1. Bias for first inspections is plotted with solid bars, and bias for subsequent inspections is plotted as striped bars. Error bars represent 1 SE of the mean.

The preceding analyses demonstrate that searches are controlled by several sources. The change in selectivity caused by Matching Subset Size demonstrates an influence of either task-specific strategy or bottom-up salience on stimulus selection. However, the fact that the overall bias, regardless of magnitude, is towards Template Matching colours in all conditions highlights the contribution of the confirmation bias in visual search.

The change in selection bias that appeared only when targets appeared in the Template Mismatching Colour suggests that participants may have opted to visually confirm the colour of the target stimulus before responding, instead of relying on inference, as inspecting the target on these trials would require at least one Template Mismatching inspection, thus lowering the bias score. This interpretation is bolstered by the finding that inspections after the first show a larger reduction in bias to Template Matching stimuli, as target inspections would naturally come at the end of the search. If searches always ended with a target inspection, this would mean that participants may have opted to conduct a cognitively simpler search, wherein inspections continued until the target stimulus was encountered, even though our task allowed for inference if searches were conducted in a strategic manner. On the other hand, the near chance bias at Matching Subset Size 6 may instead reflect a mixture of biases across trials, such that participants actually switched templates on some trials. To determine the search strategy that participants used, we calculated the proportion of trials where the target was inspected before a correct response was given. We reasoned that, for a given Matching Subset Size, the difference in the probability of target inspections reflects the use of inference. If trials are successfully terminated following a target inspection more often when the target colour matches the target template than when it does not, we can conclude that participants relied on inference to make a response more often in the template mismatching condition, and were more likely to visually inspect the template matching stimuli in the template matching condition. These target inspection data are plotted in Figure 4.

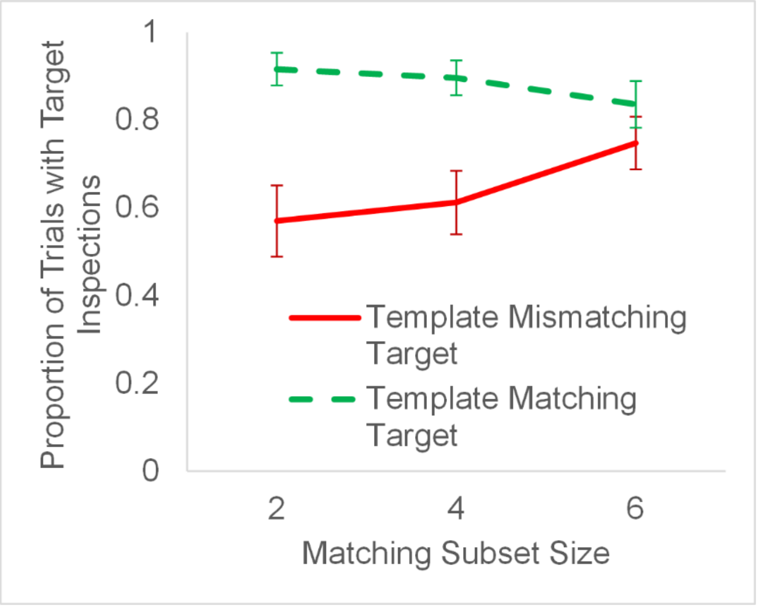

Figure 4. Proportion of trials where targets were fixated before being correctly identified in Experiment 1. Green, dashed line depicts trials with a Template Matching target, and red, solid lines depict trials with a Template Mismatching target. Error bars show one standard error of the mean.

The probability of a target inspection was affected by Target Colour, $F(1, 11) = 10.73$, $MSE = 0.95$, $p = .008$, $\eta^2 = 0.13$, with Target Fixations being overall more likely in the Template Matching Condition, $M_{\text{match}} = .88$, $SE_{\text{match}} = .04$, $M_{\text{mismatch}} = .64$, $SE_{\text{mismatch}} = .07$. This indicates an overall tendency to complete searches by visually confirming the presence of a Template Matching Target, but to report the absence of a Template Matching Target using inference. However, this effect interacted with Matching Subset Size, $F(2, 20) = 8.61$, $MSE = 0.17$, $p = .002$, $\eta^2 = 0.03$.

When the Matching Subset Size was 2, target inspections were more likely when the target matched the template colour, $t(10) = -4.06$, $p = .002$. The same was true of Matching Subset Size 4, $t(10) = -.3.54$, $p = .005$, but not of Matching Subset Size 6, where target inspections were equally likely, $t(10) = 1.09$, $p = .30$. Given that, in the Matching Subset Size 6 condition, target inspections did not reliably differ, and that target inspections occurred often for both Target Colours, it appears that participants did not consistently use colour to guide their search strategy. The variance in which subset (Template Matching or Template Mismatching) is selected is unlikely to be due to individual differences in strategy, as reported strategy (searching matching coloured stimuli or searching the minority colour, included as a Between Subjects factor) did not interact with any Target Fixation effects, $F\text{s} < 1.35$, $p\text{s} > .29$, or Selection Bias effects, $F\text{s} < 2.99$, $p\text{s} > .19$.

While the response time data reported here and in Rajsic et al. (2015) suggests that participants opted to search through the larger, Template Matching Subset Size even when that would incur a search time cost, a detailed look at search behaviour shows a mixture of search strategies. While we observed an overall bias to select stimuli that would confirm the presence of a Target Template, this tendency decreased as the Template Matching Subset Size increased. Furthermore, analyses of inference in search suggested that participants occasionally switched to a disconfirmation strategy when this was economical. Such evidence for a mixture of search strategies would account for the small quadratic trend in search slopes found in this experiment, as well as in our previous experiments (Rajsic et al., 2015). The confirmation bias, then, is stochastic; it is reduced when inefficient, but not reliably. This may be due to a relative increase in the salience of information that matches a target template, which must be overcome using acquired knowledge of the task-specific strategy in those trials where confirmatory searching would entail a longer search.

Experiment 2

The next step in determining whether confirmation bias results from a cognitive cost-benefit trade-off was to measure search when stimulus presentation was gaze-contingent. In this experiment, participants were still presented with coloured circles constituting to-be-search stimuli, but the critical target features – the letters superimposed upon the circles – were not presented unless a given stimulus was foveated. By making information accrual in search contingent on eye-position, we reduce some of the avenues available to search (namely, covert shifts of attention to peripheral and peri-foveal portions of the visual field). This is expected to increase the relative costs of inspections and template updating, and so we predicted a shift towards more strategic, and less confirmatory, searching.

Method

Participants. As in Experiment 1, 12 participants completed the experiment as partial completion of course credit. Participants were enrolled in a first-year Psychology course at the University of Toronto, and provided informed consent before participating.

Stimuli and Procedure. The task, stimuli, and procedure for Experiment 2 were identical to Experiment 1 with the following exception: search stimuli consisted only of coloured circles when not fixated. When participants' gazes fell within 1.5 degrees of the centre of one particular circle, the letter assigned to that stimulus (as in Experiment 1) was drawn on the fixated circle. When participants' gazes left a circle, the letter was removed from it, ensuring that target information was only present when a stimulus was fixated. As in Experiment 1, each participant underwent a calibration procedure prior to completing the experiment, and was recalibrated when a drift correct before each trial indicated poor calibration, in order to ensure accurate recording of eye position.

Results and Discussion

To briefly preview the results of Experiment 2, search RTs, fixation durations, stimulus selectivity, and target inspections all revealed that gaze-contingent searches were more strategic than standard searches. Overall accuracy was again high during the search task, $M = 93.1\%$, $SE = 1.0\%$, and did not differ by condition, $F\text{s} \leq 1.56$, $p\text{s} \geq .23$. These search times are depicted in Figure 5a. A visual inspection reveals that, unlike Experiment 1, the effect of Matching Subset Size, $F(2, 22) = 22.34$, $MSE = 4.95$, $p < .001$, $\eta^2 = 0.18$, was not monotonic. Instead, Matching Subset Size produced a mixture of linear and quadratic trends, $F\text{s} > 14.09$, $p\text{s} < .003$, $\eta^2\text{s} > 0.07$, indicating that participants had adopted the more flexible subset search strategy, choosing to inspect the smaller subset. Searches were faster when Target Colour matched the template, F(1, 11) = 7.44, $MSE = 1.01$, $p = .02$, $\eta^2 = 0.03$, although a marginal interaction was also observed, $F(2, 22) = 3.42$, $MSE = 0.49$, $p = .051$, $\eta^2 = 0.02$. Pairwise comparisons revealed that Template Matching Targets were only found faster than Template Mismatching Targets at Matching Subset Size 2, $t(11) = 3.97$, $p = .002$, $M_{\text{match}} = 2002\text{ms}$, $SE_{\text{match}} = 135\text{ms}$, $M_{\text{mismatch}} = 2402\text{ms}$, $SE_{\text{mismatch}} = 189\text{ms}$. At Matching Subset Size 4, a marginal difference between Target Colours existed, $t(11) = 2.12$, $p = .058$, $M_{\text{match}} = 2720\text{ms}$, $SE_{\text{match}} = 205\text{ms}$, $M_{\text{mismatch}} = 3087\text{ms}$, $SE_{\text{mismatch}} = 161\text{ms}$, but at Matching Subset Size 6, no difference between Target Colours was present, $t(11) = -0.394$, $p = .70$, $M_{\text{match}} = 2627\text{ms}$, $SE_{\text{match}} = 207\text{ms}$, $M_{\text{mismatch}} = 2682\text{ms}$, $SE_{\text{mismatch}} = 207\text{ms}$. An advantage for finding Template Matching targets was present at Matching Subset Sizes 2 and 4, but not at Matching Subset Size 6.

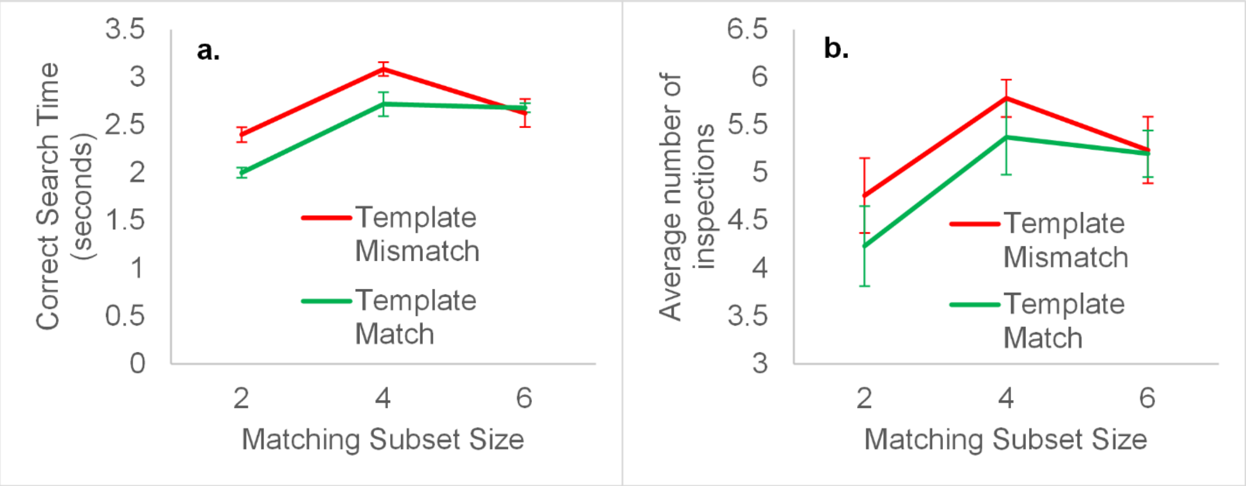

Figure 5. Panel A depicts median correct search times in Experiment 2, and Panel B depicts the average number of fixations per search. Error bars depict one within-subjects standard error.

As in Experiment 1, we measured stimulus inspections (as defined earlier) used in search to uncover how participants went about finding target stimuli. The gaze data for two participants was lost due to a computer error, and so the following analyses are of the remaining ten participants' gaze data. The resulting average number of inspections per condition are depicted in Figure 5b. As with search RT, Matching Subset Size produced a non-monotonic effect indicative of a flexible subset search strategy, $F(2, 18) = 6.93$, $MSE = 6.05$, $p = .006$, $\eta^2 = 0.03$, showing a strong quadratic trend of Matching Subset Size, $F(1, 9) = 11.78$, $MSE = 6.88$, $p = .007$, $\eta^2 = 0.02$, but only a marginal linear trend, $F(1, 9) = 4.49$, $MSE = 5.22$, $p = .063$, $\eta^2 = 0.01$ Fewer inspections were required when Target Colour matched the template, $F(1, 9) = 5.62$, $MSE = 1.60$, $p = .042$, $\eta^2 = 0.004$, and this effect did not interact with Matching Subset Size, $F(2, 18) = 1.03$, $MSE = 0.33$, $p = .38$, $\eta^2 < .001$. The number of inspections used, closely mirrored search RT data, as in Experiment 1.

To assess the selectivity in search, we again calculated the bias towards, or away, from Template-Colour Matching Stimuli for two Search Epochs: first inspections, and subsequent inspections. These scores were corrected for chance, and are plotted in Figure 6.

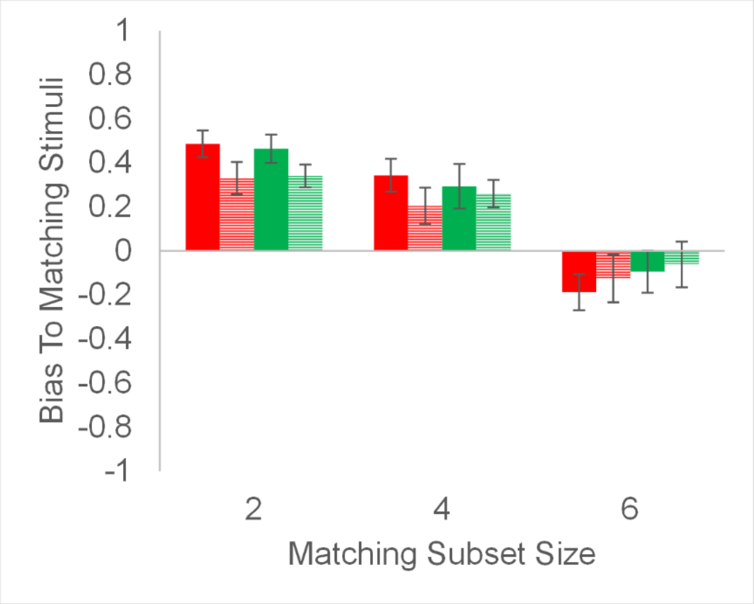

Figure 6. Bias towards Template Matching Stimuli, above (or below) chance, plotted for each Template Matching Subset Size, for Mismatching Colour Targets (red bars) and Matching Colour Targets (green bars) in Experiment 2. Bias for first inspections is plotted with solid bars, and bias for subsequent inspections is plotted as striped bars. Error bars represent 1 SE of the mean.

We observed two influences on selectivity. First, the bias towards Template Matching Colours was affected by Matching Subset Size, $F(2, 18) = 23.23$, $MSE = 2.96$, $p < .001$, $\eta^2 = 0.41$, such that the bias decreased linearly as Matching Subset Size increased, $F(1, 9) = 26.55$, $MSE = 5.47$, $p = .001$, $\eta^2 = 0.38$. A quadratic contrast, $F(1, 9) = 9.33$, $MSE = 0.46$, $p = .014$, $\eta^2 = 0.03$, showed that the change in bias was greater between Subset Sizes 4 and 6; $M_4 = 0.28$, $SE_4 = 0.07$, $M_6 = -0.12$, $SE_6 = 0.09$; than between Subset Sizes 2 and 4, $M_2 = 0.41$, $SE_2 = 0.05$. Second, Search Epoch affected the bias, $F(1, 9) = 7.73$, $MSE = 0.11$, $p = .021$, $\eta^2 = 0.007$, with the bias being overall lower after the first inspection.

Critically, comparing the effect of Matching Subset Size on the selection bias between Experiments 1 and 2 yielded an interaction, $F(2. 38) = 5.44$, $MSE = 0.56$, $p = .008$, $\eta^2 = 0.03$. Independent samples *t*-tests showed that this difference was driven by a reduction in the bias at Matching Subset Size 6 of Experiment 2, $t(19) = 2.43$, $p = .025$, indicating that gaze-contingent searching led to the strategic allocation of attention towards the Mismatching colour stimuli, unlike in Experiment 1. In contrast to Experiment 1 as well, an analysis of first inspection durations of distractors revealed no main effect of Stimulus Type, $F(1, 9) = 2.44$, $MSE = 196251$, $p = .15$, $\eta^2 = .02$, but rather an interaction between Stimulus Type and Template Matching Subset Size, $F(1, 9) = 5.19$, $MSE = 108582$, $p = .017$, $\eta^2 = .02$. Paired samples *t*-tests revealed a reliable difference between Stimulus Types at Matching Subset Size 2, $t(9) = 9.20$, $p < .001$, such that Template Matching Stimuli were inspected longer, $M_{\text{match}} = 597\text{ms}$, $SE_{\text{match}} = 36\text{ms}$, $M_{\text{mismatch}} = 292\text{ms}$, $SE_{\text{mismatch}} = 20\text{ms}$, and a marginal trend in the same direction for Matching Subset Size 4, $t(9) = 1.96$, $p = .08$, $M_{\text{match}} = 450\text{ms}$, $SE_{\text{match}} = 35\text{ms}$, $M_{\text{mismatch}} = 374\text{ms}$, $SE_{\text{mismatch}} = 27\text{ms}$, but no difference at Matching Subset Size 6, $t(9) = 1.13$, $p = .29$, $M_{\text{match}} = 463\text{ms}$, $SE_{\text{match}} = 26\text{ms}$, $M_{\text{mismatch}} = 435\text{ms}$, $SE_{\text{mismatch}} = 20\text{ms}$. Thus, the change in selectivity noted in our bias measurement was complimented by a similar change in inspection durations.

The change in selectivity observed using gaze-contingent windows might simply reflect a longer time spent planning searches, such that participants updated their template on each search as warranted by the distribution of coloured stimuli in the display. However, comparing the time between search onset and first inspections between Experiments 1 and 2 yielded no reliable differences, *F*s < 2.20, *p*s < .17. The first inspection times at Matching Subset Size 6 for

Experiments 1 and 2 were M_{Exp1} = 404ms, SE_{Exp1} = 27ms and M_{Exp2} = 416ms, SE_{Exp2} = 27ms, respectively. If the improved selection strategy seen in Experiment 2 occurs due to longer search planning and template updating, then it would appear that this additional planning only requires approximately 12ms.

Lastly, we again analysed the likelihood of fixating the target stimulus before providing a correct response. These data are plotted in Figure 7. While target fixation probability showed a main effect of Target Colour, $F(1, 9) = 7.69$, $MSE = 0.62$, $p = .02$, $\eta^2 = 0.05$, an interaction was observed, $F(2, 18) = 21.17$, $MSE = 0.84$, $p < .001$, $\eta^2 = 0.15$. Paired comparisons between Target Colours at each Matching Subset Size further supported the conclusion that participants flexibly allocated attention to either the Matching or Mismatching colour stimuli. At Matching Subset Size 2, $t(9) = 6.08$, $p < .001$, the target was fixated more often when it was Template Matching, $M_{\text{match}} = 0.998$, $SE_{\text{match}} = 0.002$, than when it was Template Mismatching, $M_{\text{mismatch}} = 0.43$, $SE_{\text{mismatch}} = 0.09$. This was also true at Matching Subset Size 4, $t(9) = 3.73$, $p = .005$; $M_{\text{match}} = 0.91$, $SE_{\text{match}} = 0.03$, $M_{\text{mismatch}} = 0.62$, $SE_{\text{mismatch}} = 0.08$. At Matching Subset Size 6, however, this difference reversed, $M_{\text{match}} = 0.63$, $SE_{\text{match}} = 0.08$, $M_{\text{mismatch}} = 0.87$, $SE_{\text{mismatch}} = 0.08$, albeit only numerically, $t(9) = 1.84$, $p = .099$.

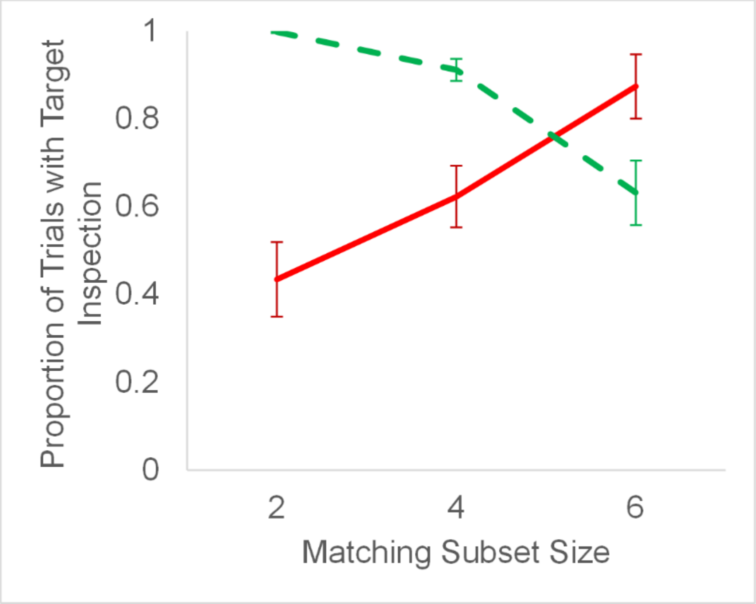

Figure 7. Proportion of trials where a correct response was given and the target was inspected before search termination in Experiment 2. Green, dashed line depicts trials with a Template Matching target, and the red, solid line depicts trials with a Template Mismatching target. Error bars show one standard error of the mean.

In sum, the results from Experiment 2 show that gaze-contingent search reduced the extent of confirmatory searching, as assessed by measurements of search time, average inspections, selectivity, and – to an extent – target fixations. These findings converge on the conclusion that, under search conditions with higher inspection costs, participants were able to prioritize the smaller subset, irrespective of the search proposition, in order to search more effectively. Despite this improvement in prioritization, the confirmation bias was still present in two ways: first, participants had a preference for selecting Template Matching stimuli at Matching Subset Size 4. Second, the bias towards Template Matching Stimuli deviated from chance at Matching Subset Size 2 more than the bias towards Template Mismatching stimuli deviated from chance at Matching Subset Size 6. Overall, however, Experiment 2 suggests confirmation bias can be reduced when the costs of accessing information are increased. In Experiment 3, we provide a stronger test of this proposal by introducing additional inspection costs.

Experiment 3

In order to test whether searches are more efficient when the costs of inspections are increased, we conducted a third experiment where these inspection costs were further increased. In this experiment, we used a mouse-contingent search, reasoning that the additional costs of control over the slower movements would increase incentives to search strategically. Compared to eye movements, arm and hand movements require the recruitment of larger muscles, involve additional degrees of freedom, and suffer larger efferent delays and contraction times. Moreover, there are additional reference frame transformations for mouse cursor control, where the cursor moves in a different spatial plane than the control device. In terms of performance, eye movement times increase less as the index of difficulty (a measure of movement difficulty in terms of speed-accuracy trade-offs) than do cursor movement times (Vertegaal, 2008). Given these additional demands, we expected that the change in guidance seen between Experiments 1 and 2 would be further exaggerated in Experiment 3.

Method

Participants. A new sample of twelve undergraduate students, enrolled in a first-year Psychology course at the University of Toronto, completed this experiment for partial fulfillment of course credit. All participants provided informed consent before participating.

Stimuli and Procedure. Stimuli and procedure were identical to Experiment 2 with the exception that a cursor, controlled by a standard USB computer mouse, was used to control the presence of search stimuli (letters). Given that the cursor was used to inspect the display, gaze positions were not recorded, and no eye tracking was performed.

Results and Discussion

Overall, the results of Experiment 3 mirrored those of Experiment 2; strategic stimulus selection of smaller subsets as revealed by search RTs, number of inspections, color selectivity, and target inspection probability. Comparisons between Experiments 1 and 2, however, revealed that the extent of strategic selection was amplified by using mouse-contingent search. Overall search accuracy was high in Experiment 3, M = 92.8%, SE = 2.1%, but was affected by Matching Subset Size, $F(2, 22) = 7.03$, $MSE = 0.03$, $p = .004$, $\eta^2 = 0.09$, and the combination of Target Colour and Matching Subset Size, $F(2, 22) = 5.71$, $MSE = 0.006$, $p = .01$, $\eta^2 = 0.02$. Accuracy for trials with a Template Matching Subset Size of 6, M = 89.0%, SE = 3.3%, was lower than for other Matching Subset Sizes, M = 94.7, SE = 1.5%, F(1, 11) = 7.79, $p = .018$, partial $\eta^2 = 0.07$, and was lower when the Target appeared in the Template Mismatching Colour, but only at Matching Subset Size 6, M_{match} = 91.9%, SE_{match} = 2.4%, M_{mismatch} = 86.1%, SE_{mismatch} = 4.4%. More response errors were made, overall, on those trials in which confirmatory searching would be most difficult.

Median correct search RTs are depicted in Figure 8a. These search times showed, like Experiment 2, that searches were more strategic. Matching Subset Size, $F(2, 11) = 30.72$, $MSE = 5.72$, $p < .001$, $\eta^2 = 0.33$, had a non-monotonic effect on search, with both a linear, $F(1, 11) = 16.21$, $MSE = 2.92$, $p = .002$, $\eta^2 = 0.09$, and a quadratic, $F(1, 11) = 44.32$, $MSE = 8.51$, $p < .001$, $\eta^2 = 0.25$, trend accounting for the effect. The presence of the quadratic trend indicated that participants again did prioritize the Template Mismatching stimuli when appropriate. A main effect of Target Colour was observed, $F(1, 11) = 7.08$, $MSE = 2.86$, $p = .022$, $\eta^2 = 0.08$, but was accompanied by an interaction, $F(2, 22) = 3.43$, $MSE = 0.46$, $p = .05$, $\eta^2 = 0.02$. We therefore compared the search RT for different Target Colours at each Matching Subset Size. Pairwise comparisons revealed that Template Colour Matching Targets were reported faster than Template Colour Mismatching Targets at Matching Subset Size 2, $t(9) = 2.62$, $p = 0.24$, M_{match} = 1950ms, SE_{match} = 62ms, M_{mismatch} = 2444ms, SE_{match} = 196ms, and Matching Subset Size 4, $t(9) = 3.37$, $p = .006$, M_{match} = 2877ms, SE_{match} = 153ms, M_{mismatch} = 3470ms, SE_{mismatch} = 177ms, but not at Matching Subset Size 6, where no difference was observed, $t(9) = 0.54$, $p = .60$; M_{match} = 2637ms, SE_{match} = 92ms, M_{mismatch} = 2744ms, SE_{mismatch} = 200ms. These results parallel Experiment 2 in demonstrating the emergence of a tendency to prioritize template mismatching stimuli when such stimuli appear in the minority, and could therefore reduce search load.

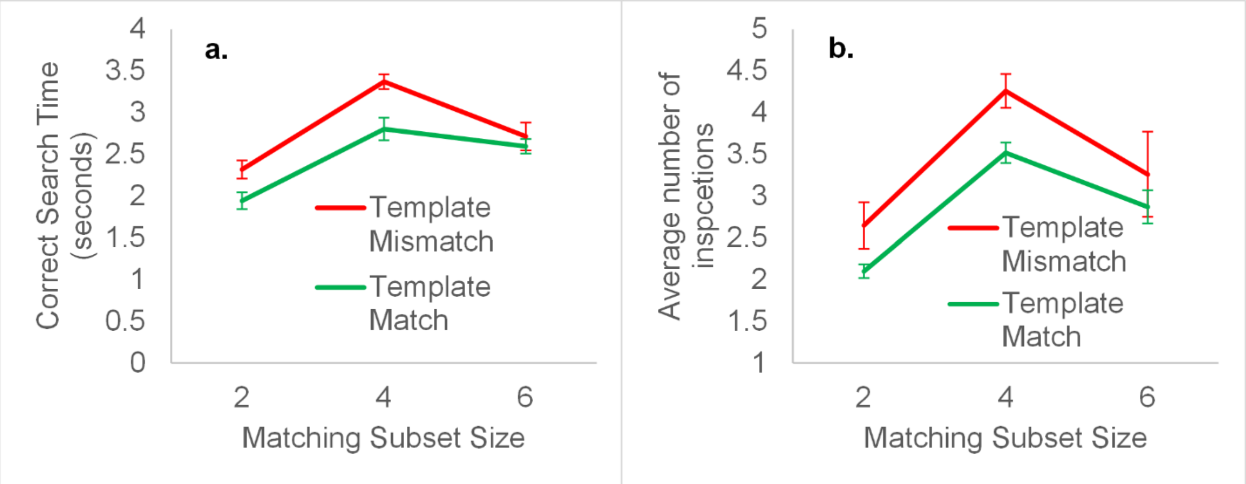

Figure 8. Panel A depicts median correct search times in Experiment 1, and Panel B depicts the average number of fixations per search. Error bars depict one within-subjects standard error.

As with Experiments 1 and 2, we analysed the dynamics of search using three metrics: total inspections, bias towards Template Matching stimuli, and likelihood of target inspection.

For the first metric, we defined an inspection as instances where the cursor was placed over a target stimulus. If the same stimulus was revealed with the cursor as the previous revealed stimulus, this was considered as a single inspection, in order to prevent over-counting by poor cursor control. Unfortunately, inspection durations could not be analysed due to a coding error that resulted the times of each inspection being improperly recorded. The resulting average number of inspections are depicted in Figure 8b.

As with search RT, Matching Subset Size had a non-monotonic effect on the average number of inspections, $F(2, 22) = 19.48$, $MSE = 13.81$, $p < .001$, $\eta^2 = 0.27$, as evidenced by a mixture of a linear, $F(1, 9) = 6.86$, $MSE = 5.72$, $p = .024$, $\eta^2 = 0.06$, and a quadratic, $F(1, 9) = 37.54$, $MSE = 21.88$, $p < .001$, $\eta^2 = 0.21$, trend. The effect of Target Colour did not reach statistical significance, $F(1, 11) = 4.07$, $MSE = 5.64$, $p = .07$, $\eta^2 = 0.05$, and no interaction was observed, $F(2, 22) = 1.23$, $MSE = 0.19$, $p = .31$, $\eta^2 = 0.004$. These results did not differ markedly from those observed in Experiment 2, and show a strategic, rather than confirmatory, search strategy. To provide a direct comparison, however, we included Experiment as a between-subjects factor. This analysis revealed no interactions between the Effector (eye or mouse) and Target Colour, Matching Subset Size, or their interaction, $F\text{s} \leq 0.95$, $p\text{s} \geq .40$, $\eta^2\text{s} \leq 0.003$. However, a main effect of Effector was present, $F(1, 20) = 40.93$, $MSE = 129.60$, $p < .001$, $\eta^2 = 0.13$, with mouse contingent searches requiring fewer overall inspections than gaze contingent searches, $M_{\text{mouse}} = 3.11$, $SE_{\text{mouse}} = 0.21$, $M_{\text{gaze}} = 5.10$, $SE_{\text{gaze}} = 0.23$.

We next analysed the selectivity bias, calculated using inspections, which is plotted in Figure 9. Matching Subset Size affected selectivity, $F(2, 22) = 35.88$, $MSE = 17.43$, $p < .001$, $\eta^2 = 0.59$, such that the bias towards Template Matching Stimuli reduced as the Template Matching Subset Size increased, $F(1, 11) = 39.98$, $MSE = 33.76$, $p < .001$, $\eta^2 = 0.58$. A quadratic trend was also present, $F(1, 11) = 8.71$, $MSE = 1.11$, $p = .013$, $\eta^2 = 0.02$, reflecting a larger drop in confirmatory selection between Subset Size 4, $M = 0.40$, $SE = 0.09$, and Subset Size 6, $M = -0.38$, $SE = 0.16$, than from Subset Size 2, $M = 0.81$, $SE = 0.06$, to Subset Size 4. In addition, Subset Size interacted with Search Epoch (first inspections vs. all other inspections), $F(2, 22) = 13.57$, $MSE = 0.16$, $p < .001$, $\eta^2 = 0.006$. However, a three-way interaction between Search Epoch, Matching Subset Size, and Target Colour was present, $F(2, 22) = 4.91$, $MSE = 0.03$, $p = .017$, $\eta^2 = 0.001$, and so we analysed changes in selectivity by Search Epoch and Target Colour separately for each Subset Size. At Subset Size 2, there was a main effect of Search Epoch, $F(1, 11) = 6.68$, $MSE = 0.20$, $p = .025$, $\eta^2 = 0.\ 07$, and no other effects, Fs ≤ 1.15, ps $\geq .31$, η^2s ≤ 0.01, reflecting a decrease in the bias after the first inspection. However, for Matching Subset Sizes 4 and 6, no changes in selectivity were observed by Search Epoch or Target Colour, Fs ≤ 2.93, ps $\geq .12$, η^2s $\leq .\ 008$. Overall, the most striking result is that colour selectivity was enhanced in the mouse-contingent compared to gaze-contingent search, as evidenced by an interaction between Matching Subset Size and Experiment (2 vs. 3), $F(2, 40) = 7.46$, $MSE = 2.42$, $p = .002$, $\eta^2 = .07$.

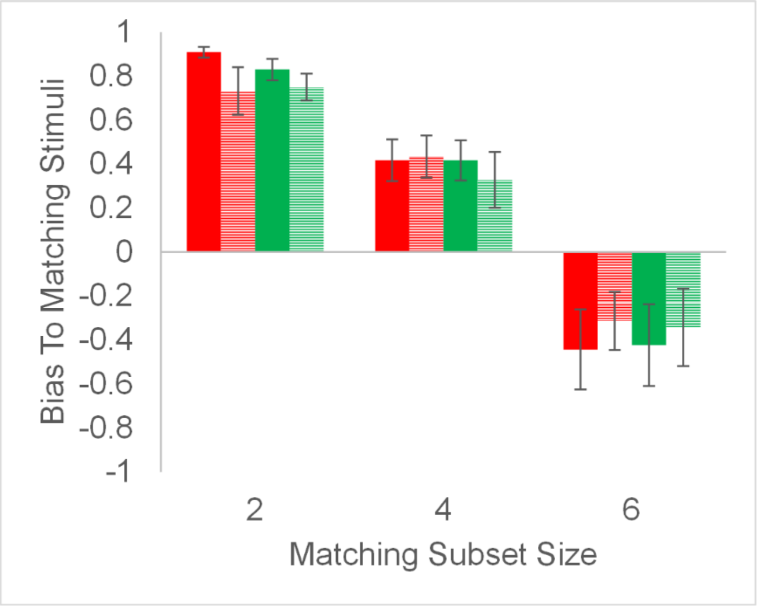

Figure 9. Bias towards Template Matching Stimuli, above (or below) chance, plotted for each Template Matching Subset Size, for Mismatching Colour Targets (red bars) and Matching Colour Targets (green bars) in Experiment 3. Bias for first inspections is plotted with solid bars, and bias for subsequent inspections is plotted as striped bars. Error bars represent 1 SE of the mean.

As a final analysis, we examined the likelihood of correctly completing a search after visually inspecting the target, which is plotted in Figure 10. Main effects of Target Colour, $F(1, 11) = 7.24$, $MSE = 0.88$, $p = .02$, $\eta^2 = 0.08$, and Matching Subset Size, $F(2, 11) = 4.40$, $MSE = .004$, $p = .025$, $\eta^2 < 0.001$, as well as an interaction between Target Colour and Matching Subset Size were observed, $F(2, 22) = 37.53$, $MSE = 2.76$, $p < .001$, $\eta^2 = 0.50$. Comparing target fixation frequency between Target Colours (Template Matching and Template Mismatching) for Matching Subset Sizes revealed a higher probability of fixating the Target on the Template Matching Target trials when the Matching Subset Size was 2 or 4, $ts(11) \leq 3.41$, $ps < .006$, but that this pattern reversed at Matching Subset Size 6, $t(11) = 3.20$, $p = .008$. This indicates that, overall, participants inspected Template Mismatching stimuli first when the Template Matching stimuli were more numerous, and relied on inference to report the presence of a Template Matching Target in these conditions more often than not. In addition, the use of inference was more pronounced in Experiment 3 than in Experiment 2, as indicated by a three-way interaction between Target Colour, Matching Subset Size, and Experiment, $F(2, 40) = 3.37$, $MSE = 0.20$, $p = .044$, $\eta^2 = 0.02$. This supports our speculation that increasing inspection costs, and using limb movements instead of saccades, improved participants' ability to minimize their inspections in search on a trial-to-trial basis.

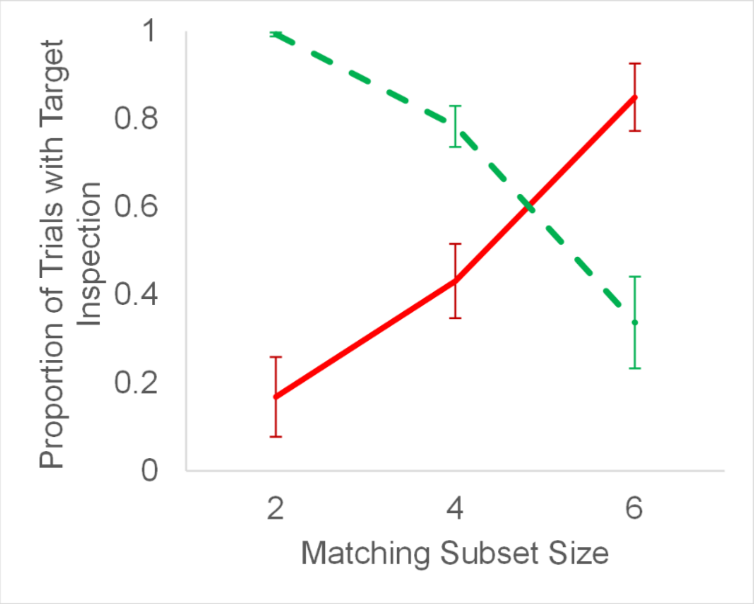

Figure 10. Proportion of trials where a correct response was given and the target was inspected before search termination in Experiment 3. Green, dashed line depicts trials with a Template Matching target, and the red, solid line depicts trials with a Template Mismatching target. Error bars show one standard error of the mean.

Although the results of Experiment 3 show that increases in inspection costs lead to reductions in confirmatory searching, one remaining issue is that, thus far, it is unclear whether it is motor costs, information costs, or simply time costs that underlie the changes in search strategy. In Experiment 2, we used a gaze-contingent search to limit the perceptual information, which we expected to increase the costs of poorly planned search inspections in terms of lost information (from the visual periphery). In Experiment 3, we used a mouse-contingent search to increase the costs in terms of motor control – every inspection required larger limb movements and additional reference frame transformations. However, both of these manipulations also increased the overall time required to acquire information, as can be seen in the average different in RT between the Subset Size 2 and Subset Size 4, Template Present conditions, which reflects the extra time taken to search through two extra items to find the target: M_{Exp1} = 300ms, SE_{Exp1} = 53ms, M_{Exp2} = 718ms, SE_{Exp1} = 133ms, M_{Exp3} = 861ms, SE_{Exp3} = 100ms. In fact, one could argue that no strategy shift occurred at all; if strategic search control, which relies on an analysis of the properties of the display to choose the optimal guiding colour, simply takes longer to emerge than confirmatory search biases within a given trial, the longer inspection times may entirely account for our findings. To test this possibility, a fourth experiment was conducted.

Experiment 4

Experiment 4 tested whether the improvements in search strategy seen thus far can be attributed solely to the time required to plan inspections within a search. To test this, we introduced intermittent masks into the search display, which controlled the amount of time that target-defining information was visible. By doing so, we directly controlled the amount of time available for participants to plan their subsequent inspections within a given search. If improvements in search strategy are not actually strategic but are entirely due to the time taken to plan inspections, then searches displays with high information rates should exhibit confirmatory searching and search displays with low information rates should exhibit strategic searches. Of course, lost time can also be considered an inspection cost, which could lead to sort of shifts in control that would properly be considered a strategy shift. If this were the case, participants who practiced searching with low Information rates would show a transfer of strategic searching to fast Information rate displays, whereas participants who practiced searching with high Information rates may show a transfer of confirmatory searching to slow Information rate display. To test this alternative, we ran two groups of participants through a blocked design experiment, where half of participants searched through low Information Rate displays before switching to high Information Rate displays, and the other half of participants experienced the opposite. If information rate plays a key role in determining the manner of search, we would expect that high Information rate displays would lead to confirmatory searching and low Information rate display would lead to strategic searching.

Method

Participants. Eighteen undergraduate, first year psychology students participated in this experiment in exchange for course credit. All provided informed consent, and were naïve to the purposes of the study.

Stimuli and Procedure. The stimuli were identical to those of Experiment 1 with the following exception. Where in Experiment 1, search stimuli consisted of lowercase letter (p, q, d, and b) printed on top of coloured discs, search stimuli in Experiment 4 were dynamic. Stimuli oscillated between being drawn as individual lowercase letters on top of coloured discs and overlapping lowercase letters drawn on top of coloured discs. These overlapping lowercase letters served as masks, which prevented letter from being recognized during periods of masking. For a given search stimulus, the letter presented on its coloured disc did not change between masking periods.

Two Information Rates were used. High Information Rate trials were those in which search stimuli alternated between 235ms of letter presentation and 65ms of mask presentation. Low Information Rate trials were those in which search stimuli alternated between 235ms of letter presentation and 765ms of mask presentation. A depiction of this method can be seen I Figure 11. Half of participants completed six blocks with High Information Rate trials first, followed by six blocks of Low Information Rate trials first. The other half of participants completed the opposite block order. Participants were assigned to the Information Rate Order conditions in alternating order. Eye position was not monitored in this experiment.

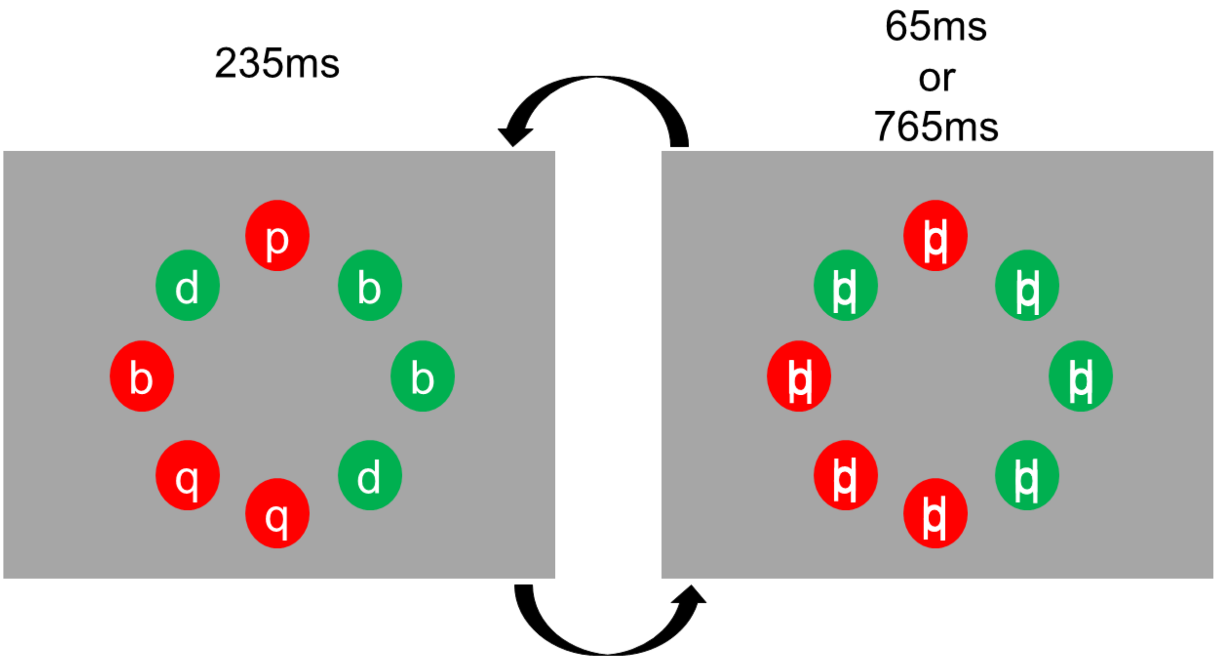

Figure 11. An example illustration of the stimuli and procedure used in Experiment 4. Note that the difference between high and low information rate trials corresponds to the duration of the mask display on the right (these possible durations are shown above the mask display).

Results and Discussion

The overall results of Experiment 4 showed that searches were consistently strategic when the information rate was low, but also showed confirmatory search patterns when information rate was high and when this was the first condition experienced. Interestingly, when high-information rates searches were performed after first experiencing low information rate searches, participants continued to search strategically despite the change in information rate.

Median correct RTs were analysed for three conditions: Matching Subset Size, Target Colour, and Information Rate. As expected, each had a main effect on RT, Fs > 22.44, ps < .001, η^2s > .04. Importantly, the interaction between Information Rate and Matching Subset Size was significant, $F(2, 34) = 8.10$, $MSE = 2.93$, $p = .001$, $\eta^2 = 0.03$. While this supports the possibility that the improved search strategy in Experiments 1, 2 and 3 merely reflect the extra time needed to plan inspections strategically during search, Matching Subset Size was quadratically related to Correct RT for both High Information Rate trials, $F(1, 15) = 5.76$, $MSE = 0.69$, $p < .03$, $\eta^2 = 0.06$, and Low Information Rate, $F(1, 17) = 27.51$, $MSE = 16.71$, $p < .001$, $\eta^2 = 0.20$. Therefore, we analysed search performance for High and Low Information Rate trials with added factor of Information Rate Order.

For High Information Rate trials, Information Rate Order interacted with Matching Subset Size, $F(2, 34) = 8.975$, $MSE = 0.53$, $p =.001$, $\eta^2 = 0.1$. For those who completed High Information Rate trials first, Matching Subset Size affected RT linearly, $F(1, 8) = 43.01$, $MSE = 2.37$, $p < .001$, $\eta^2 = 0.45$, with no quadratic trend, $F(1, 8) = 0.612$, $MSE = 0.01$, $p = .81$, $\eta^2 = 0.002$, showing confirmatory searching. When Low Information Rate trials were experienced first, Matching Subset Size on High Information Rate trials affected RT with both a linear trend, $F(1, 8) = 25.34$, $MSE = 1.25$, $p = .001$, $\eta^2 = 0.21$, and a quadratic trend, $F(1, 8) = 14.59$, $MSE =$

1.66, *p* = .005, $\eta^2 = 0.29$, demonstrating the presence of strategic searching despite identical time available for planning inspections within a trial (see Figure 12). Participants who began the experiment with Low Information Rate trials likely learned to use the distribution of colours to inform their search strategies, given the amount of planning time available within each trial. This practice and strategy development transferred over to performance on later High Information Rate trials, as seen above, where less confirmatory searching occurred. Therefore, it appears that search strategies are indeed sensitive to inspection costs, which, in this case, were opportunity costs – the time used inspecting one stimulus that could have been spent inspecting another.

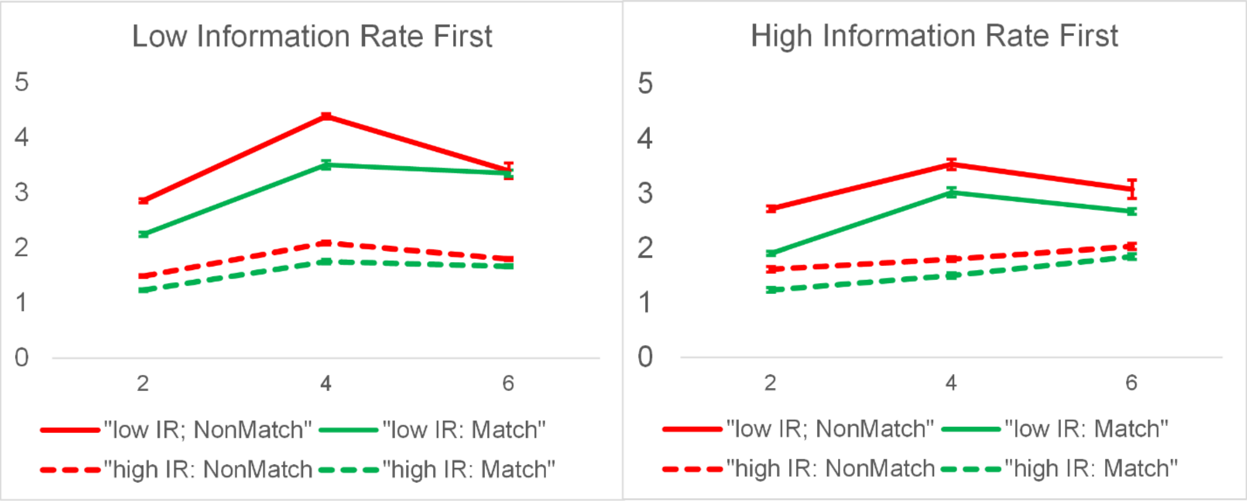

Figure 12. Correct average median search RTs, split by participants who completed Low Information Rate searches first (left) and who completed High Information Rate searches first (right). Red lines depict Template Non-Matching Target trials, and Green lines depict Template Matching Target trials. Solid lines depict Low Information Rate trials and dashed lines depict High Information Rate trials.

General Discussion

Visual search can be viewed as a process of testing whether a particular visual state (the presence or absence of a target) is true or false. Earlier we showed that, in a multiple-target conjunction search, search is biased towards whichever target conjunction is framed as the search template, which we described as a confirmation bias (Rajsic et al., 2015). In this task, searchers will place higher priority on search stimuli that match the target template, despite the fact that template assignment is arbitrary, and inspect more stimuli in the completion of a given search than an optimal search strategy requires. To account for this bias, Rajsic et al. suggested that the cognitive costs of updating guidance on each trial may outweigh the costs of over-searching a display. Our goal in the present study was to provide direct evidence for the speculation that confirmatory searching results from a cost-benefit trade-off between determining the most efficient manner of testing a visual hypothesis and simply matching input to a goal state (i.e., a template) regardless of the current environmental statistics (Rajsic et al., 2015).

The current four experiments converged on the conclusion that more efficient visual hypothesis testing – that is, adopting templates that reduced the number of inspections necessary to find the target – was used when the costs of individual inspections were increased. In Experiment 1, we replicated our earlier findings of a confirmation bias in visual search with eye tracking, demonstrating that the confirmation bias in standard visual search is evident in oculomotor behavior: stimuli matching the confirmatory template were fixated more often, and participants often concluded that a Template Mismatching target was present after exhaustively searching for a Template Matching target, rather than searching the Template Mismatching set. Experiment 2 investigated searches when response features of stimuli, but not guiding features (i.e., colour), were gaze-contingent. In this case, when covert attention directed to the periphery could not contribute to search – either through covert shifts of attention or peripheral saccade planning (Geisler, Perry, & Najemnik, 2006) -- participants were relatively more successful at prioritizing the smaller colour subset, regardless of whether the subset contained confirmatory or falsifying information about the target proposition. In Experiment 3, when mouse-contingent searches were used, requiring more costly limb movements to inspect the search display, the balance between confirmation bias and strategic searching was further shifted towards the latter. Finally, in Experiment 4, by controlling the rate of information availability during searches, we determined that the change in strategy was indeed a response to inspection costs. Taken together, these results provide strong evidence that the tendency to adopt simpler visual search strategies is a result of the cognitive costs of more sophisticated search strategies.

An important finding that emerged from an analysis of eye tracking data in Experiment 1 is that, even in standard visual search, a mixture of the two search strategies was evident. As stated earlier, this likely accounts for our finding (Rajsic et al., 2015) that search slopes between Template Matching and Template Mismatching searches are not 2:1, as would be the case if search involved an exhaustive search of the Template Matching subset. It is not yet clear whether this mixture is due to a difference between participants in search strategies, or within participants' own performance, or a combination of both. However, our results nonetheless show that the confirmation bias manifests as an advantage for Template Matching stimuli in selection, but that this advantage is probabilistic, and can be supplanted by a more efficient search strategy.

The notion that cognitive operations incur costs, and that those costs affect how tasks are performed, is not new to cognitive psychology (see Kool, McGuire, Rosen, & Botvinick, 2010 for a review). Nor is it new to visual search; Zelinsky (1996) remarked that the effort required to guide individual shifts of attention and gaze by visual appearance may not pay off. Similarly, Võ and Wolfe (2013) have stated that the contribution of memory to search likely depends on the utility of including it as a source of guidance; if feature-based guidance suffices to find a target, memory will not guide search. In a clever demonstration of the cost-benefit approach to guidance, Solman and Kingstone (2014) have recently reported that memory contributes more to search when searching involves effectors that incur a greater energetic cost. In their study, memory played a larger role in search when search required movement of the head than movements of the eye. Our results, then, extend the contention that the costs of search affect the degree to which cognitive resources are leveraged in search, further demonstrating that guidance of attention is need-based, rather than stereotyped. In our searches, more flexible guidance was used and more inferences were made when searching using the hand than the eye.

In suggesting that search relies more on cognitive resources when inspection costs are increased, we assert that guidance by global visual statistics is a flexible cognitive process. Confirmation bias is a case of visual attention being guided to stimuli possessing a specific feature—those matching a target template. The more effective, minimal search strategy – exemplified in Experiments 2 and 3 – is a case of visual attention being guided not by a specific feature (i.e., a particular colour), but instead by the ratio between features. Selecting the smallest subset cannot be achieved by relying on a particular feature value, but instead requires an initial comparison of the size of colour sets. The results of our study suggest that visual attention is more readily guided by specific features, but that increasing search costs can shift guidance to include higher-order features. This is consistent with Wolfe et al.'s (2004; see also: Vickery, King, & Jiang, 2005) finding that specific templates more effectively guide attention than do general (i.e., categorically defined) templates. While the idea that specific templates guide attention more effectively is not new, our finding of a confirmation bias in visual search is novel in that the tendency to guide by specific templates cannot be attributed to a difference in specificity of these templates (e.g., the benefit for exemplar-based over categorical search templates); participants simply tended to choose to guide attention to the colour that was framed as the affirmative case of the search instructions. The confirmation bias in visual search is, we believe, among the strongest examples of a top-down search strategy directed by a factor outside of performance incentives.

From an implementation standpoint, one could account for the confirmation bias as an amplification of the bottom-up salience of Template Matching features in an integrated salience map, with the result being guidance of attention towards stimuli possessing Template Matching features. In the context of Guided Search, this has been described as adding additional weight to the output of the feature channels that code for features matching the target template (Wolfe, 2007). Alternatively, in the context of the Target Acquisition Model (TAM; Zelinsky, 2008), one could consider the template conjunction (e.g., a green P, as in Figure 1) to be used in constructing the target feature vector, which is then correlated with the available perceptual information across the visual field. This could account for the reduction in confirmatory searching in Experiment 2, since the correlations across the visual field with the target template (the Target Map, as implemented in TAM) would likely drop as letter forms are removed from the periphery in the gaze-contingent task. However, we are not aware of any models of search that could account for the results of Experiment 3, given that the critical difference was non-visual (the effector used to reveal information), or Experiment 4, where the temporal dynamics of to-be-searched stimuli affected guidance.

The temporal dynamics of confirmatory search can have, as we see it, three possible explanations. A purely top-down perspective would suggest that the active maintenance of a particular hypothesis, or template, in working memory could be the source of bias signals, such that the active framing of the search task leads to prioritized selection of template-matching stimuli (Olivers, Meijer, & Theeuwes, 2006). An alternative, purely bottom-up perspective would suggest that initial priming from the search instructions, in wherein the template color, but not the non-template color, is presented, could produce the measured bias via priming through selection history (Awh, Belopolsky, & Theeuwes, 2012; Theeuews, Reimann, & Mortier, 2006; Krouijne & Meeter, 2016). A third option, which we prefer, is a mixture of both, where top-down attentional sets are automatized through priming mechanisms (Woodman, Carlisle, & Reinhart, 2013; Wolfe, Butcher, Lee, & Hyle, 2003). In our initial study (Rajsic et al., 2015), we found confirmatory searches both when a single search was performed per template and when one template was used for all searches. In addition, we found that self-reported strategy did not relate well to the strategy revealed from search RT analyses. These findings are compatible with a priming explanation. On the other hand, some recent experiments that we have conducted suggest that priming – at least visual priming – cannot entirely explain these search patterns, as similar searching occurs when instructions are purely linguistic (i.e., participants are asked whether the target letter is on the red stimulus, without showing a red stimulus; Rajsic, Taylor, & Pratt, accepted). All things considered, a hybrid account, where attentional sets are bootstrapped as initial templates are automatized through use, appears most promising. One interesting implication of this account is that tasks like ours, where no particular attentional set clearly the most efficient for task completion, may produce the largest variety in attentional styles, and indeed the most pronounced effects of task-irrelevant factors like instructions and stimulus salience.

Returning to the primary finding of our study, reduction in confirmatory searching with increased inspection costs points to the possibility that the type of guidance in a given search is a balance of the costs of computing guidance and the costs of gathering information, over and above the nature of the stimuli being searched. Indeed, search efficiency is affected by more than just the stimuli in a display: selection history (Maljkovic & Nakayama, 1994; Wang, Kristjansson, & Nakayama, 2005), instructions (Sobel & Cave, 2002; Smilek, Enns, Eastwood, & Merikle, 2006), and the contents of working memory (Olivers, Meijer, & Theeuwes, 2006; Soto, Hodsoll, Rotchstein, & Humphreys, 2008) all affect guidance in visual search. How each of these factors influence search in a given situation may depend on a cost-benefit analysis between the performance gain afforded by more flexible guidance, and the time taken to realize the flexible guidance. However, an important issue to be resolved is the flexibility of cost-benefit computations, if they are indeed explicitly calculated. Given that search costs tend to be temporal in nature, a race-model approach between guidance computation and implementation would be a simple heuristic for achieving strategic search guidance (Võ & Wolfe, 2013), and thus represents a good null hypothesis for tests of flexibility. However, as Experiment 4 shows, the effects of practice and strategy learning complicate this issue. Indeed, research on visual search is actively being extended towards the topic of visual foraging, showing a role for the foraging effector in selection strategies (Jóhannesson et al., 2015), balancing between opportunity and priming in target selection (Wolfe, Aizenman, Boettcher, & Cain, 2016), and variations in self-imposed search path structure when less information is available in the search environment (Solman & Kingstone, 2016).

It is worth noting that the present results do not fit with the notion that working memory limitations alone are responsible for the inefficient confirmatory search found in unrestricted versions of our task (Rajsic et al., 2015). Across the current four experiments, instructions and stimuli remained similar, and we introduced no manipulations expected to affect working memory availability. Nonetheless, search strategy varied reliably. If anything, one would expect that gaze- and mouse-contingent tasks might tax working memory more than a standard visual search task, albeit, not visual working memory (see Roper and Vecera, 2013 for an example of how different types of memory load can affect search in different ways). Yet, the ability to efficiently guide attention was improved in these conditions. It is perhaps unusual to find an improvement in strategy when additional constraints are placed on the participant; a large body of research supports the general conclusion that as tasks become more difficult, performance suffers, as difficulty strains capacity-limited controlled processes (Schiffrin & Schneider, 1977). Relatedly, one might argue that, in light of demonstrations that guidance from working memory tends to reduce as more items are remembered (van Moorselar, Theeuwes, & Olivers, 2014), a higher working memory load in experiments 2 and 3 reduced template-based guidance, allowing attention to be driven more by bottom-up salience (i.e., the smaller subset). However, the increasing use of inference that accompanied the same manipulations, which would also rely on cognitive processes, contradicts this possibility. Instead, we believe that the primary change induced by the gaze- and mouse-contingent search manipulations was not difficulty per se, but the cost of each sample taken from the display in search. This does not make the task more difficult, cognitively, but instead changes the relative payoff of different search strategies.

With respect to the confirmation bias, our results support a view of the confirmation bias that contextualizes it in terms of performance, not in terms of truth (Friedrich, 1993; Arkes, 1991). Decision makers are assumed to have the intention to seek truth and make optimal decisions, but their decisions must satisfy more constraints than the maximization of accuracy. In accounting for the presence of biases and heuristics in decision-making, it is critical to consider the costs of implementing a given analysis; spending hours choosing where to go for dinner is only sensible if the difference in the meals' quality offsets the costs of the deliberation. A given action policy should be judged both in terms of its likelihood of success and its simplicity, and human decision making indeed incorporates both of these goals (Meier & Blair, 2012). Our results demonstrate that the minimization of planning costs dictates search policy not only in explicit decision-making, but also in visual search policy. This result is perhaps surprising: visual information is phenomenologically characterized by its immediacy and availability, and so it is hard to imagine that it would not be maximally exploited to improve performance. However, even shifts of gaze come at a cost – incurred at planning and motor stages, but also in terms of lost time – and these costs affect the guidance of search (Araujo, Kowler, & Pavel, 2001).

References

Araujo, C., Kowler, E., & Pavel, M. (2001). Eye movements during visual search: the costs of choosing the optimal path. *Vision Research, 41*, 3613-3625.

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin, 110*(3), 486-498.

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in cognitive sciences*, *16*(8), 437-443.

Bacon, W. F., & Egeth, H. E. (1997). Goal-directed guidance of attention: evidence from conjunctive visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(4), 948-961.

Barret, D. J. K. & Zobay, O. (2014). Attentional control via parallel target-templates in dual-target search. *PLoS One, 9*(1), e86848.

Beck, V. M., Hollingworth, A., & Luck, S. J. (2012). Simultaneous control of attention by multiple working memory representations. *Psychological Science*, *23*(8), 887-898.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, *10*, 433-436.

Bundesen, C. (1990). A theory of visual attention. *Psychological review*, *97*(4), 523.

Bundesen, C. (1998). A computational theory of visual attention. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *353*(1373), 1271-1281.

Buttaccio, D. R., Lange, N. D., Thomas, R. P., & Dougherty, M. R. (2015). Using a model of hypothesis generation to predict eye movements in a visual search task. *Memory and Cogniton*, *43*(2), 247-265.

Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 353*(1373), 1245-1255.

Dombrowe, I., Olivers, C. N., & Donk, M. (2010). The time course of working memory effects on visual attention. Visual Cognition, 18(8), 1089-1112.

Fan, J. E. & Turk-Browne, N. B. (2015). Incidental biasing of attention from visual long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* doi: 10.1037/xlm0000209

Friedrich, J. (1993). Primary error detection and minimization (PEDMIN) strategies in social cognition: a reinterpretation of confirmation bias phenomena. *Psychological review*, *100*(2), 298-319.

Geisler, W. S., Perry, J. S., & Najemnik, J. (2006). Visual search: The role of peripheral information measured using gaze-contingent displays. *Journal of Vision*, *6*(9), 1.

Gregory, R. L. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society of London, 352*, 1121-1128.

Jóhannsesson, O. I., Thornton, I. M., Smith, I. J., Chetverikov, A., & Kristjánsson, À. (2015). Visual foraging with fingers and eye gaze. *i-Perception, 7*(2), 1-18.

Klayman, J. (1995). Varieties of Confirmation Bias. *Decision Making from a Cognitive Perspective: Advances in Research and Theory*, *32*, 385.

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94*(2), 211–228.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2011). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General, 139(4)*, 665-682.

Krouijne, W. & Meeter, M. (2016). Long-term priming of visual search prevails against the passage of time and counteracting instructions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(8), 1293-1303.

Liversedge, S. P. & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4(1)*, 6-14.

McKenzie, C. R. (2004). Hypothesis testing and evaluation. In Koehler, D. R., & Harvey, N. (Ed.), Blackwell handbook of judgment and decision making. Wiley-Blackwell.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & cognition*, *22*(6), 657-672.

Meier, K. M. & Blair, M. R. (2013). Waiting and weighting: Information sampling is a balance between efficiency and error-reduction. *Cognition, 126(2)*, 319-325.

Mynatt, C. R., Doherty, M. E., & Dragan, W. (1993). Information relevance, working memory, and the consideration of alternatives. *The Quarterly Journal of Experimental Psychology, 46A(4)*, 759-778.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2(2)*, 175-220.

Oaksford, M., & Chater, N. (1994). A Rational Analysis of the Selection Task as Optimal Data Selection. Psychological Review, 101(4), 608–631.

Olivers, C. N., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1243-1265.

Olivers, C. N., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: when it guides attention and when it does not. *Trends in cognitive sciences, 15*(7), 327-334.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*(4), 437-442.

Rajsic, J., Taylor, J. E. T., & Pratt, J. (under review). Out of sight, out of mind: Matching bias underlies confirmatory visual search. *Attention, Perception, & Psychophysics.*

Rajsic, J., Wilson, D. E., & Pratt, J. (2015). Confirmation bias in visual search. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. doi: 10.1037/xhp0000090

Roper, Z. J., & Vecera, S. P. (2014). Visual short-term memory load strengthens selective attention. *Psychonomic bulletin & review, 21(2),* 549-556.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review, 84(2),* 127-190.

Smilek, D., Enns, J. T., Eastwood, J. D., & Merikle, P. M. (2006). Relax! Cognitive strategy influences visual search. *Visual Cognition, 14(4-8)*, 543-564.

Sobel, K. V., & Cave, K. R. (2002). Roles of salience and strategy in conjunction search. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1055-1070.

Solman, G. J. & Kingstone, A. (2014). Balancing energetic and cognitive resources: Memory use during search depends on the orienting effector. *Cognition*, *132*(3), 443-454.

Solman, G. J. & Kingstone, A. (2016).. Endogenous strategy in exploration. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1634-1649.

Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in cognitive sciences*, *12*(9), 342-348.

Stroud, M. J., Menner, T, Cave, K. R, & Donnelly, N. (2012). Using the dual-target cost to explore the nature of search target representations. *Journal of Experimental Psychology: Human Perception and Performance, 38*(1), 113-122

Theeuews, J., Reimann, B., & Mortier, K. (2006). Visual search for featural singletons: No top-down modulation, only bottom-up priming. *Visual Cognition, 14*(4-8), 466-489.

Thomas, R., Dougherty, M. R., & Buttaccio, D. R. (2014). Memory Constraints on Hypothesis Generation and Decision Making. *Current Directions in Psychological Science, 23(4),* 264-270.

van Moorselaar, D., Theeuwes, J., & Olivers, C. N. (2014). In competition for the attentional template: Can multiple items within visual working memory guide attention*?. Journal of experimental psychology: human perception and performance, 40(4),* 1450-1464.

Vertegaal, R. (2008). A Fitts law comparison of eye tracking and manual input in the selection of visual target. *Proceedings of the 10th International Conference on Multimodal Interfaces,* 241-248.

Vickery, T. J., King, L., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision, 5,* 81-92.

Võ, M. L. H., & Wolfe, J. M. (2013). The interplay of episodic and semantic memory in guiding repeated search in scenes. *Cognition, 126(2)*, 198-212.

Wang, D., Kristjansson, A., & Nakayama, K. (2005). Efficient visual search without top-down or bottom-up guidance. *Perception & Psychophysics*, *67*(2), 239-253.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20*(3), 273–281.

Wolfe, J. M. (2007). Guided search 4.0. *Integrated models of cognitive systems*, 99-119.

Wolfe, J. M., Aizenman, A. M., Boettcher, S. E. P., & Cain, M. S. (2016). Hybrid foraging search: Searching for multiple instances of multiple types of target. *Vision Research, 119,* 50-59.

Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 483-502.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15(3),* 419-433.

Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research, 44*, 1411-1426.

Wolfe, J. M., Võ, M. L. H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in cognitive sciences, 15(2)*, 77-84.

Woodman, G. F., Carlisle, N. B., & Reinhart, R. M. G. (2013). Where do we store the memory representations that guide attention? *Journal of Vision*, *13*(3), 1-17.

Zelinsky, G. J. (1996). Using eye saccades to assess the selectivity of search movements. *Vision research, 36(14)*, 2177-2187.

Zelinsky, G. J. (2008). A Theory of Eye Movements during Target Acquisition. *Psychological Review, 114(4)*, 787-835.