Confirmation Bias in Visual Attention

by

Jason Daniel Rajsic

A thesis submitted in conformity with the requirements for the degree of Doctorate of Philosophy
Department of Psychology
University of Toronto

© Copyright by Jason Rajsic 2017
Confirmation Bias in Visual Attention

Jason Rajsic

Doctorate of Philosophy

Department of Psychology
University of Toronto

2017

Abstract

This dissertation evaluates the proposal that visual attention is biased towards stimuli that could potentially verify a perceptual hypothesis, compared to stimuli that might falsify it. To do so, I use a visual search task with two, mutually exclusive, coloured targets (e.g., a red p or a blue p). Before each set of searches, one target colour is designated positive, or confirmatory, information and one is designated negative information by requiring affirmative responses or negative responses to one of the two targets (e.g., press Z if the p is red, press M if it is not). The first set of experiments, reported in Chapter 2, showed that confirmatory search patterns were observed for both novel and practiced hypotheses, were observed despite instructions to the contrary, and were mitigated by a search colour-preview. Chapter 3 measured the ocular correlates of the search bias, showing that more fixations indeed go towards confirmatory stimuli, but showed that when stimulus inspections took longer, search became less confirmatory and more efficient. Finally, Chapter 4 demonstrated that confirmatory search did not occur when hypotheses referred to visually heterogeneous stimuli and also did not occur when hypotheses included a negation, suggesting that a visual matching bias underlies confirmatory search. Confirmation bias thus appears to be a side-effect of goal-directed cognition in general, and the present results are discussed in relation to theories likening reasoning to foraging and information search.
Acknowledgments

I owe a debt of gratitude to a great number of people for making this dissertation possible. First, my advisor, Jay Pratt. Thank you for believing in me, for pushing me to reach my potential, and for your tireless support. I also wish to thank my former advisor, Daryl Wilson, for encouraging me to pursue research, and giving me the independence and support to follow my interests. Thank you to my lab mates, who have given me inspiration, motivation, and friendship – Davood Gozli, David Chan, Greg Huffman, Eric Taylor, Sol Sun, Matt Lowe, Kristin Wilson, Merryn Constable, Geoff Harrison, Jess Maxwell, Matt Hilchey, Mark Mills, and the many other wonderful personalities from the labs, too many to name! Thank you to my committee; Susanne Ferber for your generosity, encouragement, and advice, and Daphna Buchsbaum for your enthusiasm, feedback, and for broadening the scope of my reading.

I would like to thank my friends, who have been there to keep me sane and given me such joy during my PhD; Adam and Jessica Prewer, David Kelusky, Gaylen Lindal, Brock Criger, Jessica Spape, Jen Ng, Mike McNeely, Donna Lightfoot, Peter Sinclair, Adrian Nasagar, Justin Deonarine, Jenny Shen, Liza Igochine, and Liam Karry.

Finally, and most of all, I would like to thank my parents, Dan and Sue Rajsic for their unwavering and selfless support through the many years I’ve spent in universities – I’m still not gone! Thank you for understanding and supporting my pursuit of research, I could not possibly have done it without you. Thank you too to my sister, Jaclyn Rajsic, and my brother-in-law, Richard Walters, fellow doctors with whom I can share the peaks and troughs of research and who have been so incredibly generous to me.
Table of Contents

Contents

Acknowledgments.................................................................................................................................................. iii
Table of Contents ................................................................................................................................................ iv
List of Figures .................................................................................................................................................... viii
Chapter 1 Confirmation Bias in Visual Attention................................................................................................. 1
  1 Confirmation Bias ............................................................................................................................................. 2
  2 Selection of information: Attention .................................................................................................................. 5
  3 Overview of the Present Work ......................................................................................................................... 8
Chapter 2 Confirmatory Visual Search ................................................................................................................ 11
  4 Experiment 1 .................................................................................................................................................. 12
    4.1.1 Methods .............................................................................................................................................. 15
    4.1.2 Participants ......................................................................................................................................... 15
    4.1.3 Stimuli ................................................................................................................................................. 15
    4.1.4 Procedure ........................................................................................................................................... 15
    4.2 Results and Discussion .............................................................................................................................. 16
  5 Experiment 2 .................................................................................................................................................. 18
    5.1 Methods .................................................................................................................................................. 19
      5.1.1 Participants ....................................................................................................................................... 19
      5.1.2 Stimuli and Procedure ..................................................................................................................... 19
    5.2 Results and Discussion .............................................................................................................................. 19
  6 Experiment 3 .................................................................................................................................................. 20
    6.1 Methods .................................................................................................................................................. 21
      6.1.1 Participants ....................................................................................................................................... 21
      6.1.2 Stimuli and Procedure ..................................................................................................................... 21
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2 Results and Discussion</td>
<td>21</td>
</tr>
<tr>
<td>Experiment 4</td>
<td>23</td>
</tr>
<tr>
<td>7.1 Methods</td>
<td>23</td>
</tr>
<tr>
<td>7.1.1 Participants</td>
<td>23</td>
</tr>
<tr>
<td>7.1.2 Stimuli and Procedure</td>
<td>24</td>
</tr>
<tr>
<td>7.2 Results and Discussion</td>
<td>24</td>
</tr>
<tr>
<td>Experiment 5</td>
<td>26</td>
</tr>
<tr>
<td>8.1 Methods</td>
<td>26</td>
</tr>
<tr>
<td>8.1.1 Participants</td>
<td>26</td>
</tr>
<tr>
<td>8.1.2 Stimuli and Procedure</td>
<td>27</td>
</tr>
<tr>
<td>8.2 Results and Discussion</td>
<td>27</td>
</tr>
<tr>
<td>General Discussion</td>
<td>29</td>
</tr>
<tr>
<td>Chapter 3 The Price of Information</td>
<td>34</td>
</tr>
<tr>
<td>Experiment 6</td>
<td>35</td>
</tr>
<tr>
<td>10.1 Methods</td>
<td>36</td>
</tr>
<tr>
<td>10.1.1 Participants</td>
<td>36</td>
</tr>
<tr>
<td>10.1.2 Stimuli</td>
<td>36</td>
</tr>
<tr>
<td>10.1.3 Procedure</td>
<td>36</td>
</tr>
<tr>
<td>10.2 Results and Discussion</td>
<td>37</td>
</tr>
<tr>
<td>Experiment 7</td>
<td>44</td>
</tr>
<tr>
<td>11.1 Method</td>
<td>45</td>
</tr>
<tr>
<td>11.1.1 Participants</td>
<td>45</td>
</tr>
<tr>
<td>11.1.2 Stimuli and Procedure</td>
<td>45</td>
</tr>
<tr>
<td>11.2 Results and Discussion</td>
<td>45</td>
</tr>
<tr>
<td>Experiment 8</td>
<td>50</td>
</tr>
<tr>
<td>12.1 Method</td>
<td>51</td>
</tr>
<tr>
<td>Chapter</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>12</td>
<td>Results and Discussion</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Experiment 9</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>General Discussion</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Experiment 10</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Experiment 11</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Experiment 12</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
18.1 Methods

18.1.1 Participants

18.1.2 Stimuli and Procedure

18.2 Results and Discussion

19 General Discussion

Chapter 5 Looking for Answers

References

Copyright Acknowledgements
List of Figures

Figure 1. A sample of stimuli used, with predictions for the confirmatory search strategy and minimal search strategy. Panel (a) depicts a sample instruction for a block of searches where the target letter is p, and the template colour is blue. For this search block, then, a Colour Match trial would be a trial where the target p appeared in blue, and a Colour Mismatch trial would be when the target p appeared in red. Panel (b) illustrates two possible searches for this search block where predictions for the two hypothetical strategies most strongly differ; when the majority colour of stimuli does not match the template colour. Potential search paths are overlaid for each search strategy. Predictions in (c) are derived by counting the number of expected inspections, depicted as dashed circles in panel (b), for each possible search display type. In panel (c), the expected results for confirmatory searching are shown in the top graph and for minimal searching in the bottom graph. .......................... 14

Figure 2. Median Response Times (left) and Mean Accuracy (right) for the search task in Experiment 1. Error bars in this and all other figures depict one within-subjects standard error (Cousineau, 2005). .................................................. 17

Figure 3. Median Response Times (left) and Mean Accuracy (right) for the search task in Experiment 2 .......................................................... 20

Figure 4. Median Response Times (left) and Mean Accuracy (right) for the search task in Experiment 3 .......................................................... 22

Figure 5. Median Response Times (left) and Mean Accuracy (right) for the search task in Experiment 4 .......................................................... 25

Figure 6. Median Response Times (left) and Mean Accuracy (right) for the search task in Experiment 5 .......................................................... 28

Figure 7. Panel A depicts median correct search times in Experiment 6, and Panel B depicts the average number of fixations per search. Error bars depict one within-subjects standard error. .......................... 38

Figure 8. Sample inspections for Experiments 6, 7, and 8. Trials depicted are Template Mismatch trials for Matching Subset Size 2 (upper row) and 6 (lower row). Yellow arrows indicate paths
between successive fixations (Experiments 6 and 7), and green arrows indicate paths between successive stimulus-reveals (Experiments 7 and 8).

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 6. 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.

Figure 10. Proportion of trials where targets were fixated before being correctly identified in Experiment 6. 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.

Figure 11. Panel A depicts median correct search times in Experiment 7, and Panel B depicts the average number of fixations per search. Error bars depict one within-subjects standard error.

Figure 12. 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 7. 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.

Figure 13. Proportion of trials where a correct response was given and the target was inspected before search termination in Experiment 7. 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.

Figure 14. Panel A depicts median correct search times in Experiment 8, and Panel B depicts the average number of fixations per search. Error bars depict one within-subjects standard error.

Figure 15. 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 8. 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.

Figure 16. Proportion of trials where a correct response was given and the target was inspected before search termination in Experiment 8. 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.

Figure 17. An example illustration of the stimuli and procedure used in Experiment 9. 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).

Figure 18. 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.

Figure 19. A sample search instruction (upper row) and sample search array (lower row) for Experiments 10, 11, and 12 (columns a, b, and c). Stimuli are not drawn to scale.

Figure 20. Median Response Times in Experiment 10. Error bars in this and all other figures depict one within-subjects standard error (Cousineau, 2005).

Figure 21. Median Response Times Experiment 11 for Standard Prompts (left) and Negation Prompts (right).

Figure 22. Median correct Response Times in Experiment 3 for Gray Templates (left) and Colourful Templates (right).

Figure 23. Median search times for Experiment 4. Error bars reflect one, within-subjects standard deviation of the mean.
Chapter 1
Confirmation Bias in Visual Attention

Although the world we perceive is rich and full of detail, at any given moment we are only aware of a fraction of this information. The content of our awareness can be determined by factors beyond our control, such as the physical salience of objects or their learned value, or by voluntary factors, guided by the goals of the individual. When it comes to distraction — having one’s attention oriented away from critical information — we often think of involuntary sources of attentional control. Flashing lights, loud noises, or familiar sights pull our attention away from the task at hand. However, voluntary control can prove equally problematic in our failures to select the most important visual information. A most striking example was provided by Simons and Chabris (1999), who demonstrated that many observers trying to count passes of a basketball within a small group of people were completely unaware of a gorilla walking through the same scene. Clearly, our voluntarily allocated attention can lead us astray as well if important events fall outside the scope of our momentary interests.

In this dissertation, I document a novel failure of top-down attention. In doing so, I draw upon findings in psychology remote from the field of visual cognition. Specifically, I treat the use of top-down, goal-driven attention as a form of visual hypothesis testing. From this perspective, one can determine whether the approach that observers adopt in testing hypotheses matches those that would be ideal from a normative perspective. Research in decision making and reasoning has demonstrated many ways in which human cognition departs from prescriptive norms (Evans, 2003; Stanovich & West, 2000; Tversky & Kahneman, 1973). One of the most notorious of such cognitive biases is the confirmation bias (Nickerson, 1998), described as the tendency to preferentially inspect or analyse information that would confirm a focal hypothesis. The confirmation bias violates normative rationality by neglecting alternative hypotheses in the reasoning process. The bulk of research on confirmation bias is concerned with what might be called “offline cognition”, or cognition about situations other than the individual’s immediate situation – often imagined or remembered. However, if the confirmation bias is a fundamental principle of human reasoning, it should manifest equally in the analyses of one’s immediate surroundings. As such, the goal of this dissertation is to determine if, and if so, how, the confirmation bias occurs in visual processing. Specifically, this dissertation determines whether
visual information that is congruent with a target proposition is processed more quickly or thoroughly than information incongruent with it. Before outlining the proposed set of experiments, I will briefly review research on confirmation bias. Following that, I will review the similarities between cognitive mechanisms underlying reasoning and visual attention that justify the predictions of a confirmation bias in visual attention.

1 Confirmation Bias

The tendency to interpret evidence in a one-sided fashion has been long known to thinkers (Nickerson, 1998). However, the scientific investigation of this bias began to flourish in the latter half of the 20th century, when Peter Wason (1960) used the Rule Discovery task to show that individuals tend to test their hypotheses in a confirmatory manner. Inspired by Popper’s doctrine of falsificationism, Wason gave participants an instance of a rule – in this case, the set “2, 4, 8” – and asked participants to generate triplets that would let help them discover the rule that defined the allowable set of triplets. Wason noticed that participants overwhelmingly opted to test sets that belonged to their hypothesized rule set, as opposed to those that did not, which directly contradicted the normative behaviour prescribed by falsificationism. A second, equally influential demonstration of the confirmation bias came again from Wason (1968), this time in the form of the selection task. In the selection task, individuals are asked to choose which of four cards must be checked (turned over) in order to test the validity of a particular rule. The visible faces of the cards may show a consonant (e.g., B), a vowel, (e.g., A), an even number (e.g., 4) and an odd number (e.g., 7). The rule may then be stated as “If a card has a vowel on one side then it has an even number on the other”. Logically, the cards that must be examined are the A and the 7, for if the A does not reveal an even number, the rule is falsified. Additionally, if the 7 card should reveal a vowel, the rule is also falsified. Results typically show that all, or nearly all, individuals select the vowel as necessary, but comparatively few choose the odd number, with a similar number choosing the even number as being diagnostic. These results have also been taken to suggest a confirmatory testing strategy – participants prefer to resolve uncertainties that may be congruent with the rule being evaluated, but not those that could not be (i.e., the 7, which could only reveal a case that is inconsistent with the rule (a vowel) or a case that is irrelevant to the rule (a consonant)).
Confirmation bias has also been heavily studied in the area of social cognition. For example, Snyder and Swann (1978) asked participants to generate questions that could be asked to determine whether someone was extraverted. Given this prompt, participants were more likely to generate questions that would confirm the trait of extraversion, if answered in the affirmative – for example, “do you like big parties” – as opposed to those that could disconfirm extraversion if answered in the affirmative. The confirmation bias, thus, extends beyond abstract reasoning tasks. However, social cognition researchers have shown that the confirmation bias can be reduced when problems are framed as “cheater detection”, where a social norm may be violated (Gigerenzer & Hug, 1992). In these contexts, for example, when the selection task is reframed to be about detection of underage drinking, participants are far more likely to select both the positive antecedent (modus ponens: is drinking) and negative consequent (modus tollens: isn’t over 19). This has led some researchers to suggest that human hypothesis testing has been designed by evolution for tracking social dynamics, thereby accounting for its apparently poor performance in non-social problems (Cosmides & Tooby, 1992).

While it is easy to see how the confirmation bias (indeed, any cognitive bias) is undesirable, several researchers have pointed out how this approach to hypothesis testing may be globally optimal. Klayman and Ha (1987) point out that confirmatory searching, or a positive test strategy, is optimal when hypotheses are “sparse”; that is, when a given hypothesis makes fewer positive claims than negative claims. For example, the claim “if it is a furry animals, then it is a cat” has only one positive consequent (that, given a furry animal, it is a cat), but many negative consequents (e.g., given a furry animal, it is a dog, bear, or an otter), and so it is a sparse hypothesis. If most of the hypotheses that humans deal with are sparse, then confirmatory searching (or a positive test strategy, in Klayman and Ha’s terminology) is more efficient at gaining information in a hypothesis-testing context than falsification. A similar argument has been made by Oaksford and Chater (1994), who showed that, if Wason’s selection task is analysed using conditional probabilities, that the expected information gain for confirmatory tests (the “A” card in the example above) is greater than falsification tests (the “7” card in the above example), so long as the probability of positive antecedents and consequents (vowels and even numbers, respectively) is low. Further arguments in favour of the adaptive nature of positive testing have been made from the standpoint of the expected utility of hypothesis tests. Friedrich (1993) has argued that most hypotheses are about desirable outcomes (e.g., when I
drink coffee, I feel more energetic), and so testing hypotheses in a confirmatory fashion is sensible. For example, if what I really value, or need, is an energy boost, then it is more important that I receive the energy boost than that I learn whether coffee truly increases my energy. As Friedrich argues, hypothesis testing may be designed to avoid “costly errors” in addition to seeking truth. In this light, rejecting hypotheses that are spuriously true (e.g., placebo effects) is of relatively low priority, given that it these tests would risk the costly error of losing the desired benefits of an object or action if the hypothesized relation is indeed true. Given that humans are biological organisms, with a constant need for resources that ensure survival, this sort of short-sighted, opportunistic hypothesis testing is rational.

Optimal or not, the confirmation bias still presents a problem, as it can lead to the persistence of ill-supported beliefs. To understand why the confirmation bias occurs in situations where it is not ideal, it is important to understand the cognitive mechanisms that underlie confirmation biases. First of all, evaluating the evidence in relation to a set of hypotheses requires comparing given data, however they are represented, to hypothetical data predicted by a hypothesis. Such interpretive settings have been described as conditional reference frames (Koehler, 1991). Evaluating evidence that can support multiple alternative hypotheses requires that these alternatives be represented in such a way that they can be updated upon the analysis of new information. This places a noteworthy demand on any memory processes involved in representing hypotheses. While long-term memory has a very high capacity for information, it is not suited for the sort of dynamic updating required in explicit hypothesis evaluation. Instead, short-term memory processes (i.e., working memory) are central in the evaluation of hypotheses (Mynatt, Doherty, & Dragan, 1993; Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002). Working memory has a notoriously small capacity for information storage with estimates ranging from between 3 to 7 items (Cowan, 2001; Luck & Vogel, 1997; Miller & Miller, 1956), with some arguing that only one item can be “used” at a time (Oberauer, 2002). These cognitive constraints provide a basis for the observation that reasoners typically evaluate possible hypothesis in relative isolation (as noted by Mynatt, Doherty, & Dragan, 1993). A consequence of this limitation is that ambiguous verification may appear as support; an observation that is probable under two different hypotheses may be taken to strengthen the hypothesis being considered.
Neglecting the evaluation of competing hypotheses does not only hinder one’s ability to determine how to interpret a new observation, but also leads to biases in the selection of information. Having a particular hypothesis in mind causes the selection of information that is “pseudodiagnostic” (Doherty, Mynatt, Tweney, & Schiavo, 1979), that is, information that allows one to know whether a given set of observations is likely or not under a certain hypothesis, but does not indicate which of two (or more) hypotheses is more likely. This can potentially exacerbate the confirmation bias, given that incomplete data will be collected with which hypothesis evaluation can proceed. Indeed, Fiedler (2000) has argued that incomplete, or biased, sampling of information alone can lead to cognitive biases. Thus, the selection of information is a critical determinant of whether hypothesis evaluation will be balanced.

2 Selection of information: Attention

In cognitive psychology, the word attention is synonymous with information selection. A typical introduction to a research article in the field of attention is motivated by noting that the rate of information flow through the senses is too great for all information to be analysed, and so only a subset can be analysed. Decades of research have been dedicated to understanding how the subset of information is chosen at any given moment (Broadbent, 1958; Deutsch & Deutsch, 1963; Awh, Belopolsky, & Theeuwes, 2013). Which subset is selected for further processing, is often controlled by relatively simple aspects of the sensory input, such as sudden onsets (Yantis & Jonides, 1984), motion (Abrams & Christ, 2003), or uniqueness (Theeuwes, 1992); it can also be actively controlled by intentions and goals (Bacon & Egeth, 1994; Folk, Remington, & Johnston, 1992). This ability to flexibly adjust the processing of perceptual information based on goals is a necessary component of arguably all goal-oriented behaviour, enabling the execution of everyday tasks (Land & Hayhoe, 2001). In order to control one’s behaviour in the service of a given goal, attentional mechanisms must ensure that the perceptual information required to choose the correct actions is available for decision making. For example, during highway driving, one must ensure that the positions and trajectories of nearby vehicles are known when attempting to change lanes, whereas information about the other vehicles’ colours, or their passengers, is not strictly necessary to drive successfully.

In the context of goal-driven behaviour, we often seek very specific information. For example, when I make breakfast in the morning, I may choose to use my sensory apparatus to
find an answer to the question “where is the peanut butter?” Depending on the task and my prior knowledge, I may even go about these tasks by testing hypotheses, for example, covertly asking “is the peanut butter on the refrigerator door?”, which would guide my subsequent information-seeking behaviour (i.e., opening the door and fixating various locations on the door). Seen this way, the controlled selection of information in vision (i.e., top-down visual attention) can be described as perception that is guided towards the goal of verifying or falsifying some hypothetical state of the perceiver’s environment. Indeed, the most frequently used laboratory task in the study of visual attention has this structure; visual search (Duncan & Humphreys, 1989; Neisser, 1964; Treisman & Sato, 1990; Wolfe, Cave, & Franzel, 1989).

In a typical visual search, the observer is presented an array of discrete stimuli, and asked to determine whether one particular type of stimulus is present in the display. Observers must search the array of stimuli to determine whether, on any given instance of a visual search, the target is present or absent. Typically, the stimulus that is being searched for is called the “target”, and all other stimuli in an array are “distractors”. Given that the term “target” can ambiguously refer to either a physically present stimulus that can be found, or the more abstract description of a target, the term “target template” is useful for referring to the description of the target of a visual search. This template can be used to selectively process those aspects of the visual input that are likely to yield information about a target (Wolfe, Cave, & Franzel, 1989). While not referred to as such, one can consider the template to be a sort of visual hypothesis.

Many theories of the top-down mechanisms used for visual selection in visual search propose that information processing is biased towards goal-relevant information, and that core information processing units have a limited capacity, similar to theories of explicit hypothesis testing (Mynatt, Doherty, & Dragan, 1993; Süß, Oberauer, Wittman, Wilhelm, & Schulze, 2002). Guided Search (Wolfe, 1998; 2004), Feature Integration Theory (Treisman & Sato, 1990), the Theory of Visual Attention (Bundesen, 1990), the Boolean Map Theory of Visual Attention (Huang & Pashler, 2007), and Biased Competition (Desimone & Duncan, 1995) all state that visual search is guided, biased, or otherwise prioritized towards stimuli matching some template of the stimulus that is being searched for, whether that be through applying gain to template-matching stimuli or through suppressing of template-mismatching stimuli. In addition, research on the role of working memory in visual search guidance has led some investigators to conclude that visual selection can only be guided by a single template at a time (Olivers, Peters,
Houtkamp, & Roelfsema, 2011, but see Beck, Hollingworth, & Luck, 2012). These theorized mechanisms allow for more economical processing of visual information, so that aspects of the visual input that are task-relevant can receive preferential processing. However, prioritizing stimuli that are similar to a target template, paired with the limitation of a single template being maintained, is theoretically sufficient to produce a confirmation bias in visual search. If we consider the target template as a hypothetical visual state, evidence for alternative visual states will take longer to accumulate to the extent that this information is incongruent with features in the template. This sort of visual guidance is confirmatory; information that supports the presence of goal-relevant information is increased in salience. Moreover, in the case where critical visual information is fleeting, alternative states may never reach awareness during episodes of heightened top-down guidance. The latter possibility has been demonstrated in studies of inattentional blindness, where conspicuous events go unnoticed while one is engaged in a demanding visual task, despite no change in sensory input (Mack & Rock, 1998; Simons & Chabris, 1999). For example, observers are readily able to maintain the mundane percept of “people playing basketball” in the face of visual information that is grossly inconsistent with this interpretation; namely, a gorilla walking through the middle of the group.

Though the evidence to date supports the possibility that top-down visual selection mechanisms may automatically lead to confirmatory searching, the design of visual search tasks encourages confirmatory selection as a useful strategy, and so confirmatory searching could be a voluntarily adopted strategy. In the typical visual search task, the task goal is to report whether a target is present or absent in a given array of stimuli, any one of which could be the target stimuli. In such tasks, where a target can be present or absent, distractors (or non-targets) provide no information about the potential target, making a confirmatory search strategy optimal. The proposition “there is a target” can be verified in less time (on target-present trials) than it can be falsified because registration of the presence of a target stimulus can occur before every stimulus in the display is fully processed, providing sufficient information to execute the correct response. Falsification of target presence, however, cannot be completed until all stimuli are analysed to the level of response-discriminating categories. Therefore, it is unclear whether visual selection is confirmatory by its nature, or whether confirmatory selection is simply adopted as a useful strategy for visual search.
To determine whether top-down visual guidance has a confirmation bias by default, or whether this potential bias may simply result from tasks demands, it is necessary to provide a direct measurement of the perceptual hypothesis testing used in visual search. To this end I used a task where on each trial a target stimulus could have one of two features, with each conjunction of target identity and target feature being equally probable. To assess confirmatory searching, one particular conjunction was designated as a target template by framing the search as a question that was to be answered about a given search display. In each search, a variable proportion of distractor stimuli possessed the template-matching feature, with the remaining proportion possessing the template-mismatching feature. By measuring search times, I am able to determine whether participants perseverate on searching through template matching stimuli – those that could confirm the presence of the target defined in the template, if inspected with covert attention – instead of opting to search through template mismatching stimuli – those that could, when inspected, disconfirm the presence of the target defined in the template. Through the chapters of this dissertation, I will determine if the processes involved in selective attention are indeed confirmatory, and seek to shed light on the factors that underlie such behaviour.

3 Overview of the Present Work

To evaluate the possibility that visual attention, like reasoning, exhibits a confirmation bias, the first chapter of this dissertation serves three purposes: to establish a method for measuring visual bias caused by positively framing one of two possible targets, to provide evidence for or against the presence of such a visual bias, and to gain initial information about the primary factors (strategic failures or automatic heuristics) underlying the bias. First, in Experiment 1, to test whether information consistent with a hypothesis attracts attention, I used a visual search in which two possible targets could appear. For example, participants could search for a target that would either be a red P or a blue P among red and blue non-Ps. One of these two possible targets, however, was established as hypothesis-congruent by implying one particular target template in the search instructions. Participants were instructed to respond with one key when the target letter (e.g., “p”) was red, and press a second key when the target letter is another colour. By specifying one of the two possible conjunctions in advance, I hypothesized that participants would adopt this conjunction as a template, and show a bias to use positive tests (Klayman & Ha, 1987), that is, to inspect items that could be an instance of this target conjunction. To test this possibility, I manipulated two factors across trials: whether the target was indeed the
“hypothesized” target (i.e., whether or not the hypothesis was true or false for a given trial), based on whether the target was template-matching or mismatching, and the proportion of stimuli that match the template suggested by the instructions. Assuming that processing additional stimuli incurs a time cost, as is true for visual search, response time should be proportional to the size of the subset of stimuli that matched the participant’s target template.

Using this method, it is possible to observe how participants select information in visual search (see Bacon & Egeth, 1997; Sobel & Cave, 2002). If participants’ search is biased towards information that could confirm the presence of a target defined by the template suggested by the search instructions, then search time will be proportional to the subset of stimuli that matches the template. If, instead, participants select the most diagnostic information, response times will be proportional to the smallest subset, regardless of whether this subset matches the template. This is because, since one of two targets is always present, searching exhaustively through either set provides sufficient information for inferring the colour of the target. If the target letter “p” can be red or blue, for example, then failing to find it in the red subset allows one to infer that it is in the blue subset. These two strategies make distinct predictions about the shape of the slope relating search time to template-matching subset size, and so response time can be used to infer the participant’s selection criteria. I predicted that search would indeed be biased towards information that could confirm the presence of a target template.

Provided that search is indeed confirmatory, the next step is to determine its source. Four broad possibilities were tested in Experiments 2-5: that confirmatory selection is due to a resistance to updating one’s goals (template inertia; Experiment 2), that the working memory load required to maintain a template interferes with cognitive flexibility (Experiment 3), that it is due to a strategic error stemming from not understanding task structure (Experiment 4), or that it is due to involuntary heuristics (Experiment 5). In order to test these different accounts, I measured performance using the same search task under conditions where each source should be minimized. To assess the contribution of template inertia, the search was conducted with a new template on every trial, so that updating is always required. To assess the contribution of memory load, the search was conducted with the same template for an entire trial, which ensured that search can be guided by long-term memory. To assess the contribution of strategic errors, participants conducted searches after having the most economic strategy (i.e., selecting the smallest colour subset) described to them. Finally, to assess the role of automatic heuristics, the
search was conducted with a colour preview display leading the actual search stimuli. This should have allowed time for strategic control to overcome any rapid, automatic selection biases.

In Chapter Two, I evaluate a functional reason for the confirmation bias in attention: cognitive costs. As suggested by Friedrich (1993), hypothesis-testing strategies ought to be sensitive to costs and benefits. In the context of search, both the heuristic (confirmation bias) and strategic approaches to finding the target will lead to the correct response. The difference is only in the time taken; the heuristic approach entails a longer search on trials where more template-matching stimuli appear. However, the strategic approach requires determining the statistics of the display in order to choose the most efficient template. If this processing requires time, and if time, or cognitive effort, is required to update selection templates, then increasing the costs of acquiring information in search should offset any cognitive costs involved in adopting the most efficient strategy for a given trial. To manipulate the costs of inspection, I compared search performance when the eyes were used to collect information and when the hand – via computer mouse – was used to collect information in mouse- and gaze-contingent search (Experiments 7 and 8, respectively). Given that mouse-contingent search relies on larger muscle groups with slower temporal dynamics than eye movements, inspection costs (both temporal and physical) should be higher in this search, and so I expected less confirmation bias. Experiment 9 tested the possibility that searchers minimize time-based costs by controlling the time that information was available in search. The chapter overall concludes that search strategies are indeed sensitive to the opportunity costs of information acquisition.

In Chapter Three, I compare three accounts of the level of representation responsible for confirmatory attentional biases: a completely involuntary priming account, where repeated presentation of a specific colour produces the bias; a visual category account, where a top-down visual template produces the bias; and an abstract relevance code that tags confirmatory stimuli in the environment as relevant, regardless of their visual appearance. To do this, instructions and search stimuli were varied, contrasting search strategies when the template was conveyed verbally versus visually, when the template contained a negation, and when the template referred to a dimension (coloured versus monochromatic stimuli) instead of a specific feature (Experiments 10-12). The results were consistent with the conclusion that confirmatory selection occurs when visual hypotheses, or templates, refer to a single visual feature (e.g., the colour red).
In order to test the possibility that visual attention is preferentially allocated towards information that might confirm a focal hypothesis, I created a visual search task wherein a target was always present, but it varied between one of two colours. Critically, participants were asked to produce one response if the target letter was one specific colour, or where it was not that colour. By framing the target colour discrimination as a yes or no question, as opposed to a colour classification (e.g., is the “p” blue or is the “p” red), I expected that one target colour would become a focal, perceptual hypothesis. To measure this presumed effect on attention, I varied the proportion of search items that matched the colour of this “perceptual template.” Given that visual search time scales with the number of items in a set being searched, this allowed me to track, using search times, the colour being prioritized. Most importantly, this colour manipulation leads to several distinct predictions that allow me to dissociate confirmatory searching from other search strategies. For example, imagine a search where one is instructed to report whether the letter “p” that appeared once in a given display was blue, knowing that a lone p, amongst other letters, would be present in the display in some colour. Eight letters onset, two of which are red, and six of which are blue. If one takes “p is blue” as a perceptual hypothesis, and searches so as to confirm this hypothesis, then one will prioritize the blue stimuli in search, as they are potential exemplars of the target template. On the other hand, a clever observer may realize that, in this situation, disconfirming the hypothesis “p is blue” would require less work, as only two stimuli need be expected before sufficient information has been collected to provide a response. If this searcher scans the red letters and finds a p, they may report that the p is not blue. If a red p is not found, one can then conclude that the p must be blue.

Because the target letter is always present in a display, participants can always infer the feature of the target stimulus by an exhaustive search through the smallest subset of coloured stimuli. Participants are able to restrict their attention to particular, colour-based subsets when targets are defined by conjunctions of colour and form (Green & Anderson, 1956; Williams,

---

1966; Egeth, Virzi, & Garbart, 1984; Kaptein, Theeuwes, & van der Heijden, 1995; Friedmann-Hill & Wolfe, 1995), with some evidence that participants will select colour subsets when instructed to even if this entails searching a larger subset (Bacon & Egeth, 1997) and some evidence suggesting that participants select whichever salient dimension is rarest (Zohary & Hochstein, 1989; Poisson & Wilkinson, 1992; Sobel & Cave, 2002). Although such studies have investigated the guidance of attention to a subset of a display, in the search tasks used, the target is either present or absent, and so no stimuli on their own can provide information against the target’s presence, therefore it is not possible for the observer to adopt a strategy of disconfirmation.

My goal in using this paradigm was to determine whether visual search exhibited confirmation bias. A confirmation bias would comprise any tendency to prioritize search stimuli that matched the target template over those that did not; in other words, a bias to search stimuli that would lead to a “yes” response, with respect to the question of whether the target defined by the template was present in the display. In Experiment 1, I demonstrate that indeed it does, and subsequent experiments were conducted in order to home in on the locus of the bias. In doing so, I entertained two broad possibilities: that confirmation bias in search results from failure of a top-down search strategy (i.e., a failure to recognize that the minimal search strategy exists), or that confirmation bias in search results from relatively automatic search heuristics, possibly arising from mechanisms underlying intertrial priming (Kristjánsson, Wang, & Nakayama, 2002). Specifically, intertrial priming could lead to confirmatory searching if stimulus selection is primarily driven by a biased selection history for template-matching colours initiated by early confirmatory searches.

4 Experiment 1

Experiment 1 was designed to assess whether confirmation bias exists in visual search. More specifically, my goal was to determine whether participants would perseverate in searching using a given target template even when this strategy was inefficient (i.e., more stimuli had to be examined). By manipulating the proportion of search stimuli possessing a template-matching feature, I was able to track participants’ search behavior by measuring the response time cost associated with increasing the size of the target-matching stimulus subset. This allows me to contrast two theoretically possible search styles; a confirmatory search strategy (i.e.,
confirmation bias; search through the template-matching colour), and a minimal search strategy (i.e., an ideal performer; search through the minority colour). Although both strategies allow for the eventual, veridical confirmation or disconfirmation of the target template, they differ in the priority of the two conclusions: under a confirmatory strategy, confirmation of the presence of stimuli matching the target template will take less time than disconfirmation. This difference can be seen in the predicted search times that follow from the two search strategies.

The confirmatory search strategy, in which stimulus selection is biased towards those stimuli that would confirm the target template, predicts a monotonic increase in search time as the proportion of template-colour matching stimuli increase, and a response time benefit when the search target matches the search template. The minimal search strategy, in which stimulus selection is intended to minimize the number of stimulus inspections necessary to produce a response, predicts a quadratic relationship between the proportion of template-colour matching stimuli and response time, with the longest searches occurring when there is an equal proportion of template-colour matching stimuli and template-colour mismatching stimuli, and a reduction in search time as the smaller subset of stimuli reduces in size. In addition, the minimal search strategy predicts no consistent relationship between whether the target stimulus matches or does not match the search template, as the template adopted for a given search would depend on which stimulus colour was in the minority. The two factors, Colour Proportion and Colour Match, should therefore produce a cross-over interaction effect on search times with a minimal search strategy. A sample search instruction and illustration of the predictions of these two strategies is provided in Figure 1.

A third possible strategy, not pictured, is that participants will not use colour to guide search at all, but instead inspect items randomly and, after finding the target letter, report its colour. Because this strategy is insensitive to colour in the selection stage, it predicts a flat search

---

2 In addition, a 2:1 ratio between search slopes for trials where the target appears in the template mismatching colour and trials where the target appears in the template matching colour could occur. However, this prediction requires the very strict assumption that selection of stimuli is completely colour-based, i.e., that information is accumulated about items in one colour subset exclusively. Given the relatively small display size (eight items), and the fact that all items are in view, it is not clear that this assumption can be upheld, and so I put forward a more robust prediction that response times should be proportional to the number of template-matching stimuli. I thank Derrick Watson for bringing this issue to my attention.
slope across the target-matching subset conditions, and an overall response time cost for reporting the target when it appears in a template-mismatching colour.

**Figure 1.** A sample of stimuli used, with predictions for the confirmatory search strategy and minimal search strategy. Panel (a) depicts a sample instruction for a block of searches where the target letter is p, and the template colour is blue. For this search block, then, a Colour Match trial would be a trial where the target p appeared in blue, and a Colour Mismatch trial would be when the target p appeared in red. Panel (b) illustrates two possible searches for this search block where predictions for the two hypothetical strategies most strongly differ; when the majority colour of stimuli does not match the template colour. Potential search paths are overlaid for each search strategy. Predictions in (c) are derived by counting the number of expected inspections, depicted as dashed circles in panel (b), for each possible search display type. In panel (c), the expected results for confirmatory searching are shown in the top graph and for minimal searching in the bottom graph.
4.1.1 Methods

4.1.2 Participants

Twelve undergraduate students volunteered to participate for course credit. All participants provided informed consent.

4.1.3 Stimuli

Search displays were composed of eight letters, spread evenly along the perimeter of an imaginary circle centred on a fixation cross. Each letter in a search display was a lowercase p, q, b, or d, approximately 2° in height and 1° in width, and was drawn approximately 8° from fixation using Arial font. The letters were always one of four similar letters (lowercase b’s, p’s, q’s, and d’s), chosen to discourage the possibility of target pop-out. Search displays had a dark grey background display. In addition, stimulus colours were selected from a pool of seven possible colours; purple, yellow, green, orange, pink, blue, and red (RGB values, respectively: 200, 0, 255; 200, 200, 0; 0, 255, 0; 255, 128, 0; 255, 128, 255; 50, 50, 255; 255, 50, 50). Before each block of searches, a set of instructions were presented on the screen.

4.1.4 Procedure

In a given block, one letter was selected as the target letter, and two stimulus colours were selected from the aforementioned pool of eight colours. At the outset of each block, participants were instructed to report whether the chosen target letter was one of the two colours, or not, using one of two keys (M and Z, on a standard keyboard) for each response. For example, if the target letter on a given block was p, and the two selected colours were red and blue, the participant may have seen the instruction, “For each search, respond as follows: Press Z if the p is this colour (blue), press M if the p is another colour.” The particular response mapping changed from block to block, such that no key was constantly mapped to either type of response. Once participants had read and memorized the search rule, they initiated a block of 30 searches by pressing the Enter key. Each participant completed 16 different blocks.

For each search, the target letter was always present, accompanied by seven distractor letters. The participant’s task was to determine, for a given display, which of two possible colours the target letter appeared in. Distractor letters were each coloured with one of the two colours selected for the block, which I will refer to as the template matching colour and template
mismatching colour, with the former referring to the colour explicitly mentioned in the block’s search rule. Two factors were manipulated within search blocks: which colour the target was drawn in (template matching or template mismatching; each equally likely), and the proportion of search stimuli of the template matching colour (0.25 – two of eight letters, 0.5 – four of eight letters, and 0.75 – six of eight letters; each equally likely). Each block contained an equal number of trials from each condition, and their order was randomized so that participants’ global strategy could be measured. Distractor letters were randomly sampled with replacement from the pool of non-target letters. Search stimuli remained on screen until a response was provided, after which the word “Correct” or “Incorrect” was presented at fixation as feedback.

4.2 Results and Discussion

To determine which strategy was implemented by participants, I conducted a 3 X 2 repeated measures ANOVA on median response time (RT) for trials with a correct response, with Colour Proportion (0.25, 0.5, 0.75) and Match-Colour (template match, template mismatch) as factors. Unless otherwise noted, all RT analyses include only trials with a correct response. Predicted results for the confirmatory search strategy are a monotonic effect of Colour Proportion and a main effect of Match-Colour, whereas the minimal search strategy predicted a quadratic (i.e., non-monotonic) effect of Colour Proportion and an interaction between Match-Colour and Colour Proportion, as searches should terminate with template matching targets and template mismatching targets, respectively, when the template matching colour is in the minority and majority, respectively. My results (Figure 2) showed that search indeed slowed when proportionally more hypothesis-confirming stimuli were present in the display. Response times increased as the proportion of template-matching colours increased, $F(2, 22) = 28.37, p < .001$, $\eta^2_p = .72$; a linear contrast proved statistically reliable, $F(1, 11) = 47.03, p < .001$, $\eta^2_p = .81$, accompanied by a marginally significant quadratic contrast, $F(1, 11) = 4.33, p = .06$, $\eta^2_p = .28$. These results show that search time indeed increased as more template-matching colours appeared in the search display, although the increase in search time was not completely linear (we return to this point in Experiment 3).
Responses were also overall slower for template mismatching colours, $F(1, 11) = 58.39$, $p < .001$, $\eta^2_p = .84$, and, crucially, this effect did not interact with Colour Proportion, $F(2, 22) = 1.21$, $p = .32$, $\eta^2_p = .10$. Although I observed a quadratic trend of Colour Proportion on search time, the lack of an interaction between Colour Proportion and Colour Match contradicts the possibility that participants adaptively switched search strategies to minimize their searches when the target-mismatching colour was the smaller colour subset, as in that case, the Colour Mismatch trials would now be those where the target matched the updated template, and therefore ought to have exhibited a reduction in search time. I return to the issue of this quadratic trend in the results section of Experiment 3. In addition to supporting the confirmatory selection strategy, these results rule out a post-selection strategy, where stimuli are selected randomly and colour is only analysed after the target letter is identified.

I analysed overall accuracy using a similar repeated measures ANOVA to determine whether results may have been due to a speed-accuracy trade-off. No main effects or interactions were observed, $F$s < 1.75, $ps > .20$, $\eta^2_p$s < .14, ruling out the possibility of a trade-off between speed and accuracy.

One possible source of perseveration on selection of the template-matching colour is inter-trial priming. Inter-trial priming refers to the facilitation of target selection when one of the
targets’ features repeat across sequential trials. Such priming is known to occur when a target’s presence varies (Olivers & Meeter, 2006), and Experiment 1’s design meets that criterion if we consider template matching and template mismatching targets to be distinct target representations.

To assess the possibility that inter-trial priming contributed to confirmatory selection, I divided trials into those that were preceded by a template matching target trial, and those that were preceded by a template mismatching target trial, which I will refer to as the Priming condition. Prime X Match–Colour X Colour Proportion repeated measures ANOVA revealed a main effect of Prime, $F(1, 11) = 14.96, p = .003$, but no interactions between Prime and other factors, $Fs < 1.92, ps > .17$. It therefore appears that inter-trial priming did not greatly affect selection performance, but that trials requiring falsification of the target template led to a small overall reduction ($M = 59$ms, $SE = 11$ms) in search time on the subsequent trial.

5 Experiment 2

The results of Experiment 1 showed a robust effect of confirmatory searching. Overall, search tended to be biased by the target template provided in initial search instructions, despite the fact that both bottom-up saliency and top-down strategy should have encouraged selection of the smaller subset when the template matching stimuli were more numerous. In the present experiment, I sought to determine whether search may have been confirmatory because it is less cognitively demanding to retain a single template across a number of trials (Shiffrin & Schneider, 1977), rather than switch templates based on the properties of any one search display. To examine this, a new search target template was presented before each search trial in Experiment 2. This allows me to test the possibility that maintaining a consistent mapping between a given target letter and colour as a search template was the source of confirmatory search behavior in the previous experiment. If confirmatory searching occurs due to a resistance to template-switching across trials, then I would expect a minimal search pattern of results (see Figure 1c). However, if confirmatory searching is an automatic consequence of guiding search, then the results of Experiment 2 will mirror those of Experiment 1.
5.1 Methods

5.1.1 Participants

Twelve new undergraduate students were recruited for the present experiment. All participants provided informed consent and were compensated with course credit. I chose twelve participants in order to match Experiment 1 for statistical power, and this general approach was also adopted for all subsequent experiments.

5.1.2 Stimuli and Procedure

The stimuli and procedure for Experiment 2 were identical to those of Experiment 1, with the exception that participants now completed 300 trials where search displays on every trial were preceded by a new set of search instructions providing a new target template. With this change, the two possible stimulus colours, the target letter, template colour, and response mapping were randomized before every trial.

5.2 Results and Discussion

I analysed median response time (Figure 3) and accuracy using separate 3 X 2 repeated measures ANOVAs as in Experiment 1. Once again, a main effect of Colour Proportion was present, $F(2, 22) = 44.04, p < .001, \eta^2_p = .80$. Consistent with confirmatory searching, I observed a significant linear trend, $F(1, 11) = 59.78, p < .001, \eta^2_p = .84$, but no reliable quadratic trend, $F(1, 11) = 1.40, p = .26, \eta^2_p = .11$. In addition, a main effect of Colour-Match was again observed, $F(1, 11) = 49.97, p < .001, \eta^2_p = .82$, such that RT was faster when the target matched the template colour, and, critically, no interaction was present, $F(2, 22) = 0.16, p = .85, \eta^2_p = .02$. These results suggest that template-colour matching stimuli were prioritized for search, and that searches terminated upon the detection of a template-colour matching target. Search template repetition, then, does not appear to be necessary for a confirmatory search strategy to emerge.
Unlike Experiment 1, I observed an accuracy effect in Experiment 2. Colour Proportion did not reliably alter accuracy, $F(2, 22) = 1.21, p = .32, \eta_p^2 = .10$, nor did Colour Proportion interact with Colour-Match, $F(2, 22) = 0.92, p = .41, \eta_p^2 = .08$, but Colour-Match did, $F(1, 11) = 5.03, p < .001, \eta_p^2 = .31$, such that more errors were made in reporting a template mismatching target than in reporting a template matching target. Combined with the RT effects, the overall picture is that template-colour mismatching targets were associated with poorer performance in general.

6 Experiment 3

The results of Experiment 2 shows that confirmatory selection still occurs when the target template changed from trial to trial, meaning that confirmatory searching does not occur simply as a result of an attempt to minimize cognitive effort. However, memorizing a new target template on every trial would certainly tax working memory; for example, Carlisle and Woodman (2011) have shown that new search targets are held in visual working memory, but the working memory load decreases as searches continue with the same template. The lack of an optimal selection strategy in Experiments 1 and 2, then, may have been due to the relatively high working memory load associated with adopting new search targets. In Experiment 3, I reduced cognitive load by having participants maintain the same search template for the entire experiment. This allowed me to determine whether a flexible selection strategy could be adopted.
when working memory demands were minimized. I anticipated two possibilities; that confirmatory searching would again occur, showing that - while not necessary – template repetition could be sufficient to encourage a confirmation bias, or that confirmatory searching would cease, showing that it was the increased cognitive load associated with adopting new target templates that prevented the use of an optimal search strategy.

6.1 Methods

6.1.1 Participants

Twelve undergraduate students were recruited for the third experiment. All participants provided informed consent and were compensated with course credit.

6.1.2 Stimuli and Procedure

I used the same stimuli and procedure as Experiments 1 and 2, with the exception of the number of trials and blocks, which were changed to 300 and 1, respectively. As a consequence, each participant received a single pair of stimulus colours, target template instruction, and response mapping that persisted for all searches in the experiment.

6.2 Results and Discussion

The results of a 3 X 2 repeated measures ANOVA again showed a main effect of Colour Proportion, $F(2, 22) = 8.20, p = .002, \eta^2_p = .43$ on search times (Figure 4). Polynomial contrasts revealed that the effect of Colour Proportion was linear, $F(1, 11) = 11.12, p = .007, \eta^2_p = .50$, and not quadratic, $F(1, 11) = 1.06, p = .33, \eta^2_p = .09$. In addition, a main effect of Colour Match was also evident, $F(1, 11) = 7.36, p = .02, \eta^2_p = .40$, with no interaction between the two factors, $F(2, 22) = 0.38, p = .69, \eta^2_p = .03$. The results of the search time analyses were quite clear: search time was consistently slower when more stimuli matched the target template, and when the target did not match the template. Both of these main effects showed that despite the reduced working memory load in Experiment 3, confirmatory selection once again occurred.
A 3X2 repeated measures ANOVA on accuracy showed no main effects or interactions, \( F_s < 1.015, \quad ps > .38, \quad \eta^2_p s < .085 \), thus ruling out the possibility of speed-accuracy trade-offs. Furthermore, adding reported strategy as a between-subjects factor yielded no reliable interactions with any factors for either RT or accuracy, \( F_s < 0.69, \quad ps > .52, \quad \eta^2_p s < .06 \), suggesting that explicit search strategies did not substantially alter participant’s searches. Overall, the results of Experiment 3 suggest that repeated search does not reduce the confirmatory nature of visual search. Thus, the results of Experiment 2 cannot be attributed solely to the increased cognitive load of the switching search targets.

As can be seen from Figures 3 and 4, the Colour Proportion X RT slopes were more shallow in Experiment 3 than in Experiment 2; the average RT cost incurred for each additional template-colour matching stimulus in Experiment 2 was 151 ms, whereas in Experiment 3, this cost was reduced to 56 ms, \( t(11) = 2.68, \quad p = .02 \). Experiment 3, then, shows that experience with a given template increases search efficiency, driven presumably by accumulated priming (Kristjánsson, Wang, & Nakayama, 2002; Becker & Horstmann, 2009). Coupled with the results of Experiment 2, I tentatively suggest that the quadratic trend in Experiment 1 reflects the contribution of economical colour selection – that is, selection of the Colour Mismatching subset when it is smaller – once sufficient experience has been acquired with a given template. This suggestion may seem paradoxical given that in Experiment 3, where only one template is ever used, no quadratic trend emerged, but this could reflect the fact that, as guidance is practiced, the
costs of switching to the Colour Mismatching subset exceed the costs of searching through the larger, Colour Matching subset.

7 Experiment 4

When the cognitive load of juggling multiple target templates was eliminated in Experiment 3, confirmatory searching nonetheless persisted. In Experiment 4, I evaluated the role of search strategy. In my previous experiments, the participants were simply instructed to respond to searches as set out in the instructions, and I aimed to observe which strategy they would adopt. Although the strategy evident in the search behavior appeared to be a confirmatory strategy, it is possible that this strategy was adopted because participants did not recognize the other strategies made available by task structure; namely, that if a target was not observed in a given colour set, one could infer that, on that trial, it appeared in the opposite colour set. In Experiment 4, I explicitly told this fact to participants at the outset of the experiment. In addition, participants were informed that the fastest way to complete a search would be to examine the stimuli in the smallest colour subset to check for a target letter. I expected that, if confirmatory search was the default, or preferred, strategy, these instructions would not affect search behavior. However, if confirmatory searches were simply an artifact of participants’ lack of familiarity with the task and its idiosyncrasies, these instructions would eliminate confirmatory searching. In the former case, a linear effect of Colour Proportion on search time and a main effect of Colour Match on search time should again be found. However, if instructions are able to curb the use of confirmatory selection, then a quadratic effect of Colour Proportion on search time, and an interaction between Colour Proportion and Colour Match should be found.

7.1 Methods

7.1.1 Participants

Twelve undergraduate students were again recruited for the present experiment. All participants provided informed consent and were compensated with course credit.
7.1.2 Stimuli and Procedure

The stimuli and procedure for Experiment 4 were identical to those of Experiment 1. Only the instructions given at the outset were modified. This consisted of the addition of the following sentences:

“The fastest way to do these searches is to look through whichever coloured letters there are fewer of. If you see the target letter, you can respond appropriately, but if you don’t, you will know it must be in the other group, and can make the opposite response.”

Participants were then led through an example where the template mismatching colour was in the minority, and told that if the target letter was in that set, they could immediately report the absence of the target-letter in the template-matching colour. In addition to verbally describing the strategy, participants were asked to identify the stimulus that would be best to inspect first in the example mentioned above. If the participant indicated that a template mismatching stimulus would be the best to inspect first, this was taken to mean that the participant had understood the strategy. However, if the participant failed to identify the mismatching stimulus, the optimal strategy and illustrative example were reiterated until the participant chose a template mismatching stimulus in the example.

7.2 Results and Discussion

Search behavior once again exhibited a confirmatory search pattern. A 3X2 repeated measures ANOVA on search RT (Figure 5) showed a main effect of Colour Proportion, $F(2, 22) = 8.66, p = .002, \eta^2_p = .44$, which consisted of a linear trend, $F(1, 11) = 9.00, p = .012, \eta^2_p = .45$, and a marginally significant quadratic trend, $F(1, 11) = 4.74, p = .052, \eta^2_p = .30$. In addition, a main effect of Colour Match was again present, $F(1, 11) = 35.36, p < .001, \eta^2_p = .76$, and Colour Match did not interact with Colour Proportion, $F(2, 22) = 1.30, p = .29, \eta^2_p = .11$. A 3X2 repeated measures ANOVA on search accuracy showed no main effects or interactions, $Fs < 0.10, ps > .91, \eta^2_p < .01$.

These results of Experiment 4 were qualitatively identical to Experiment 1, supporting the notion that confirmatory search was not an artifact of a lack of awareness of proper strategy. Although the results suggest that a quadratic relationship between the number of template-colour matching stimuli and search time was present, the lack of an interaction between Colour Match
and Colour Proportion again indicates that participants were not consistently switching templates when the template-colour matching stimuli outnumbered the template-colour matching stimuli. As outlined in the discussion of Experiment 3, I suggest that this may reflect the contribution of some economical searches, which become possible once the template becomes learned through use.

Further supporting the conclusion that explicit strategy did not reduce confirmatory searching is the observation that only five of twelve participants reported using the minimal search strategy when debriefed, despite having been informed of it at the outset. Indeed, reported search strategy did not reliably interact with any factors for either RT, $F_s < 0.63$, $p_s > .55$, $\eta^2_{ps} < .06$, or for accuracy, $F_s < 0.71$, $p_s > .42$, $\eta^2_{ps} < .07$.

![Figure 5. Median Response Times (left) and Mean Accuracy (right) for the search task in Experiment 4.](image)

Given that explicitly instructing participants to use a particular strategy did not affect their search performance, I combined the data from Experiment 1 and 4, where the procedure had been otherwise identical, to provide a more powerful analysis of the effect of learned strategy on search behavior. Although the cost of reporting template-colour mismatching targets was slightly attenuated for those reporting a minimal search strategy, $F(1, 22) = 4.05$, $p = .06$, $\eta^2_{ps} < .16$, as in Experiment 1, all other interactions were still unreliable for both search time, $F_s < 1.29$, $p_s > .29$, $\eta^2_{ps} < .06$, and for accuracy, $F_s < 2.43$, $p_s > .13$, $\eta^2_{ps} < .10$. If anything, it seems that a reported minimal search strategy manifests only in a reduced RT and accuracy cost associated with finding the template-colour mismatching target. This demonstrates that, at least
in this task, search performance and metacognitive strategy are dissociable; participants appear to know how to complete the task most efficiently, but do not behave in accordance with this approach.

8 Experiment 5

Thus far, my results have consistently provided evidence for a confirmatory search bias. The effect is largely insensitive to the presence or lack of repetitions of template use, as well as knowledge of the task. In Experiment 5, I tested whether the confirmatory search bias could be reduced by presenting a preview of the colour of search stimuli in advance of the search. Given the robustness of confirmatory selection observed thus far, it is tempting to conclude that confirmatory selection of stimuli matching a template is a relatively automatic process. In previous experiments, participants often did report using a minimal selection strategy, even though their search times told a different story. It may be the case that the strategic guidance of search lags behind the more automatic orienting towards template matching stimuli. An alternative to automatic guidance to template matching stimuli that one could reasonably expect is for automatic orienting to be towards the fewest, and therefore most perceptually salient stimuli. However, at least for my search task, confirmation bias seems to be the default tendency that must be overcome.

To test this possibility, I presented a colour preview in advance of the search stimuli on every trial. My reasoning is that, if given the chance to observe the statistics of the colours while not having the ability to begin searching, participants might more appropriately plan their search in advance. Relatedly, Kunar, Flusberg, & Wolfe (2006) found that color-cues only successfully oriented attention to target locations with explicit knowledge and a 1500ms lead-time. If strategic control of selection is simply slower than template-guided selection, I predict a quadratic trend between search time and the proportion of Colour Proportion, and an interaction between Colour Match and Colour Proportion.

8.1 Methods

8.1.1 Participants

Twelve undergraduate students were recruited for the present experiment. All participants provided informed consent and were compensated with course credit.
8.1.2 Stimuli and Procedure

The stimuli and procedure were very similar to Experiment 1; 16 blocks of 30 trials were again implemented, and a search template was provided prior to each search block, but search stimuli themselves were slightly changed. For each search, a colour preview display was presented for 1000 ms in which coloured squares, approximately $1.2^\circ \times 1.2^\circ$, appeared centered on the positions of their respectively coloured search stimuli. After 1000 ms had elapsed, the letters used as search stimuli onset in front of the coloured squares. These letters were uniformly coloured in white. Instructions were changed accordingly, such that participants were now asked to respond regarding whether the target letter was on a particular colour.

8.2 Results and Discussion

Preliminary RT analyses showed a number of outlying trials, consisting of both suspiciously long search times (>10s, 0.013% of all trials), and anticipatory responses (<100ms, 0.06% of all trials). Trials with search times falling outside of either of the aforementioned bounds were excluded before conducting the following analyses. The extended colour preview display led to a change in the pattern of search RT (Figure 6), but this change was also accompanied by changes in search accuracy. A 3X2 repeated measures ANOVA on RT revealed a main effect of Colour Proportion, $F(2, 22) = 17.05, p < .001, \eta^2_p = .61$, which was comprised of a linear, $F(1, 11) = 8.70, p = .013, \eta^2_p = .44$, and quadratic, $F(1, 11) = 46.97, p < .001, \eta^2_p = .81$, trend. The effect of Colour Proportion was accompanied by a main effect of Colour Match, $F(1, 11) = 15.62, p = .002, \eta^2_p = .59$, but no interaction was observed, $F(2, 22) = 2.34, p = .12, \eta^2_p = .18$. 
The RT data alone is suggestive of a flexible selection strategy, but a repeated measures ANOVA on accuracy revealed speed-accuracy trade-offs. A main effect of Colour Match, with template-colour matching targets being reported with lower accuracy, approached significance, $F(1, 11) = 4.41$, $p = .06$, $\eta^2_p = .29$. In addition, Colour Proportion decreased search accuracy, $F(2, 2) = 13.88$, $p < .001$, $\eta^2_p = .56$, in a monotonic fashion, $F(1, 11) = 37.38$, $p < .001$, $\eta^2_p = .77$, such that search responses were less accurate as more stimuli matched the template colour. In addition, Colour Match and Colour Proportion interacted, $F(2, 22) = 3.71$, $p = .04$, $\eta^2_p = .25$, such that the difference in accuracy between Colour Match conditions increased as the proportion of template-matching colours increased, $F(1, 11) = 6.71$, $p = .025$, $\eta^2_p = .38$, with accuracy suffering most when the target appeared in the template matching colour. Combined with the increase in the number of participants reporting a minimal search strategy (10 of 12), these results suggest that while participants improved their search speed on trials when the task demands encouraged prioritizing the colour not mentioned in the instructions, this also led to more search errors. Participants may have opted to switch templates for these searches, but the increased cognitive load of these switches led to more errors at the response planning stage.

To further clarify the effect of colour previews on search strategy, I supplemented these analyses by computing efficiency scores (mean accuracy divided by median response time for correct and incorrect trials). Efficiency was greatest with 2 of 8 stimuli matching the template colour ($M_{\text{Colour Match}} = 1.23$, $SE = 0.10$; $M_{\text{Colour Mismatch}} = 1.10$, $SE = 0.09$), worst with 4/8 matching
the template colour ($M_{\text{Colour Match}} = 0.79, SE = 0.06; M_{\text{Colour Mismatch}} = 0.68, SE = 0.04$), and slightly better with 6/8 matching the template colour ($M_{\text{Colour Match}} = 0.81, SE = 0.08; M_{\text{Colour Mismatch}} = 0.88, SE = 0.10$). Estimates of efficiency were affected by Colour Proportion, $F(2, 22) = 74.95, p < .001, \eta^2_p = .51$, with both linear, $F(1, 11) = 27.31, p < .001, \eta^2_p = .71$, and quadratic, $F(1, 11) = 40.97, p < .001, \eta^2_p = .79$, components. Although Colour Match was only marginally significant, $F(1, 11) = 4.23, p = .062, \eta^2_p = .28$, it critically interacted with Colour Proportion, $F(2, 22) = 7.98, p = .002, \eta^2_p = .42$. This interaction was such that, when the number of template-matching colours was greater than the number of template-mismatching colours, template-colour matching targets were less efficiently detected than template-colour mismatching targets, suggesting that participants indeed did adapt their search strategy to the proportion of colours in a search display. However, that this effect was accompanied by a linear effect of Colour Proportion suggests that the strategic search strategy coexisted with a confirmatory tendency. Searches where confirmatory searching was strategically optimal were overall more efficient than trials where it was not, as indicated by a pairwise comparison between 25% template-matching colour trials and 75% template-matching colour trials, $F(1, 11) = 27.31, p < .001, \eta^2_p = .71$. Nonetheless, the previews did alter the extent of confirmatory searching; a Mixed Model ANOVA comparing efficiency scores between Experiments 1 and 5 showed that Experiment interacted with the effect of Colour Proportion, $F(2, 44) = 15.88, p < .001, \eta^2_p = .42$, as well as with the interaction between Colour Proportion and Colour Match, $F(2, 44) = 4.12, p = .023, \eta^2_p = .16$. Overall, these results show that colour preview displays can attenuate confirmatory searching, but not completely.

9 General Discussion

In the present chapter, I measured visual search performance in a novel task designed to examine the use of confirmatory search strategies. Participants were asked to report whether a target letter was a target colour or not, and across searches I varied the proportion of stimuli in search displays that were the target colour or non-target colour. Despite the fact that this search task allowed participants to adopt a strategy of target disconfirmation – that is, they could examine the stimuli in a non-target colour – the target confirmation strategy dominated performance. I conclude, therefore, that a confirmation bias exists even in simple visual search tasks. This conclusion is supported by the conjunction of two general findings; first, that participants were slower to report the target identity when proportionally more search stimuli
matched the target template, as defined in search instructions, even though a more economical selection strategy was available, and second, that participants were faster to report the target identity when the target was the colour of the template than when the target was another colour.

The first finding is reminiscent of studies of subset search (e.g., Bacon & Egeth, 1997), where participants perseverate in selecting stimuli possessing a particular guiding feature if instructed, even when this feature is present in the larger of two equally useful search subsets. While subset search research speaks to the strength of instructions in dictating top-down selection, these results cannot, by themselves, demonstrate confirmatory searching, as no other search strategy (i.e., disconfirmation of target presence) is available in these tasks as a viable alternative to confirmatory searching. Similarly, the finding of slower search times for targets not matching the template may be considered to be an instantiation of the classic finding of slower searches for targets defined by the absence of a feature (Wolfe, 2001). If this comparison is valid, then my results provide a demonstration that feature absence may be relative to task set. That is, the particular stimulus that would be considered “feature-present” and “feature-absent” in my task was a consequence of an arbitrary assignment of one of two colours in the search instructions.

The results of these experiments converge on the conclusion that the default, or preferred, search strategy is one in which searcher prioritizes stimuli that share features with a target template, and opts to determine the status of a target by matching it to the template rather than by switching to a disconfirmation strategy of searching for stimuli that would provide evidence against the presence of a template-matching target. My results therefore rule out the possibility that confirmatory searching is task-contingent selection strategy. In my search task, matching search stimuli to a single colour-letter conjunction entailed conducting more analysis than strictly necessary to complete a search. However, search times indicated that this is how participants opted to search. It is important to note that, because my data rely only on overall response time, increases in response time caused by increases in the proportion of template-matching stimuli may not simply reflect increases in the total number of stimuli inspected in search, but may also reflect increases in time spent processing the colour statistics in the display to plan searches, updating templates, processing individual stimuli, or selecting responses. Additional stimulus manipulations and within-trial search metrics (e.g., eye tracking) are needed to resolve this uncertainty, and will be supplied in the forthcoming chapters. Nonetheless, from a purely
performance based perspective, one may still conclude that visual search is successfully terminated faster when a target’s presence is confirmed, not disconfirmed.

Confirmatory search may stem from a number of underlying sources. The first possibility is, as has been suggested before, that visual search can only be guided by one template at a time. A number of studies investigating the control of visual search guidance by representations in visual working memory have demonstrated that only one representation appears to be prioritized to guide search at a time (reviewed in Olivers et al., 2011). Although contrary findings exist (Beck, Hollingworth, & Luck, 2013; Irons, Folk, & Remington, 2012), a sufficiently sophisticated notion of search templates, such as the Boolean Map Theory of Visual Selection (BMTVS) can accommodate guidance by multiple features, but via a single template. In the BMTVS, multiple features may be combined using Boolean (conjunctive and disjunctive) operations, with the critical consequence that any stimuli selected using a given Boolean setting cannot be distinguished from each other on the selection dimension (e.g., colour); further template adjustments would need to be made in order to distinguish these selected stimuli from each other on any a particular dimension. In my selection task, because colour is a dimension that is necessary to select the appropriate response, BMTVS predicts that only a single colour can be used to guide selection and analysis, because colour is necessary for deciding between responses in addition to selecting potential target stimuli. A single template architecture introduces costs associated with updating the target template to the appropriate template for a particular display. The costs associated with calculating and updating to the appropriate template to use may simply outweigh the benefits of updating in terms of overall search time. Thus, capacity limitations in search template guidance from working memory are a potential culprit in the source of confirmatory searching. As suggested earlier, a similar limitation has been suggested in reasoning (Mynatt, Doherty, & Dragan, 1993), which requires search for information through a possibility space.

On the other hand, confirmatory searching in my task may be due to difficulties in guiding search with negative information (see Arita, Carlisle, & Woodman, 2012). In the instructions, I framed the search such that the target could either be one particular colour, or not that particular colour. While the search performance appears to reflect a confirmatory, template matching stimulus prioritization, it may be that search cannot be strategically guided by negative information (e.g., “not blue”). As noted earlier, search for absent features is well-known to be
difficult. If it is the case that my mere framing of the non-template colour as the absence of the template colour was sufficient to recode the non-template colour as an absent feature, this may account for why selection was preferentially guided towards the template colour. In this case, the template colour would have been treated as a “present” feature, and therefore would lead to easier selection, making the perseveration of selection on this colour optimal for participants. While plausible, this account would require an additional interpretation of feature-absent effects: until now, these effects have been taken to reflect a property of the visual system’s coding, as opposed to task demands. The nearest approximation of a cognitive, rather than perceptual, interpretation of feature-absent effects that I am aware of stems from work on familiarity as a feature in visual search (e.g., Wang, Cavanagh, & Green, 1994; Shen & Reingold, 2001). These studies have shown that stimuli whose low-level visual properties are otherwise identical interact with search efficiency depending on whether they are meaningful stimuli: finding an unfamiliar stimulus (a rotated letter) among familiar stimuli (un-rotated stimuli) is easier than finding a familiar stimulus amongst many unfamiliar stimuli. If the negative-framing of the search task in my instructions is indeed the reason for confirmatory searching, then I will have incidentally provided a demonstration of the top-down construction of what defines a feature in visual search.

Although in this task, participants performed in a way that was not strategically optimal, confirmatory searching is likely a globally optimal strategy for visual search. For falsification to be an economical strategy requires some features or stimuli exist that are negatively correlated with the presence of whatever is being searched for. When target presence and absence is independent of other environmental features, only confirmation of the target’s presence can reduce search times compared to an exhaustive search. In light of arguments that visual search may be optimized for foraging (Klein & MacInnes, 1998; Cain, Vul, Clark, & Mitroff, 2012), confirmatory searching would prove beneficial, in that the analysis of the environment would be tailored towards the goal of finding any extant resources, and promote the sustained pursuit of a goal even in situations where positive evidence is scant. Friedrich (1993) has argued for a related basis of confirmation bias in reasoning, noting that different types of errors produce more or less costs in the context of particular goals, and that it is more pragmatic to minimize costly errors than to simply minimize all errors, irrespective of their consequences. In the context of my visual search task, additional covert costs – such as switching templates in working memory – may simply be more costly than the additional time spent searching in those cases where more stimuli
match target templates. More broadly, expending cognitive (and motoric, in the case of saccades) resources to sustain a purely visual search for signs of prey is a relatively low cost investment, given the potential payoffs. Mechanisms in visual search, therefore, may be tuned to allow perseveration on the possibility that a real resource is indeed present, so that visual inspections can be sufficiently thorough. Cognition in general may be seen as a (relatively) biologically cheap way of tuning our actions in advance so as to acquire proportionally greater survival gains.

Speculations aside, the possibility that confirmation bias in visual search results from the low cost of information acquisition relative to the costs of computing the most efficient search strategy is a hypothesis that lends itself to testing. Chapter 3 will thus test the hypothesis that confirmatory search results from the ease of conducting, relative to planning, a search.
Chapter 3
The Price of Information

In Chapter 2\(^3\), my novel visual search task revealed that visual attention is biased towards information that might confirm a visual hypothesis, as 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. Thus, 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. I (Rajsic et al.), however, did not directly test the cost-benefit account of confirmatory searching.

In this third chapter, I 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, I 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, I measured searches in three experiments that varied the dynamics of inspections used to scan search displays. Experiment 6 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 7 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 8, I 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. I predicted that, as inspection costs increased from Experiments 6 to 8, I would observe a complimentary reduction in the confirmation bias in visual search. In Experiment 9, I 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.

10 Experiment 6

My goal for Experiment 6 was to replicate the design of Chapter 2 (Rajsic et al., 2015) with the addition of eye-tracking, and with the slightly modified stimuli that were to be used in Experiment 7’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, I expected to replicate my 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, I expected to find that more saccades would be made to Template Matching stimuli, especially early in search.

10.1 Methods

10.1.1 Participants.

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

10.1.2 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).

10.1.3 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 “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, I 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.

10.2 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.

I first analysed median correct search times to assess whether search exhibited a confirmation bias. These search times are depicted in Figure 7a. Search RT overall increased with Template Matching Subset Size, linear contrast: $F(1, 11) = 18.93$, $MSE = 1.66$, $p = .001$, $\eta^2 = 0.15$, 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 Chapter 2 (Rajsic et al., 2015 in showing a confirmation bias in visual search.

![Figure 7](image.png)

**Figure 7.** Panel A depicts median correct search times in Experiment 6, and Panel B depicts the average number of fixations per search. Error bars depict one within-subjects standard error.

Given that I collected eye movement data in Experiment 6, I 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 my 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. Sample scanpaths can be seen in Figure 8. I 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 7b. 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.
Figure 8. Sample inspections for Experiments 6, 7, and 8. Trials depicted are Template Mismatch trials for Matching Subset Size 2 (upper row) and 6 (lower row). Yellow arrows indicate paths between successive fixations (Experiments 6 and 7), and green arrows indicate paths between successive stimulus-reveals (Experiments 7 and 8). In each case, inspections for one randomly selected trial are plotted.

I 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 I assessed a bias towards confirmatory stimuli, I 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, I used a guessing correction of $p(\text{Bias}) = (p(\text{Obs}) - p(\text{Chance}))/\left(1 - p(\text{Chance})\right)$, where $p(\text{Obs})$ was the measured probability of inspecting the Template Matching colour and $p(\text{Chance})$ was 0.25, 0.5, and 0.75.
for the Matching Subset Sizes 2, 4, and 6, respectively. Importantly, when \( p(\text{Obs}) \) was below \( p(\text{Chance}) \), \( p(\text{Chance}) \) was adjusted to the proportion of Template Mismatching colours in the display. The resulting stimulus inspection tendencies are plotted in Figure 8.

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, \( F(1, 10) = 16.36, MSE = 0.85, p = .002, \eta^2 = .07 \), but was qualified by an interaction with Search Epoch, \( F(1, 10) = 12.58, MSE = 0.39, p = .005, \eta^2 = 0.03 \). 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, I also analysed the duration of first inspection on trials where the target was not the first fixated item. This allowed me 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, \( F(1, 11) = 8.72, MSE = 10247 p = .014, \eta^2 = .01 \), such that Template Matching Stimuli were inspected for more time, \( M = 221 \text{ ms}, SE = 7 \text{ ms}, \) than Template Mismatching Stimuli, \( M = 203 \text{ ms}, SE = 8 \text{ ms} \). All other factors and interactions did not reliably affect first inspection durations, \( Fs < 1.87, ps < .18, \eta^2$s < .004.\)
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 6. 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 my 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, I calculated the proportion of trials where the target was inspected before a correct response was given. I 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, I 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 9.

![Figure 9](image)

**Figure 9.** Proportion of trials where targets were fixated before being correctly identified in Experiment 6. 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, $Fs < 1.35, ps > .29$, or Selection Bias effects, $Fs < 2.99, ps > .19$.

While the response time data reported here and in Chapter 2 (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 I 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 my 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.

11 Experiment 7

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, I 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 I predicted a shift towards more strategic, and less confirmatory, searching.

11.1 Method

11.1.1 Participants.

As in Experiment 6, 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.

11.1.2 Stimuli and Procedure.

The task, stimuli, and procedure for Experiment 7 were identical to Experiment 6 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 6) 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 6, 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.

11.2 Results and Discussion

To briefly preview the results of Experiment 7, 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, $F_s \leq 1.56$, $ps \geq .23$. These search times are depicted in Figure 10a. A visual inspection reveals that, unlike Experiment 6, the effect of Matching Subset
Size, $F(2, 22) = 22.34, \text{ MSE} = 4.95, p < .001, \eta^2 = 0.18$, was not monotonic. Instead, Matching Subset Size produced a mixture of linear and quadratic trends, $Fs > 14.09, ps < .003, \eta^2s > 0.07$, 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, $F(1, 11) = 7.44, \text{ MSE} = 1.01, p = .02, \eta^2 = 0.03$, although a marginal interaction was also observed, $F(2, 22) = 3.42, \text{ MSE} = 0.49, p = .051, \eta^2 = 0.02$. 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 11](image-url)

**Figure 11.** Panel A depicts median correct search times in Experiment 7, and Panel B depicts the average number of fixations per search. Error bars depict one within-subjects standard error.

As in Experiment 6, I 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 10b. 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 6.

To assess the selectivity in search, I 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 11.

![Figure 11](image)

**Figure 12.** 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 7. 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.

I 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, \( F(1, 9) = 9.33, MSE = 0.46, p = .014, \eta^2 = 0.38 \), 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 6 and 7 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 7, \( t(19) = 2.43, p = .025 \), indicating that gaze-contingent searching led to the strategic allocation of attention towards the Mismatching colour stimuli, unlike in Experiment 6. In contrast to Experiment 6 as well, an analysis of first inspection durations of distractors revealed no main effect of Stimulus Type, \( F(1, 9) = 2.44, MSE = 196251, p = .15, \eta^2 = 0.02 \), but rather an interaction between Stimulus Type and Template Matching Subset Size, \( F(1, 9) = 5.19, MSE = 108582, p = .017, \eta^2 = 0.02 \). 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 my 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 warrant by the distribution of coloured stimuli in the display. However, comparing the time between search onset and first inspections between Experiments 6 and 7 yielded no reliable differences, \( Fs < 2.20, ps < .17 \). The first inspection times at Matching Subset Size 6 for Experiments 6 and 7 were \( M_{\text{Exp1}} = 404\text{ms}, SE_{\text{Exp1}} = 27\text{ms} \) and \( M_{\text{Exp2}} = 416\text{ms}, SE_{\text{Exp2}} = 27\text{ms} \), respectively. If the improved selection strategy seen in Experiment 7 occurs due to longer search planning and template updating, then it would appear that this additional planning only requires approximately 12ms.
Lastly, I again analysed the likelihood of fixating the target stimulus before providing a correct response. These data are plotted in Figure 12. While target fixation probability showed a main effect of Target Colour, $F(1, 9) = 7.69, \text{MSE} = 0.62, p = .02, \eta^2 = 0.05$, an interaction was observed, $F(2, 18) = 21.17, \text{MSE} = 0.84, p < .001, \eta^2 = 0.15$. 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 13. Proportion of trials where a correct response was given and the target was inspected before search termination in Experiment 7. 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 7 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 7 suggests confirmation bias can be reduced when the costs of accessing information are increased. In Experiment 8, I provide a stronger test of this proposal by introducing additional inspection costs.

12 Experiment 8

In order to test whether searches are more efficient when the costs of inspections are increased, I conducted a third experiment where these inspection costs were further increased. In this experiment, I 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, I expected that the change in guidance seen between Experiments 6 and 7 would be further exaggerated in Experiment 8.
12.1 Method

12.1.1 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.

12.1.2 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.

12.2 Results and Discussion

Overall, the results of Experiment 8 mirrored those of Experiment 7; strategic stimulus selection of smaller subsets as revealed by search RTs, number of inspections, colour selectivity, and target inspection probability. Comparisons between Experiments 7 and 8, however, revealed that the extent of strategic selection was amplified by using mouse-contingent search. Overall search accuracy was high in Experiment 8, \( M = 92.8\% \), \( SE = 2.1\% \), but was affected by Matching Subset Size, \( F(2, 22) = 7.03, \text{MSE} = 0.03, p = .004, \eta^2 = 0.09 \), and the combination of Target Colour and Matching Subset Size, \( F(2, 22) = 5.71, \text{MSE} = 0.006, p = .01, \eta^2 = 0.02 \).

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, \( M_{\text{match}} = 91.9\% \), \( SE_{\text{match}} = 2.4\% \), \( M_{\text{mismatch}} = 86.1\% \), \( SE_{\text{mismatch}} = 4.4\% \). 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 13a. These search times showed, like Experiment 7, that searches were more strategic. Matching Subset Size, \( F(2, 11) = 30.72, \text{MSE} = 5.72, p < .001, \eta^2 = 0.33 \), had a non-monotonic effect on search, with both a linear, \( F(1, 11) = 16.21, \text{MSE} = 2.92, p = .002, \eta^2 = 0.09 \), and a quadratic, \( F(1, 11) = 44.32, \text{MSE} = 8.51, p < .001 \),
η² = 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, $F(1, 11) = 7.08$, $MSE = 2.86$, $p = .022$, $η² = 0.08$, but was accompanied by an interaction, $F(2, 22) = 3.43$, $MSE = 0.46$, $p = .05$, $η² = 0.02$. I 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, $t(9) = 2.62$, $p = .024$, $M_{\text{match}} = 1950ms$, $SE_{\text{match}} = 62ms$, $M_{\text{mismatch}} = 2444ms$, $SE_{\text{mismatch}} = 196ms$, and Matching Subset Size 4, $t(9) = 3.37$, $p = .006$, $M_{\text{match}} = 2877ms$, $SE_{\text{match}} = 153ms$, $M_{\text{mismatch}} = 3470ms$, $SE_{\text{mismatch}} = 177ms$, but not at Matching Subset Size 6, where no difference was observed, $t(9) = 0.54$, $p = .60$; $M_{\text{match}} = 2637ms$, $SE_{\text{match}} = 92ms$, $M_{\text{mismatch}} = 2744ms$, $SE_{\text{mismatch}} = 200ms$. These results parallel Experiment 7 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 14](image.png)

**Figure 14.** Panel A depicts median correct search times in Experiment 8, and Panel B depicts the average number of fixations per search. Error bars depict one within-subjects standard error.

As with Experiments 6 and 7, I analysed the dynamics of search using three metrics: total inspections, bias towards Template Matching stimuli, and likelihood of target inspection. For the first metric, I 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 13b.

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 8, and show a strategic, rather than confirmatory, search strategy. To provide a direct comparison, however, I 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, $Fs \leq 0.95, ps \geq .40, \eta^2s \leq 0.003$. 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, $M_{mouse} = 3.11, SE_{mouse} = 0.21, M_{gaze} = 5.10, SE_{gaze} = 0.23$.

I next analysed the selectivity bias, calculated using inspections, which is plotted in Figure 14. 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 I 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, $Fs \leq 1.15, ps \geq .31, \eta^2s \leq 0.01$,
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, $F_{s} \leq 2.93, ps \geq .12, \eta^{2}_{s} \leq .008$. 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 15.** 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 8. 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, I examined the likelihood of correctly completing a search after visually inspecting the target, which is plotted in Figure 15. 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, \( t_{11} \leq 3.41, p_s < .006 \), 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 8 than in Experiment 7, 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 my 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 16. Proportion of trials where a correct response was given and the target was inspected before search termination in Experiment 8. 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.](image)

Although the results of Experiment 8 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 7, I used a gaze-contingent search to limit the perceptual information, which I expected to increase the costs of poorly planned search inspections in terms of lost information (from the visual periphery). In Experiment 8, I 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: $M_{Exp1} = 300\text{ms}$, $SE_{Exp1} = 53\text{ms}$, $M_{Exp2} = 718\text{ms}$, $SE_{Exp1} = 133\text{ms}$, $M_{Exp3} = 861\text{ms}$, $SE_{Exp3} = 100\text{ms}$. 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 my findings. To test this possibility, a final experiment was conducted.

13  **Experiment 9**

Experiment 9 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, I introduced intermittent masks into the search display, which controlled the amount of time that target-defining information was visible. By doing so, I 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, I 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, I would expect that high Information rate displays
would lead to confirmatory searching and low Information rate display would lead to strategic searching.

13.1 Method

13.1.1 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.

13.1.2 Stimuli and Procedure.

The stimuli were identical to those of Experiment 6 with the following exception. Where in Experiment 6 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 in Figure 16. 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 17. An example illustration of the stimuli and procedure used in Experiment 9. 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).

13.2 Results and Discussion

The overall results of Experiment 9 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, $F$s > 22.44, $p$s < .001, $\eta^2$s > .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 7 and 8 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, I 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 17). 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 18. 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.

14 General Discussion

In Chapter 2, I showed that, in a multiple-target conjunction search, search is biased towards whichever target conjunction is framed as the search template, which I 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, I suggested that the cognitive costs of updating guidance on each trial may outweigh the costs of over-searching a display. My goal in the present chapter 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 6, I replicated my 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 7 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 8, 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 9, by controlling the rate of information availability during searches, I 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 6 is that, even in standard visual search, a mixture of the two search strategies was evident. As stated earlier, this likely accounts for my 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, my 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. My 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 my 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, I 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 7 and 8 – 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 my 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, my 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, I believe, among the strongest examples of a top-down search strategy directed by a factor outside of performance incentives.

The reduction in confirmatory searching with increased inspection costs points to the possibility that the type of guidance used 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, & Merkle, 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 9 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 my task (Chapter 2; Rajsic et al., 2015). Across the current four experiments, instructions and stimuli remained similar, and I 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 7 and 8 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, I 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.

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, I am not aware of any models of search that could account for the results of Experiment 8, given that the critical difference was non-visual (the effector used to reveal information), or Experiment 9, where the temporal dynamics of to-be-searched stimuli affected guidance.

The temporal dynamics of confirmatory search can have, as I 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 colour,
but not the non-template colour, is presented, could produce the measured bias via priming through selection history (Awh, Belopolsky, & Theeuwes, 2012; Theeuwes, Reimann, & Mortier, 2006; Krouijne & Meeter, 2016). A third option, which I 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 Chapter 2 (Rajsic et al., 2015), I found confirmatory searches both when a single search was performed per template and when one template was used for all searches. In addition, I found that self-reported strategy did not relate well to the strategy revealed from search RT analyses. However, these findings are nonetheless compatible with a bottom-up priming explanation. The broader issue, however, is whether confirmation bias occurs due to relatively local, low-level attentional dynamics, or a more abstract preference for any visual information that matches a perceptual hypothesis. In the following chapter, I address the question of how concrete, or how abstract, the perceptual templates underlying confirmatory search are.
Chapter 4
Matching Bias

In the search task used in the preceding chapters, I have endeavoured to isolate the contribution of subjective perceptual expectations by using two equiprobable colours and two equiprobable targets, thus ensuring that search biases could not result from statistics of the perceptual input. The critical manipulation is in the task instructions: by requiring a “yes” response to one coloured target and a “no” response to another coloured target, I sought to create a “focal” perceptual hypothesis. What is it about these instructions in my search task that leads to selection biases? One possibility is that the instructions bias search because they present participants with a specific visual input that matches one of the stimulus colours. For example, in Chapter 1, I consistently instructed participants using a coloured rectangle to depict the positive template. Thus, one possibility is that confirmatory searching results from simple, bottom-up intra-trial priming of the confirmatory colour (e.g., Theeuwes, Reimann, & Mortier, 2007).

Another possibility is that mentioning one of two possible target features in the instructions primes categorical attentional guidance processes. Guided Search, for example, proposes that the selection of relevant coloured stimuli in a search array depends on broadly tuned, categorical colour channels (Wolfe, Cave, & Franzel, 1989; Wolfe, 2007). A categorically tuned architecture is ideal for top-down control, given that goals of a search would often begin with a linguistic code in everyday situations (e.g., saying to a friend “that blue car looks expensive”), but especially in the context of psychology experiments where participants are instructed with written or spoken guidelines. If the confirmation bias results from a heuristic matching process between elements named in the instructions and this categorical guidance apparatus, then the confirmation bias should be observed when templates are specified only using words, not visual depictions. Experiment 10 tests this account against the possibility that confirmatory search biases are due to bottom-up priming.

If visual attention is truly attracted to confirmatory stimuli, confirmation biases should extend beyond situations in which stimuli match a particular template on a single, explicitly

---

mentioned, homogenous visual feature. Instead, stimuli should attract attention because of their ability to verify a proposition per se, even when this proposition involves more abstract classes of stimuli. Although searching for red stimuli when asked whether a target is red or not could reflect a preference to find information that would yield an affirmative answer – a true confirmation bias -- it could also be due to a heuristic of relevance, such that stimulus features mentioned in the rule are heuristically deemed more important, or informative (Sperber, Cara, & Girotto, 1995). Experiments 11 and 12 were conducted to distinguish true confirmatory search from a relevance heuristic by measuring whether biases occur when confirmatory stimuli are defined using negation (Experiment 11) and when confirmatory stimuli are visually heterogeneous (Experiment 12).

15 Experiment 10

Experiment 10 was conducted to determine whether confirmation biases in visual search are mere instances of bottom-up priming of visual features or whether they can occur when a template is described verbally. To do so, I adapted the methods and stimuli from Chapter 1 (Rajsic et al., 2015). Participants were instructed that, on each trial, they should evaluate whether a particular question about the display should be answered in the affirmative or negative. Specifically, all trials asked whether a particular letter was on a circle of a particular colour. Instead of using a coloured square to communicate the particular colour, as in Chapter 1, the present chapter used a verbal label for each colour (e.g., “red”). If participants search in a biased manner, they should preferentially search the template-matching (confirmatory) colour, resulting in increased search times when the template-matching group is more numerous. If participants search in a strategic manner – ignoring confirmation bias – they should preferentially search the colour with fewer circles on a trial-to-trial basis.

15.1 Methods

15.1.1 Participants

Sixteen undergraduate students volunteered to participate for course credit. All participants provided informed consent.
15.1.2 Stimuli

Search displays consisted of eight letters, presented on the circumference of an imaginary circle centered on a central fixation cross. Each letter in a search display was a lowercase p, q, b, or d, approximately 2° in height and 1° in width, and was drawn approximately 8° from fixation using Arial font drawn in black (RGB: 0,0,0). These letters were placed on top of small discs (approximately 1° in radius) whose colours were selected from a pool of seven possible colours; purple, yellow, green, orange, pink, blue, and red (RGB values, respectively: 200, 0, 255; 200, 200, 0; 0, 255, 0; 255, 128, 0; 255, 128, 255; 50, 50, 255; 255, 50, 50), with the background set as mid-gray (RGB: 128, 128, 128). Before beginning a block of trials, participants were presented with instructions written on the computer monitor in the following form: “For each trial, answer this question: “Is the x on a y circle?” Press key 1 if yes, press key 2 if no.” For a given instruction x would be the target letter (p, b, d, or q), y would be the categorical colour name, and keys 1 and 2 would refer to either the Z or X key, which were alternately used as either response. For example, as illustrated in Figure 18a, participants may have been prompted to respond as to whether the “p” was on a red circle, using the Z key for yes and the X key for no. Subsequent searches would include distractor letters on red and blue circles, with target p’s appearing either on a red or blue circle from trial to trial. These instructions remained on screen until participants chose to begin the corresponding block. Figure 18a depicts a sample instruction and search display (at Template-Matching Subset Size 4, with a Matching Target Colour).
For each trial, answer this question:
Is the p on a red circle?
Press Z if yes, Press X if no

For each trial, answer this question:
Is the p on a non-blue circle?
Press Z if yes, Press X if no

For each trial, answer this question:
Is the p on a colorful circle?
Press Z if yes, Press X if no

Figure 19. A sample search instruction (upper row) and sample search array (lower row) for Experiments 10, 11, and 12 (columns a, b, and c). Stimuli are not drawn to scale.

15.1.3 Procedure

One experimental session consisted of 12 blocks of 24 trials, where each block consisted of four repetitions of the six experimental conditions: Target Colour (Template Matching or Template Mismatching) X Template Matching Subset Size (2, 4, or 6). For a given block, two of the seven possible colours were selected randomly as the two search colours to be used for the subsequent 24 trials. Two conditions were manipulated: the Target Colour, which was Template-Matching if it matched the colour mentioned in the instructions and Template-Mismatching if it did not, and the Template-Matching Subset Size, which could be 2, 4, or 6 stimuli. The actual target colour on a given trial was equally likely to be Template-Matching and Template-Mismatching, regardless of Template-Matching Subset Size, and participants were informed of this overall pattern.
A given trial began with the presentation of a blank screen with a fixation cross for 2000ms. Following this period, the search display was presented until a response was given. After a response was entered, using either the Z or X key, written feedback about response accuracy (“Correct” or “Incorrect”) was displayed in the center of the screen for 2000ms. After feedback offset, the next trial began.

15.2 Results and Discussion

To determine whether confirmation bias occurred with bottom-up priming concerns removed, I analysed the effect of Template Matching Subset Size and Target Colour on median correct response times (RTs), where I expect a monotonic effect of Template Matching Subset Size if selection is biased towards template-confirming stimuli. Two participants were excluded for having either lower than 80% accuracy or average RT more than two standard deviations from the group mean (i.e., greater than 2890 ms). Both Template Matching Subset Size, $F(2, 26) = 73.46, p < .001, \eta^2_p = .85$, and Target Colour, $F(1, 13) = 51.51, p < .001, \eta^2_p = .81$ affected RT, as well as an interaction, $F(2, 26) = 10.60, p < .001, \eta^2_p = .45$. Follow up contrasts on Template Matching Subset Size showed a linear trend, $F(1, 13) = 86.78, p < .001, \eta^2_p = .87$, but only a marginally significant quadratic trend, $F(1, 13) = 3.71, p = .08, \eta^2_p = .22$. Median correct RT is shown in Figure 2. An analysis of accuracy revealed only a main effect of Target Colour, $F(1, 13) = 7.66, p = .016, \eta^2_p = .37$, such that Template Mismatching Targets were reported more accurately, $M = 95\%$. $SE = 1\%$, than Template Matching Targets, $M = 92\%, SE = 1\%$. 
To ensure that my effects were not due to speed-accuracy trade-offs, I calculated an efficiency score, mean accuracy divided by median response time, for each participant in each condition. Similarly to median correct RT, efficiency declined as Template Matching Subset Size increased, $F(2, 26) = 61.52, p < .001, \eta^2_p = .83, M_{SS2} = 0.91, M_{SS4} = 0.72, M_{SS6} = 0.61$. Efficiency was also lower for Template Matching Targets, $M = 0.80$, than Template Mismatching Targets, $M = 0.70$. Thus, the confirmatory search bias I observed was not due to a speed-accuracy trade-off.

Overall, these data show that confirmatory searching occurs even when template colours are not visually presented, but instead conveyed through language. Therefore it is not the case that confirmatory search biases are simply due to bottom-up visual priming from instructions. Rather, confirmatory templates can be formed verbally, implying a level of non-perceptual, semantic abstraction.
16 Experiment 11

Experiment 10 showed that search is biased towards information that could lead to an affirmative endorsement of a visual hypothesis; when search was framed as being about the presence of one target and not another, even though both targets were equally likely, stimuli matching the colour of the framed colour attracted attention. Critically, this occurred in the absence of any visual presentation of the target colour in the instructions, leading to the conclusion that confirmation bias in search is not due to visual priming, but may derive from categorical guidance mechanisms (e.g., Wolfe, Cave, & Franzel, 1989).

In Experiment 11, I sought to determine whether the confirmatory search bias is due to a more abstract coding of relevance. In the previous experiment, all stimuli that matched a template matched by virtue of having the same feature. In research on reasoning using the Wason Selection Task, a number of researchers have emphasized a distinction between truly confirmatory data selection, where data is selected because it could be consistent with the proposition being evaluated, and a relevance heuristic wherein the objects or classes mentioned in the proposition being evaluated are rendered more salient (reviewed in Evans, 1998). A common technique for dissociating these two possibilities is to introduce negation into the proposition being evaluated, so that the positive set is no longer explicitly mentioned (e.g., “If there is an A on the front of a card, there is not a 7 on the back” does not mention a particular stimulus as a true consequent). Thus, in this experiment I pursued the question of whether confirmatory search patterns result from a matching bias by including blocks where one stimulus colour was referred to by negation (i.e., in a block of red and blue stimuli, asking participants whether a target letter was on a “non-red” circle in lieu of “blue” circle\(^5\)). Notably, the visual stimuli in this experiment are identical to Experiment 11. Moreover, the information provided in the prompt is equivalent. The only difference is the negative condition. Thus, the visual information and the logical information available to observers in Experiments 10 and 11 are the same. The question is whether the negation disrupts observers’ ability to use the template to guide search. If confirmatory selection is based on the ability of stimuli to yield an affirmative response, then I should observe similar search patterns between the Standard and Negation

\(^5\) I thank Todd Horowitz for suggesting this experiment.
search conditions. However, if selection is due to a matching-bias, the Negation search RT will not increase as the Template-Matching Subset Size increases.

16.1 Methods

16.1.1 Participants

Nineteen undergraduate students were recruited for a second experiment. All participants provided informed consent and were compensated with course credit. Participants were run until the post-exclusion sample size of Experiment 10 (14) was reached after using the same exclusion criteria.

16.1.2 Stimuli and Procedure

Stimuli and Procedure were identical to Experiment 10, with the following exception: blocks were divided into two types. Standard blocks included instructions in the same format as Experiment 10, whereas, in Negation blocks, participants answered questions of the form “Is the x on a non-y circle?”. These blocks were presented in a random order, determined separately for each participant. Figure 18b depicts a sample negated instruction and search display (at Template-Matching Subset Size 4, with a Matching Target Colour).

16.2 Results and Discussion

Median correct RTs were analysed, with Target Colour, Template-Matching Subset Size, and Negation as factors. Five participants were excluded for having accuracy lower than 80% or average RT more than two standard deviations above the group mean (i.e., greater than 3520ms). Overall, both Target Colour, $F(1, 13) = 9.73, p = .008, \eta^2_p = .43$, and Template-Matching Subset Size, $F(1, 13) = 9.30, p = .001, \eta^2_p = .42$, affected search time. Critically, Negation interacted with both Target Colour, $F(1, 13) = 7.66, p = .016, \eta^2_p = .37$, and Template-Matching Subset Size, $F(2, 26) = 3.93, p = .032, \eta^2_p = .23$ (see Figure 3). As such, I analysed search performance separately for the Standard and Negation. Accuracy was not affected by any factors or their interaction, and so was not analysed further, $Fs < 1.93, ps > .17, \eta^2_p < .13$. 
Figure 21. Median Response Times Experiment 11 for Standard Prompts (left) and Negation Prompts (right).

For Standard trials, Target Colour affected correct search times, $F(1, 13) = 17.30, p = .001, \eta^2_p = .57$, as did Template-Matching Subset Size, $F(2, 26) = 23.23, p < .001, \eta^2_p = .64$, accompanied by an interaction, $F(2, 26) = 3.51, p = .045, \eta^2_p = .21$. Template Matching Subset Size showed significant linear, $F(1, 13) = 35.19, p < .001, \eta^2_p = .73$, and quadratic trends, $F(1, 13) = 7.30, p = .018, \eta^2_p = .36$. Follow-up paired t-tests, while search RT increased as Template-Matching Subset Size increased for 2 to 4 for both Template-Matching, $t(13) = 3.60, p = .003$, and Template-Mismatching Targets, $t(13) = 8.22, p < .001$, increases from Subset Size 4 to 6 did not lead to a significant increase in search RT for Template-Matching, $t(13) = 0.04, p = .97$, or Template-Mismatching Targets, $t(13) = 1.83, p = .09$. However, given that the search RT was faster for Template Matching Targets than Template Mismatching targets at Subset Size 6, participants showed an overall confirmatory search tendency.

For Negation trials, neither factor, nor their interaction, affected search RT, $Fs < 1.02, ps > .37, \eta^2_p = .07$. At the end of each experimental session, participants reported their search strategies. Those who reported that, when shown a Template-Matching Subset Size 6 display, they would choose to first inspect a Template-Mismatching Target were classified as “strategic” searchers, whereas those who reported that they would choose to first inspect a Template-Matching Target (despite the larger Subset Size) were classified as “confirmatory” searchers. Overall, seven participants were classified as confirmatory searchers, and seven were classified as strategic searchers. However, an analysis of Negation trials showed that Search Strategy did
not interact with Template-Matching Subset Size, $F(2, 24) = 0.06, p = .94, \eta^2_p = .005$, Target Colour, $F(1, 12) = 2.40, p = .15, \eta^2_p = .17$, nor their combination, $F(1, 12) = 0.04, p = .96, \eta^2_p = .003$. The same was true for Standard trials, $Fs < 0.35, ps > .63$; reported search strategy did not modulate the search strategy indicated by search RT.

One reason that the Negation condition may not have shown confirmatory searching is due to an asymmetry in information between these conditions. In the Standard condition, the colour of the implied template was mentioned in the rule, whereas in the Negation condition, only the colour of the implied non-template was mentioned. As such, participants may have searched in a confirmatory manner once they knew the implied template’s colour; that is, later in a given block. To assess this possibility, I analysed search performance for both Standard and Negation trials with the additional factor of Block Half (first vs. last). For Standard trials, Block Half showed no main effect, $F(1, 13) = 3.13, p = .10, \eta^2_p = .19$, nor interactions, $Fs < 1.68, ps < .21, \eta^2_p s < .12$, with Template Matching Subset Size or Target Colour. On the other hand, in Negation trials, the interaction between Block Half and Template Matching Subset Size affected RT, $F(2, 26) = 3.77, p = .037, \eta^2_p = .23$, and Accuracy, $F(2, 26) = 4.86, p = .016, \eta^2_p = .27$. In the first half, search RTs were notably longer for Matching Subset Size 4 and 6, $Ms = [1977ms, 1937ms], SEs = [183ms, 176ms]$, compared to 2, $M = 1779ms, SE = 151ms$. In the second half, however, RTs were very similar across all Matching Subset Sizes, $M_{[2, 4, 6]} = [1844ms, 1839ms, 1799ms], SE_{[2, 4, 6]} = [150ms, 135ms, 117ms]$. As such, there is a suggestion of confirmatory searching with Negation instructions, but certainly it is not as clear or consistent as Standard instructions.

Overall, the results of Experiment 11 show that confirmatory search biases disappear when the goals of search are framed using negation. Indeed, neither Template-Matching Subset Size nor Target Colour affected search patterns when the target question included a negation. This suggests that no colour-based selection occurred in this case. It is, however, difficult to distinguish this possibility from the alternative that search strategies differed across participants. What I can conclude is that instructions that refer to a negated feature do not reliably produce confirmatory search.
17 Experiment 12

The results of Experiment 11 demonstrate that visual confirmation biases do not occur when search goals are communicated using negation (i.e., when looking for a target without a particular property). Despite the search stimuli being identical across negation and standard blocks, search strategy differed markedly. However, it possible that search is biased to stimuli that are confirmatory in an abstract sense when negation is removed. My previous demonstrations of confirmatory search have all relied on situations in which a tested proposition refers to the presence or absence of a single, visual feature, meaning that participants could create a single visual template, or expectation, in advance of a search for stimuli possessing that feature. In Experiment 12, I ask whether confirmatory search biases rely on this ability – to prepare a single visual template in advance – or whether a set of stimuli that are visually heterogenous might all attract attention solely because they could affirm a proposition. This provides a strong test of the possibility that participants select information because of its abstract ability to verify a proposition. The guidance of attention can be diluted when multiple potential target types are searched for (Menneer, Cave, & Donnelly, 2009; van Moorselaar, Theeuwes, & Olivers, 2014; but see Beck, Hollingworth, & Luck, 2012), suggesting that a confirmatory template for a visually heterogenous set of target types is unlikely unless stimuli are able to be rapidly perceived as confirmatory, and subsequently selected.

To test for attention biases towards visually heterogenous, but confirmatory, stimuli, Experiment 12 used instructions that referred not to individual colours, but instead the presence or absence of colour (i.e., saturation). Here, I expect that visual grouping processes involved with guidance (Duncan & Humphreys, 1989) will not contribute to salience, leaving only the categorical match between stimuli and the representation of search goals.

17.1 Methods

17.1.1 Participants

Seventeen undergraduates volunteered to participate in Experiment 12. All participants provided informed consent and were compensated with course credit. Participants were run until the included sample size of Experiment 10 (14) was matched after performance-based exclusions, using the same criteria as Experiment 10.
17.1.2 Stimuli and Procedure

The stimuli and procedure for Experiment 12 were identical to those of Experiment 10, with two exceptions. First, instead of using subsets of two different colours, one stimulus subset was now composed of random samples from the colours used in Experiment 10, whereas the other was composed of seven shades of gray (RGB values: 77, 77, 77; 102, 102, 102; 128, 128, 128; 153, 153, 153; 179, 179, 179; 204, 204, 204; 230, 230, 230). To ensure that all search stimuli were luminance increments relative to the background, I set the background screen colour to black (RGB: 0, 0, 0).

Second, the instructions were changed such that instead of participants answering a question about whether a target letter was on a specifically coloured circle, participants were instructed in one of two ways. The question posed to participants was either “Is the x on a colourful circle” or “Is the x on a gray circle.” Participants completed an equal number of both block types (six). Block order was again determined randomly for each participant. Figure 18c depicts a sample colourful-search instruction and search display (at Template-Matching Subset Size 4, with a Matching Target Colour).

17.2 Results and Discussion

Median correct RTs were again analysed, with the additional factor of Colour Category; that is, whether participants answered a questions about whether the target letter was on a gray circle or on a colourful circle. Three participants were excluded from analysis for either accuracy lower than 80% or average RT more than two standard deviations above the group mean (i.e., greater than 2830ms). Overall, only Target Colour, $F(1, 13) = 11.36$, $p = .005$, $\eta^2_p = .47$, affected correct search RT, such that trials that led to a “yes” response were overall faster, $M = 1793ms$, $SE = 100ms$, than trials where a “no” response was given, $M = 1981ms$, $SE = 97ms$ (see Figure 4. Critically, no effect of Template-Matching Subset Size was found, $F(2, 26) = 1.10$, $p = .38$, $\eta^2_p = .07$, indicating participants did not select stimuli on the basis of their colour category.
Furthermore, Template-Matching Subset Size did not statistically interact with Colour Category, $F(2, 26) = 0.27, p = .77, \eta^2_p = .02$, Target Colour, $F(2, 26) = 0.31, p = .73, \eta^2_p = .02$, nor their combination, $F(2, 26) = 2.26, p = .13, \eta^2_p = .15$. Finally, no factors or interactions affected search accuracy, $F s < 1.70, ps > .20, \eta^2_p < .12$.

To summarize, while a post-perceptual confirmation bias was present in this task, such that affirmation of the question being evaluated was faster than rejection, I did not find evidence that stimuli were prioritized for search on the basis of their template-matching features. This result indicates that dimension-level perceptual frames do not spontaneously guide search. In both Experiments 11 and 12, participants appear to have searched for target letters using a “brute force”, or random, search, making a decision about the target’s properties after having found it, rather than using target properties to guide attention to subsets of potential targets. At no point did the data suggest that guidance was used strategically (i.e., to search smaller subsets), despite this possible strategy. Feature-based subset searching, then, seems not to be a function of the environment, but rather of the participants’ task set. While this is clearly evident in the contrast between Experiment 10 and 11, where the same search stimuli were used, it is not clear whether grouping of subsets (by the presence or absence of hue) in Experiment 12 is even possible. Experiment 13 addressed this uncertainty.
18  Experiment 13

Although Experiment 12 did not reveal a confirmatory search tendency when stimuli are heterogeneous, this may reflect an inability to guide attention to stimuli sharing a more abstract feature, like hue, or its absence. To determine whether the lack of guidance in Experiment 12 was due to an inability to select a heterogenous group of stimuli or due to a lack of a bias, I conducted a final experiment where the target letter could be in the template-matching subset or not present at all. In this situation, selecting the template-matching subset is an ideal strategy. Thus, if heterogeneously coloured stimuli can be selectively searched when selection would improve performance, search times will increase in proportion to the size of the template-matching subset.

18.1 Methods

18.1.1  Participants

Fourteen participants, none of whom participated in any of the previous experiments, participated in Experiment 13. All of the participants were enrolled in a first-year undergraduate Psychology course at the University of Toronto, and were compensated with course credit for their participation. Participants all gave informed consent before participating.

18.1.2  Stimuli and Procedure

Stimuli and procedure were identical to those of Experiment 12, with two exceptions. First, target letters appeared on one of the template-matching search stimuli on half the trials, but on the other half of the trials, all letter stimuli were non-targets. Second, the instructions at the beginning of each block were changed to reflect this modification. The prompt for Experiment 13 was “For each trial, answer this question: Is the <target letter> on a <colourful/gray> circle? Press <key1> if yes, Press <key2> if no,” where angular brackets depict variable contents (i.e., the target letter could be p, d, b, or q).

18.2 Results and Discussion

One participant was excluded from analysis for having an average RT greater than two standard deviations from the group mean (i.e., greater than 3038 ms). Median search RTs can be seen in Figure 22. Template Matching Subset Size, \( F(2, 12) = 127.32, p < .001, \eta^2_p = .91 \), Target
Presence, $F(2, 12) = 172.16, p < .001$, $\eta^2_p = .94$, and Colour Category, $F(2, 24) = 14.37, p = .003$, $\eta^2_p = .55$, all affected search RTs, with an interaction between Target Presence and Template Matching Subset Size, $F(2, 12) = 35.02, p < .001$, $\eta^2_p = .74$. As can be seen in Figure 22, for both Colour Categories, search slopes were linear, with Target Absent searches being notably slower. A linear contrast for Template Matching Subset Size, $F(1, 12) = 151.52, p < .001$, $\eta^2_p = .93$, with no quadratic contrast, $F(1, 12) = 0.07, p = .80$, $\eta^2_p = .006$, showed that searches were restricted to appropriate category set. Searches were faster when the target was present, $M_{\text{present}} = 1335\text{ms}$, $SE_{\text{present}} = 63\text{ms}$, $M_{\text{absent}} = 1772\text{ms}$, $SE_{\text{absent}} = 88\text{ms}$. An analysis of accuracy also showed higher accuracy for Target Absent, $M = 94.8\%, SE = 1.2\%$ than Target Present, $M = 90.1\%, SE = 2.0\%$, searches, suggesting that miss errors were more common than false alarms, $F(1, 12) = 12.34, p = .004$, $\eta^2_p = .51$. Target Matching Subset Size, also affected accuracy, $F(2, 12) = 6.46, p = .006$, $\eta^2_p = .35$, such that accuracy declined as Subset Size increased, $M_{[2,4,6]} = [93.7\%, 92.9\%, 90.8\%]$, $SE_{[2,4,6]} = [1.6\%, 1.6\%, 1.6\%]$, suggesting that both misses and false alarms occurred more often when more stimuli matched the search template, a trend that was present in the confirmatory searches found in Experiments 10 and 11. Overall, however, these data show that searches can be guided towards a heterogeneous colour category (the presence or absence of hue), which, in combination with the findings of Experiment 12, show that the confirmation bias does not occur for visually heterogeneous templates.

![Figure 23. Median search times for Experiment 4. Error bars reflect one, within-subjects standard deviation of the mean.](image-url)
19 General Discussion

The goal of the Chapter 4 was to determine the level of representation at which biases in attention induced by the framing of a search goal occur. In Chapter 2, I showed that, in a search for two possible target conjunctions, simply phrasing the instructions such that one target is the absence of another target will lead to preferential selection of the latter target possibility (Rajsic et al., 2015). However, these results are attributable to a range of possible representational sources, ranging from simple visual priming to an abstract, logical target code. The present results demonstrate that confirmatory biases, as they exist in visual search, occur when one possible target type is defined by the presence of a visual feature (i.e., the colour “red”), but not when positive templates consist of a set of visual features (i.e., any coloured stimulus) or the absence of a visual feature (i.e., not red). This suggests that confirmation bias results from a sort of conceptual priming, such that propositions that can be translated into a single, categorical visual template can produce search biases for instances of this visual template. This is consistent with the finding that the presentation of verbal labels of objects speeds their entry in to awareness (Lupyan & Ward, 2013) and orients attention (Spivey, Tyler, Eberhard, & Tanenhaus, 2001), as well as findings that visually specific templates guide attention better than more abstract templates (Vickery, King, & Jiang, 2005; Maxfield & Zelinksy, 2012; Hout & Goldinger, 2014). Furthermore, it is consistent with findings that negative information tends not to guide attention in visual search (Moher & Egeth, 2012; Beck & Hollingworth, 2015; Becker, Hemsteger, & Peltier, 2016).

Given the contrast between the results of Experiment 13, which demonstrate an ability to attend to a heterogeneous subset, and the results of Experiment 12, which show no bias towards heterogeneous subsets due to the task framing, I must emphasize that the confirmation bias in visual search appears to be just that - a bias. As in Chapter 1, I interpret data from these experiments as indicating the presence of cognitive heuristics in search that can, in certain circumstances, be overcome. Indeed, Chapter 2 revealed that searches in which information is obtained more slowly shows a reduced confirmation bias (Rajsic, Wilson, & Pratt, 2017). Furthermore, Walenchok, Goldinger, & Hout (2016) have shown that confirmatory searching patterns are reversed when Template-Matching targets are less common than Template-Mismatching targets, suggesting search efficiency takes precedent over cognitive framing.
Overall, the available evidence suggests that cognitive economy is an important factor in the presence of cognitive heuristics in attention (see also: Irons & Leber, 2016).

Another important conclusion of this study is that merely framing one class of stimuli as positive instances of a hypothesis does not guarantee that they will be prioritized. What appears to be necessary for this bias to emerge is for positive instances to share a common visual feature, and for that feature to be explicitly stated in advance. As such, I speculate that the mechanism underlying this bias may be the visual representations that are constructed to encode and store the question being evaluated. This is consistent with the notion that attention is often involuntarily driven to stimuli with features that match information held in visual working memory (Soto, Hodsoll, Rotshtein, & Humphreys, 2008; Olivers, 2009). In Experiments 11 and 12, since targets were defined by the absence of a feature, or by a visually heterogenous set of features, I suspect that the search instructions could not be stored as a visual code. I note, however, that in Experiment 12, I did observe an overall RT cost for template-mismatching targets, suggesting an additional, post-perceptual confirmation bias.

The finding that confirmatory search exists only for non-negative templates is consistent with research on confirmation bias using Wason’s selection task (Wason, 1968; Evans & Lynch, 1973; Evans, 1998). Although participants often neglect to select the not-q card in their evaluation of an arbitrary rule (i.e., to use modus tollens), when participants evaluate the expression “if p then not-q”, their selection of the negated consequent (in this case, simply q) improves. Indeed, negation reduces card selections for both antecedent cases and consequent cases. These findings are consistent with the notion that evaluation performance in the standard task is a mixture of tendencies towards logical evaluation and tendencies, or heuristics, to select those cards with features that are mentioned by the rule (i.e., the p and q cards). Most theories of the matching bias explain it by appealing to some sort of relevance heuristic; at the first stage of reasoning, information must be sorted by its relevance to the evaluation of a proposition (Sperber, Cara, & Girotto, 1995). Stimuli that possess features contained in the to-be-evaluated proposition are rapidly seen as relevant, whereas stimuli that may be relevant, but are not mentioned in the proposition (i.e., a false consequent when evaluating an “if p then q” proposition) must be recognized as relevant by mentally unpacking the proposition’s implications. In this light, the visual confirmation bias does seem to be an instance of a matching bias heuristic, which is consistent with my previous work showing that it persists despite
instructions to attend the smaller subset (Rajsic, Wilson, & Pratt, 2015). Research on the matching bias has uncovered one salient limitation, however: the use of realistic materials and scenarios (Griggs & Cox, 1983; Oaksford & Stenning, 1992). In such situations, the richer knowledge base available to guide information selection and store the proposition in memory seems to reduce the effect of matching biases in data selection. As such, future research on the confirmation bias in search ought to consider using realistic materials and prompts to assess whether the matching-heuristic will still apply and lead to confirmatory search patterns, especially given the ability of object category knowledge to guide attention (Maxfield & Zelinsky, 2012; Yu, Maxfield, Zelinsky, 2016).
Chapter 5
Looking for Answers

Throughout this dissertation, I have suggested that top-down visual attention can be viewed as a process of testing whether a particular visual state (the presence or absence of a target) is true or false. In doing so, I consider visual search behaviour to be a particular sort of visual hypothesis test; determining whether a particular set contains one element. To be sure, different varieties of searches entail testing different hypotheses. Multiple-target searches increase the complexity of visual hypothesis testing by increasing the number of alternative hypotheses, and the present task adds to this the constraint that alternative hypotheses are mutually exclusive. The notion that visual processing entails hypothesis testing is not new (Enns & Di Lollo, 2000; Gregory, 1997; Lee & Mumford, 2003), however perceptual hypotheses in perception per se are often considered to be implicit, whereas the sorts of hypotheses considered in the present task may be more explicitly held (but perhaps only relatively so, as seen in Experiment 4). What I suggest here is that “top-down” visual attention can be considered the consequence of perceptual hypothesis testing – the result of cognitive systems trying to acquire specific, visually derived information. While this conceptualization does not overturn current conceptions of visual attention, it does perhaps reframe research on top-down attention as research on how we ask questions about the visual world. This specific goal – understanding how visual displays are compared to hypothetical states – has been studied in the context of the sentence verification paradigm (Clark & Chase, 1972; Underwood, Jebbett, & Roberts, 2004), which does tend to show a response-time advantage for sentence verification over falsification (but see Underwood & Green, 2003). However, perhaps surprisingly, these experiments have largely attempted to explain performance as a function of the propositional transformations required to compare linguistic and visual representations, and not as a consequence of the shifts of attention required to accumulate the necessary visual information. One recent exception is work by Franconeri in the context of graphical display comprehension (Franconeri, Scimeca, Roth, Helseth, & Kahn, 2012), which considers the verification of perceptual states as requiring a series of shifts of attention (or a “visual routine”; Ullman, 1984). This perspective on top-down attention certainly not the dominant one, and so the work in this dissertation may, I hope, lead to further attempts to interpret performance in complex visual situations as perceptual hypothesis tests.
While I have shown that visual search tends to prioritize confirmation of the presence of a target stimulus, even when seemingly suboptimal, it is not yet clear whether the confirmatory tendencies of top-down visual attention are related to confirmation bias as it exists in more cognitively complex reasoning situations. A surface relationship between visual search and reasoning has been suggested by Mercier (2012), who noted the utility of using visual search as an analogy for how individuals seek arguments in order to persuade another person. In principle, both vision and reasoning are subject to the problem of combinatorial explosion (Tsotsos, 1995; Evans, 2006) where the number of possible interpretations of observations exceeds any reasonable estimate of computability. In both cases, selection is necessary in order to arrive at any conclusion, and goal-driven selection is a sensible implementation for motivated agents.

Some very important differences, of course, exist between the kind of hypothesis testing that occurs in visual search and the kind of hypothesis testing that occurs in reasoning, especially about hypothetical problems (e.g., Wason, 1968, Tversky & Kahneman, 1973). In visual search, information that decisions are based on is continuously available, whereas information may be less available when reasoning about a hypothetical situation. Information may need to be represented in memory for decision making (for example, in estimating base-rates that are not explicitly given), or information may need to be constructed, imagined, or re-represented, as is the case in the selection task (Platt & Griggs, 1993). However, this simple depiction may give too much credit to vision. Visual information is very often ambiguous and incomplete, requiring some non-visual disambiguation (Gregory, 1997). Even when unambiguous, vision information processing is often capacity limited, both by a lack of acuity in peripheral vision (Zelinsky, 1998) and by the amount of information that can be maintained in visual working memory (even when visible; Tsubomi, Fukuda, Watanabe, & Vogel, 2013). Given that both vision and reasoning have to cope with the problem of incomplete or ambiguous data, it is possible that similar mechanisms, algorithms, and constraints are shared across these psychological domains.

While the problem of incomplete information, and thus the necessity of sampling policies, may be shared between these two contexts, differences in the type of information that is reasoned about may contribute a great deal to the strategies and heuristics used. The results of this dissertation show, at surface level, that colour-based perceptual hypotheses show confirmation bias. Chapter 4 showed that dimensional hypotheses (i.e., coloured or not) similarly did not produce a confirmation bias. As such, a general conclusion of this dissertation is that the
specific information format may contribute a great deal to the specific search strategy that is used. If information search policies do not simply involve optimizing the veracity of conclusions, but also involve optimizing the use of cognitive resources to construct search policies, then how information is represented should matter. Indeed, reasoning tends to improve when data are presented as natural frequencies compared to probabilities (Gigerenzer & Hoffrage, 1995) and informationally equivalent selection tasks lead to widely different error rates depending on whether rules concern “violations” (Cosmides & Tooby, 1992; Gigerenzer & Hug, 1992) or otherwise make negated consequents salient (Sperber, Cara, & Girotto, 1995). Likewise, visual search for a unique colour can be completed seemingly in parallel, whereas search for a unique feature conjunctions do not (Wolfe & Horowitz, 2004), which is almost certainly a consequence of how the visual system processes and codes information, and not a consequence of the search policy adopted. In predicting how information will be used to inform decisions, it is essential to know decision-makers mentally organize information.

A complicating factor in the question of representation is that both reasoning and visual attention are shaped profoundly by expertise and experience. Explicit choices between two gambles change depending on whether payoffs are experienced instead of described (Hertwig, Barron, Weber, & Erev, 2004), and learned knowledge can interfere with reasoning: logically valid syllogisms are often rejected when their conclusions conflict with semantic memory (e.g., all pets are cute - some pets are spiders - therefore, some spiders are cute; Evans, Barston, & Pollard, 1983; De Neys, 2006). Visual attention is also shaped by experience; it is captured by stimuli that have acquired associations with reward and task-relevance (Awh, Belopolsky, & Theeuwes, 2012; Anderson, Laurent, & Yantis, 2011) consistent with the notion that attentional biases that occur repeatedly become automatized (Shiffrin & Schneider, 1977; Carlisle, Arita, Pardo, & Woodman, 2011). In Chapter 2, I found that experience with a given hypothesis did not reduce confirmatory search, despite increasing the speed of search, showing that experience does not necessarily improve search efficiency, but speeds it. This was slightly surprising given recent findings that repeated searches reduce the requirement to hold search templates in visual working memory (Carlisle, Arita, Pardo, & Woodman, 2011), and may suggest that difficulties in testing multiple alternatives may not always be related to difficulties in capacity for multiple hypotheses (Mynatt, Doherty, & Dragan, 1993), but may instead emerge when competing
hypotheses require conflicting conditional reference frames (Koehler, 1991) or attentional settings (Huang & Pashler, 2007).

What did reduce confirmatory search, in Chapter 3, was practice searching in a situation where confirmatory search was artificially inefficient (because information was not rapidly available) did lead to the learning of a non-confirmatory search strategy. A second general conclusion of this dissertation, then, is that costs are important in shaping information search. Inefficient information search strategies can persist if they are not sufficiently inefficient to promote exploration of alternative search strategies; searchers may become trapped in a sort of local maximum of search efficiency unless encouraged to explore other strategy possibilities through payoffs. An interesting implication of this is that biases may be the most pronounced and idiosyncratic when environments sufficiently reward (or do not sufficiently punish) suboptimal search strategies (Friedrich, 1993). For example, in the choice of restaurants to eat at, positive experiences at one restaurant may encourage revisiting that restaurant over visiting other, novel restaurants. This in turn biases the sampling of restaurant quality, such that better restaurants are less likely to be tried, and thus a disproportionate weighting of the quality of the good restaurants that one has experienced. Ultimately, this provides a problem only when a relatively large number of restaurants are satisfactory or better – if all restaurants were poor, dissatisfaction would promote wider exploration – but it does suggest that environments wherein positive resources are abundant will lead to idiosyncratic, and essentially baseless (not objectively superior to alternatives), preferences that are sample-dependent (Hoeffler, Ariely, & West, 2006).

With respect to the confirmation bias, my 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). After all, human decision makers have only limited time and cognitive resources at their disposal in the
course of information search and evaluation (Griffiths, Lieder, & Goodman, 2015). My 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). Goal-driven selection is a broad feature of the human mind, and while globally beneficial, it can lead us to flawed beliefs or behaviors when the assumptions borne by particular goal-driven attentional settings are themselves flawed, irrespective of the domain of analysis. Of course, all is not lost; flawed beliefs and interpretations can be corrected by a sufficient amount of inconsistent evidence. The effect of selection is merely a bias towards certain conclusions, not utter hegemony of beliefs and expectations in the face of all available evidence. My primary proposal is that the effects of selection on evaluation of information will occur regardless of the domain in question. It is for these reasons that I suggest (along with others: Manohar & Husain, 2013; Meier & Blair, 2013; Hills et al., 2014; Butaccio, Lange, Thomas, & Dougherty, 2015) that visual search can be informative about general information search strategies.

A deep relationship between search and reasoning has been suggested by Hills and colleagues (2006; 2014), who argue that goal directed cognition may find its evolutionary roots in foraging behavior, and has shown that search styles can be primed across domains—for example, between a visual foraging task and a lexical search task (Hills, Todd, & Goldstone, 2008)—suggesting shared cognitive control mechanisms. Seeing reasoning as a relative of search processes that evolved for foraging may provide a novel perspective on reasoning errors. For example, Oaksford and Chater showed that selection patterns in the Wason selection task, which are logically incorrect, are consistent with a rarity assumption in the context of probabilistic reasoning—as long as “p” and “q” refer to sufficiently less frequent occurrences than “not p” and “not q”, inspecting p and q cards provides more information than the normatively valid “not q” card”. This is precisely the sort of type of context in which goal-directed, conditional action would provide the most gains in foraging—when consequents (resources) are rare, but its probability conditional on an antecedent (a cue) is high, allowing for the resource to be found or collected better than under random exploration. Goal-directed states, in the context of foraging,
are only beneficial when they improve one’s ability to collect a needed resource; falsifying a spuriously valid hypothesis about resources would rarely be expected to improve one’s resource collection (in fact, it might reduce it). In other words, sufficiency is more important than necessity when it comes to hypotheses about resources. If this asymmetry between confirmation and falsification is valid, then foraging-related evolutionary shaping of primitive cognitive processes may have provided the foundation of reasoning biases; in the present case, the relative ease of applying template-matching processes in goal-directed perception and action. Indeed, perceptual matching does seem to be a shared cognitive ability across numerous species (see Lind, Enquist, & Ghirlanda, 2015), and has been suggested to underlie the understanding of abstract relationships “same” and “different” (Hochmann, Mody, & Carey, 2016), whereas more complex ways of gaining information (e.g., inference by exclusion; Call, 2004) are less readily observed across species (Aust, Range, Steurer, & Huber, 2008). Moving forward, demonstrating a direct connection between attentional biases in visual reasoning and attentional biases in offline reasoning would help provide support for the speculation that shared mechanisms underlie the two sorts of cognitive operations.
References


http://doi.org/10.1037/0096-1523.16.3.459


Copyright Acknowledgements

Portions of this dissertation have been published in the following articles, available from their respective publishers:

