THE UNIVERSITY OF CHICAGO

NAVIGATION STRATEGIES AND HEURISTICS IN CONSUMER SEARCH

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE UNIVERSITY OF CHICAGO

BOOTH SCHOOL OF BUSINESS

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

ALEXANDER KENTARO MOORE

CHICAGO, ILLINOIS

AUGUST 2023

DEDICATION

For my family, who support me through the best and worst of times.

Contents

Acknowledgments v
List of Tables vii
List of Figures ix
1. Chapter 1. Navigation Strategies in Search 1
1.1. Introduction 1
1.2. Optimal Search Strategies 4
1.3. Value-Based Search Heuristics 10
1.4. Spatial Search Strategies 17
1.5. The Present Research 27
2. Chapter 2: Value-Based Navigation 29
2.1. Study 1: Estimating Reservation Values 29
2.2. Study 2: Search with Navigation with Numerical Stimuli 39
2.3. Study 3: Search with Navigation with Graphical Stimuli 49
2.4. General Discussion 59
3. Chapter 3. The Value-Spatial Navigation Tradeoff 60
3.1. Study 1: The Impact of the Number of Boxes 60
3.2. Study 2: The Impact of Display Features 73
3.3. General Discussion 93
4. Chapter 4. Learning 94
4.1. Study 1. Learning with 5 Boxes 94
4.2. Study 2. Learning with 12 Locations 109
4.3. General Discussion 123
5. Chapter 5: General Discussion 125
5.2. Limitations and Scope of Research 130
5.3. Conclusion 132
6. Appendix 135
6.1. Supplementary Materials for Chapter 2, Study 1 135
6.2. Supplementary Materials for Chapter 2, Study 2 and Study 3 137
6.3. Supplementary Materials for Chapter 3, Study 1 138
6.4. Supplementary Materials for Chapter 3, Study 2 144
6.5. Supplementary Materials for Chapter 4, Study 1 154
6.6. Supplementary Materials for Chapter 4, Study 2 162
7. References 166

Acknowledgments

Throughout my time at Booth and the Center for Decision Research I have learned a great deal from the faculty. I would like to offer my deepest gratitude to my advisors, Reid Hastie and Dan Bartels, who taught me more than they can know through our hours of conversation. I would also like to thank Emma Levine, George Wu, and Alex Koch who all taught me a great deal about research and academia.

My dissertation would not have been possible without the help of many people. I would like to thank my dissertation committee: Reid Hastie, Dan Bartels, Oleg Urminsky, Nicholas Reinholtz, Pradeep Chintagunta, and Alex Imas. I would also like to offer my gratitude to the leadership and staff at the CDR. I would especially like to thank Kaushal Addanki, Bryan Baird, and Rui Sun. Finally, I would like to thank the CDR, the Booth PhD Office, and the National Science Foundation for financial support.

Throughout my time as a student, I benefitted from the kindness, wisdom, and intelligence of my peers. I would especially like to thank Minkwang Jang, who has supported me through difficult times and taught me a great deal. I would also like to thank Kariyushi Rao, Kristina Wald, Sam Hirshman, Diag Davenport, Nick Herzog, Lin Fei, Yena Kim, and Danny Katz for many great conversations. And, I would like to thank the members of HOCAG for helping me develop many of my ideas.

Finally, I want to thank all the friends and family who have supported me. Thanks to my wife, Victoria, who has always been there for me in the toughest of times. Thanks to my parents, Dan and Mieko Moore, who have taught me everything I know about being a decent person. Thanks to all my grandparents, Wilma and Richard Moore and Sami and Yoshiro Kato. I particularly want to thank Yoshiro Kato for being a role model on how to live and Richard

Moore for being my intellectual role model. Thanks to my son, Isaac, who makes sure that I don’t work too much. And thanks to my friends, Dan Evans, Carrie Smith, Matthew Robinson, and Dimitar Antov. Finally, I would like to thank Patricia Nordeen, Boaz Keysar, and Linda Ginzel for setting me on this path.

List of Tables

Table 2.1.1. Regression of stated RVs on the minimum, maximum, and search costs in Study 1 32
Table 2.1.2. Regression of Stated RVs onto Ranges, EVs (expected values), and Search Costs in Study 1 34
Table 2.1.3. Heuristic fit of Stated RVs in Study 1 37
Table 2.2.1. Regression of Winnings and Outcomes as a Percent of Optimal Search Outcomes on Strategy in Study 2........ 46
Table 2.3.1. Regression of Winnings and Outcomes as a Percent of Optimal Search Outcomes on Strategy in Study 3........ 56
Table 3.1.1. Regression of Down (vs Up) Moves onto Number of Boxes Presented 70
Table 3.2.1. Number of Participants in Each Treatment in Study 2 75
Table 3.2.2. Average Duration (A), Maximum Value (B), and Profit (C) in Study 2 77
Table 3.2.3. Percent of Participants Making RV or EV Choices First in Study 2 80
Table 3.2.4. Average RV (Panel A) and EV (Panel B) of locations inspected in Study 2 .. 85
Table 4.1.1. Stimulus Sets for Study 1........ 95
Table 4.1.2. Global Descriptive Statistics for Searches in Study 1 96
Table 4.2.1. Stimulus Sets for Study 2........ 110
Table 4.2.2. Global Descriptive Statistics for Searches in Study 2 112
Table 6.1.1 Study 1 stimuli and average stated RVs (reservation values)........ 135
Table 6.2.1. Study 2 and 3 Stimuli........ 137
Table 6.3.1. Stimuli in Studies 1 and 2 138
Table 6.3.2. Individual Level Comparison of First Inspection by Position to Random Choices........ 140
Table 6.3.3. Models of Up and Down Moves on Number of Boxes 142
Table 6.3.4.Models with Number Remaining Above and Below the Most Recently Inspected Brand 143
Table 6.3.5. Coefficient on Binary Variable of Double Downward Inspections (Present = 1) for Inspection Order 1 through 4 143
Table 6.4.1. Regression of (1) Duration, (2) Maximum Value, and (3) Profit on Conditions 144
Table 6.4.2. Regression of Maximum RV first choices (Models 1 and 3) and EV first choices (Models 2 and 4) on Conditions........ 145
Table 6.4.3. Comparison of First Inspection by Spatial Position to Random Choices...... 146
Table 6.4.4.Individual Level Comparison of First Inspection by Position to Random Choices........ 147
Table 6.4.5. Regression of First Choices in Top 3 Positions on Conditions 148
Table 6.4.6. Regression of Average RV of Searches (Model 1) and Average EV of Searches (Model 2) on Conditions........ 149
Table 6.4.7. Regression of Up vs Down Moves onto Graphical/Numerical Conditions in Linear Conditions........ 150
Table 6.4.8. Regression of Up vs Down Moves onto Graphical/Numerical Conditions in Linear Conditions by Screen Position........ 151
Table 6.4.9. Regression of Clockwise vs Counterclockwise Moves onto Graphical/Numerical Conditions in Circular Conditions 152

Table 6.4.10. Regression of Clockwise vs Counterclockwise Moves onto Graphical/Numerical Conditions in Circular Conditions 153
Table 6.5.1. Regression of Profit onto Search Number in Study 1.................................... 154
Table 6.5.2. Regression of Ratio of Optimal to Participant Profit onto Search Number in Study 1 155
Table 6.5.3. Regression of Maximum Value Found onto Search Number in Study 1 155
Table 6.5.4. Regression of Search Duration onto Search Number in Study 1.................. 156
Table 6.5.5. Regression of RV and EV First Choices onto Search Number in Study 1 156
Table 6.5.6. Linear Probability Model (Table A) and Logistic Regression Models (Table B) of RV and EV First Choices Onto Block By Condition 157
Table 6.5.7. Regression of Mean RV and EV onto Search Number in Study 1 159
Table 6.5.8. Regression of Mean RV and EV onto Search Number by Tertile in Study 1 161
Table 6.6.1. Regression of Profit onto Search Number in Study 2................................... 162
Table 6.6.2. Regression of Ratio of Optimal to Participant Profit onto Search Number... 162
Table 6.6.3. Regression of Maximum Found onto Search Number in Study 1................. 163
Table 6.6.4. Regression of Search Duration onto Search Number in Study 1.................. 163
Table 6.6.5.Regression of (A) RV and (B) EV First Box Choice and onto Search Number in Study 2 164
Table 6.6.6. Regression of Average RV and EV onto Search Number in Study 2............ 165

List of Figures

Figure 1.4.1. Comparison of Search With No Working Memory Constraints (left) with a Simple Pairwise Navigation Model that Extracts Reservation Values (RVs) Randomly (right) ... 24
Figure 1.4.2. Optimal Pairwise Navigation with Scanning (Left) and Satisficing Pairwise Navigation with Scanning (Right) ... 26
Figure 2.1.1. Average stated RVs (reservation values) by Box Minimum, Box Maximum, and Search Costs in Study 1 (Horizontal black lines are box ranges; light gray lines are risk-neutral normative model predictions) in Study 1. ... 33
Figure 2.1.2. Scatter plot and histograms of participants by individual regression weights on the Minima and Maxima of the values defining the boxes for which they estimated RVs in Study 1. ... 37
Figure 2.2.1. Interface display shown to participants in Study 2 ... 40
Figure 2.2.2. Navigation Heuristics Consistent with Participant First Choices (Panel A) and by Total Length of Search for Full Paths (Panel B) in Study 2. ... 43
Figure 2.2.3. Proportion of Contingent Navigation Decisions Consistent with RV and EV Strategies by Inspection Number in Study 2 ... 44
Figure 2.2.4. Stop/Continue Decisions by Inspection and Strategy Classified based on First Choices (Panel A) and Full Paths in Study 2 (Panel B) ... 46
Figure 2.2.5. Participant Percent of Optimal Model Profits, Costs, and Maximum Values for Strategies Classified by First Choices (Panel A) and Full Paths (Panel B) in Study 2 ... 47
Figure 2.2.6. Percent of Searches with the Same Outcome As The Optimal Model For Strategies Classified by First Choices (Panel A) and Full Paths (Panel B) in Study 2.. 48
Figure 2.3.1. Stimulus Display in Study 3 ... 51
Figure 2.3.2. Navigation Heuristics Consistent with Participant First Choices (Panel A) and by Total Length of Search for Full Paths (Panel B) in Study 3 ... 53
Figure 2.3.3. Proportion of Contingent Navigation Decisions Consistent with RV and EV Strategies by Inspection Number in Study 3 ... 54
Figure 2.3.4. Continuation Decisions by Round and Strategy based on First Choices (Panel A) and Full Paths (Panel B) in Study 3 ... 54
Figure 2.3.5. Participant Percent of Optimal Model Profits, Costs, and Maximum Values for Strategies Classified by First Choices (Panel A) and Full Paths (Panel B) in Study 3 ... 57
Figure 2.3.6. Percent of Searches With Same Outcome As The Optimal Model For Strategies Classified by First Choices (Panel A) and Full Paths (Panel B) in Study 3.. 58
Figure 3.1.1. Percentage of Sequences of 1, 2, and 3 Strategy Consistent Inspections by Number of Brands Presented in Study 1 ... 65
Figure 3.1.2. First Inspections at Each Vertical Position for Each Condition in Study 1.... 67
Figure 3.1.3. Log of Ratio of Up vs Down Choices by Condition and Vertical Position in Study 1 ... 69
Figure 3.1.4. Percentage of Two Inspection Sequences that are Both Downward from the Most Recently Inspected Brand by Current Inspection Number in Study 1 ... 72
Figure 3.2.1. Examples of the (A) Numerical-Circular and (B) Graphical-Linear Layout from Study 2 ... 77

Figure 3.2.2. Spatial Position of First Choices in Study 2 82
Figure 3.2.3. Proportion of Participant First Choices by Screen Position and Condition in Study 2 83
Figure 3.2.4. Average RV (Panel A) and EV (Panel B) of locations inspected in Study 2 . 86
Figure 3.2.5. Difference between Percent of Participants' Downward Moves and Percent of Random Model Downward Moves 89
Figure 3.2.6. Counterclockwise vs Clockwise Sequential Choices in Circular Conditions by Screen Position in Study 2 91
Figure 4.1.1. Mean Profits by Search Number 97
Figure 4.1.2.Percent of Optimal Model Profits by Search Number and Condition............. 99
Figure 4.1.3. Mean Search Duration by Search Number 102
Figure 4.1.4. Relationships Between Individual Level Coefficients for (A) Maximum and Profit, (B) Duration and Profit, and (C) Duration and Maximum 103
Figure 4.1.5. Choice of First Box by Search Number 105
Figure 4.1.6. Mean RV (Panel A) and Mean EV (Panel B) by Search Number................ 106
Figure 4.1.7. Mean RV (Panel A) and Mean EV (Panel B) by Search Number and Learning Tertile 108
Figure 4.2.1. Mean Profits by Search Number in Study 2 113
Figure 4.2.2. Percent of Optimal Model Profits by Search Number and Condition in Study 2 114
Figure 4.2.3.Mean Maximum Values Found by Search Number in Study 2..................... 115
Figure 4.2.4. Mean Search Duration by Search Number in Study 2 116
Figure 4.2.5. Choice of First Location by Search Number and Condition in Study 2....... 118
Figure 4.2.6. Mean RV (Panel A) and Mean EV (Panel B) by Search Number in Study 2 119
Figure 6.1.1. Mean Stated RVs By Cluster in Study 1 136
Figure 6.3.1. Log of Ratio of Up vs Down Choices by Condition and Vertical Position after the First Inspection in Study 1 141
Figure 6.5.1. Mean Maximum Values Found by Search Number in Study 1.................... 154
Figure 6.5.2. Mean RV (Panel A) and Mean EV (Panel B) by Search Number and Learning Tertile by Condition in Study 1 160

1. Chapter 1. Navigation Strategies in Search

1.1. Introduction

The ability to search is an essential cognitive skill that underlies many human activities, from survival in primordial environments to finding good deals in modern economic markets. As the availability of information and tangible resources expand with the growth of electronic environments, search becomes ever more important for human achievements. From the specific perspective of marketers, this abundance of options makes it increasingly important to understand how consumers search for products they want. The present research uses laboratory-based experiments to explore the strategies and information that consumers use to navigate as they search for products.

Many scientific disciplines are concerned with search, contributing both normative models for effective or optimal search and descriptive models of how intelligent agents (many non-human) search for resources and information. Relying heavily on research by behavioral ecologists and computer scientists, in addition to psychologists and economists, we have identified three abstract principles for goal-directed search: value-based search, proximity-based search, and socially-based search. The present research focuses on value-driven, and proximity-driven search strategies. Value-driven strategies rely on inferences about distributions of value to guide and terminate search. The most relevant models for our purposes are from Economics (discussed extensively below). Proximity-driven strategies rely on the spatial or conceptual layout of the environment to make decisions about where to search and when to stop, the most relevant models are from Cognitive Psychology, Behavioral Ecology, and Computer Science (e.g., Danchin, Giraldeau, & Cezilly, 2008; Pirolli, 2010; Smith & De Lillo, 2022). Socially-driven strategies rely on information about the actions and outcomes for similar co-specifics and the most useful models come from Behavioral Ecology, Economics, and Social Psychology. (In the present research, we do not address socially-driven strategies, only because of limited time and resources.)

Based on theoretical models of search from economics (e.g., Weitzman, 1979), we conceptualize search as involving three components: navigating to options (aka selection), evaluating options, and stopping by choosing options. Consumers sequentially inspect available products for a search cost, ultimately choosing one of the products that they have sampled so far. In economics and marketing, there is a rich history of rational, optimal models that describe such sequential search tasks (B. P. McCall & McCall, 2008). Earlier theoretical and empirical work focused on models that describe optimal stopping rules (J. J. McCall, 1970; Mortensen, 1970; Stigler, 1962). A smaller body of empirical research builds on theoretical developments to consider optimal navigation in addition to optimal stopping (e.g., Gabaix et al., 2006; Weitzman, 1979).

While these models provide important insights about optimal ways of searching, they require strong assumptions. To search in accordance with these models, consumers must have full information as well as unrealistically powerful computational abilities. Rather than fully optimizing a search, consumers rely on simplifying heuristics to help them navigate. Research in both marketing and psychology has found that consumers use a variety of sub-optimal heuristics that take advantage of the statistical structure of the environment to simplify complex decisions (Gigerenzer & Goldstein, 1996; Payne et al., 1993). And several studies of sequential search have proposed heuristics to explain non-optimal stopping decisions (Hey, 1982; Schunk & Winter, 2009; Sonnemans, 1998). Very little research has examined how consumers with limited cognitive capacities *navigate* in sequential search situations where they have differing expectations about products (c.f., Gabaix et al., 2006; Urbany, 1986).

Economic accounts of search usually omit environment factors that do not directly impact the information available about products. The product attribute structures and spatial layout (outside those explicitly impacting search costs) of products are typically not modeled as factors impacting search. For example, in standard search models, adding more products to a display may change the order in which consumers sample products because more attractive prospects are added. But adding these products should not change the *manner* in which people search. The addition of new products should not lead consumers to change how they acquire and process information about each product. Similarly, changing the spatial layout in which products are displayed should not impact the information that consumers use to navigate. Whether products are arrayed across a single shelf, across multiple shelves, or in a geometrically more complex display should not impact the products that consumers choose to sample.

To some extent, the impacts of environment factors have been integrated into classic models of search by varying search costs associated with locations (e.g., Ursu, 2018). While including these factors can lead to improved prediction, they do not offer a psychological account of *why* consumers might search in different ways. Further, cognitive accounts of how people simplify complex tasks (e.g., Johnson & Payne, 1985; Payne et al., 1993) and use their physical environments to reduce the cost of thinking (e.g., Clark & Chalmers, 1998; Dunn & Risko, 2016; Hutchins, 1995; Kirsh, 2010) offer a principled account of how these environment factors impact search strategies in predictable ways.

The environment factors described above are one form of complexity. Other forms of complexity are also likely to shift search behaviors away from optimal strategies to more cognitively efficient heuristic habits. For example, different distributions of expectations can lead to more complexity; skewed expectations may be more difficult to think about than symmetrical distributions; and different forms of feedback from the environment might change the ease with which consumers can learn to search effectively (see Hogarth et al., 2015).

The presence of multiple attributes that determine value also makes search more complex. Consumers usually integrate multiple product attributes to infer the value of a product. In a standard sequential search model, expectations based on multiple attributes are assumed to be aggregated into an overall utility. However, given limited attention and cognitive resources, in practice adding an additional step of information gathering or computation will impact search strategies. Prior decision making research has identified a variety of effort-reducing heuristics that consumers use to make decisions involving products with multiple attributes (e.g., Payne et al., 1993; Valenzuela et al., 2009). For example, consumers often engage in non-compensatory attribute-by-attribute searches instead of calculating the overall value of each possible option (Bettman et al., 1998). It is likely that consumers use similar heuristic strategies in sequential search when information about locations is presented across multiple attributes.

1.2. Optimal Search Strategies

As a starting point for understanding consumer search, we begin with optimal search models from economics. We expect that search behavior is likely to approximate the optimal strategy in simple, familiar situations. But, consumers are unlikely to use these models to calculate optimal navigation and stopping decisions especially when the search task is complex and unfamiliar. Nonetheless, the optimal models highlight important features of any search environment that consumers might focus on if they want to find good products. In addition, the models provide a performance benchmark against which we can compare consumers' decisions.

Further, they suggest a family of heuristic "stopping threshold models" which are more computationally plausible than the fully rational Basic Model.

1.2.1. The Basic Search Model

In the Basic Search Model (J. J. McCall, 1970; Mortensen, 1970; Stigler, 1962), a consumer is described as inspecting products one at a time in a random order from an infinite set of options. The consumer has full knowledge of the distribution from which the values of all the products are drawn. After inspecting each product, the consumer learns the exact value of the product. She can pay a search cost to inspect another product or stop searching. When she stops searching, she can either acquire any product that was inspected during the search for no additional search costs (i.e., recall) or go home empty-handed. The consumer's objective is to maximize her expected net value from conducting the search.

Faced with this situation, a rational consumer will continue searching until the marginal cost of an additional inspection exceeds the expected benefit of further search. Thus, a risk neutral consumer will stop searching as soon as her next inspection meets the following condition:

$$max(S_n)\int_{-\infty}^{max(S_n)} f(x)dx + \int_{max(S_n)}^{\infty} xf(x)dx - max(S_n) - C < 0 \qquad (1)$$

where S_n is the set of n products sampled from the distribution so far, f(x) is the PDF of the known distribution of values, and C is the search cost. Intuitively, the integral on the left states that if the n+1st sample has a value less than that of the highest value product found so far, a searcher would choose the higher value product that she found earlier were she to stop searching. The integral on the right is the expected value given that another product is found that has a higher value than the best product found so far. When these two integrals are summed together, we have the expected benefit from sampling the n+1st product. When we subtract the value of the best product found so far, we have the expected increase in value between sampling the nth and n+1st product. Finally, when we subtract the search cost, we have the expected gain from sampling the n+1st product. When the expected gain is positive, the expected value of sampling again, given the values of the products sampled so far, is positive, and vice versa when the expected gain is negative.

One notable feature of the Basic Search Model is that for a given distribution of outcomes and search costs, consumers can use a single threshold value to determine whether it is optimal to stop searching or search again. This threshold is called the Reservation Value (RV). The RV is the value of a hypothetical product a consumer could find during a search which would make the expected gain from additional searches negative. When the expected gain of continuing is negative, a rational consumer will stop searching. The RV is calculated by solving for R in the following equation:

$$R\int_{-\infty}^{R} f(x)dx + \int_{R}^{\infty} xf(x)dx - R - C = 0 \tag{2}$$

Although the RV is difficult to calculate, once it is calculated it offers a simple rule for optimal stopping. When a consumer samples a product with a value less than the RV, she ought to sample again. When a consumer finds a product with a value greater than the RV, she ought to stop searching and select that product.

Note that this model and its optimal rule for stopping imply that products ought to be evaluated based on the expected gains from inspecting them. However, since the Basic Model assumes that all products are drawn from a single market-level distribution, there is no information to guide navigation by discriminating among sources of products. In short, the Basic Model ignores navigation.

Nonetheless, the Basic Search Model has implications for our study of navigation. First, holding all else constant, when search costs are high, the average length of searches and selectivity will decrease because the RV decreases. Second, holding all else constant, when variance of the distribution of outcomes increases, consumers will increase average search duration and be more selective because the RV increases. Third, the upper part of a distribution of outcomes will be more important in determining selectivity and average search duration than the lower part. This implication stems from the fact that losses from drawing a product with a value in the low end of a distribution are limited because a consumer can always return to the best product encountered so far (i.e., recall) instead of accepting a low value product.

Theoretical research has identified conditions that require extensions of the Basic Search Model. For example, some sequential search models allow consumers to learn a distribution as search proceeds, rather than having perfect information ex ante (Rothschild, 1974). Other models apply to search tasks in which the horizon is realistically limited to a finite set of products (Benhabib & Bull, 1983). In these cases, the Basic Model's optimal stopping rule needs to be modified. When a consumer does not have full information ex ante, but learns from sampling, her RV will change based on the products she encounters. And, when consumers are searching a finite set of products, their RV declines as they approach the search horizon.

The Basic Search Model provides many insights into the behaviors of workers looking for jobs (for a review, see Mortensen, 1987) and consumers looking for products (e.g., De los Santos et al., 2012) in real markets. Although we doubt that typical consumer thought processes exactly compute the reservation value, we believe that consumers have a grasp of the conceptual principle of making stopping decisions based on expected gain. We often find ourselves asking, "Is it 'worth' visiting one more appliance store? What is the chance I will find a better deal if I drive 30 minutes to Walmart?"

Studies have been conducted to compare behavior to model predictions in artificial laboratory search tasks. Several studies have concluded that participants search optimally across variations in search costs (Brannon & Gorman, 2002; Braunstein & Schotter, 1982; Cox & Oaxaca, 1989) and value distributions (Dellaert & Häubl, 2012; Schotter & Braunstein, 1981). However, studies have also found that participants stop earlier than predicted by the optimal model (Brown et al., 2011; Cox & Oaxaca, 1989, 1992; Einav, 2005; Häubl et al., 2010, among many others). Some theorists have suggested that risk aversion explains this early stopping; but the evidence on whether risk aversion can account for early stopping is weak (See Cox and Oaxaca, 1989, 1992 in support; Sonnemans, 1998 against). And still other studies have found that participants do not respond to variance as predicted by the Basic Model (Brannon & Gorman, 2002; Reinholtz, 2015). Our summary is that experimental participants sometimes exhibit search behavior that is roughly consistent with the Basic Model, but there is considerable heterogeneity in search strategies and a comprehensive account needs to include both essentially rational strategies, and non-optimal heuristic strategies.

1.2.2. The Weitzman Model

The Weitzman Model extends the Basic Search Model by proposing that navigation is guided by RVs. The model starts with the assumptions of the Basic Model, but stipulates that consumers are no longer drawing products repeatedly from a single distribution of values. Instead, the consumer can inspect single products from different locations (analogous to stores or websites), with individual value distributions associated with each location. This means the rational searcher has a different RV for each location. The following rules define the behavior of a consumer adhering to the Weitzman model:

(A) Selection Rule: Navigate to locations in descending order of RV;

(B) Stopping Rule: Stop searching once the value of the best product found so far exceeds the RV of all locations that have not been visited;

(C) Choice Rule: Pick the product with the highest value.

This final rule implies that consumers can return to any location they have already visited without incurring additional costs (recall). This set of rules implies the searcher will navigate guided by location RVs. And, holding all else constant, a consumer ought to visit a location with more variation prior to one with less.

Before continuing, we want to clarify our terminology. We think of navigation as any rule or principle that consumers use to decide *where* to search. In keeping with Weitzman's terminology, we use the term "Selection" to describe his specific rational navigation rule. In general, throughout this article, however, we use the word "Navigation" as a broader term to describe all rules (both optimal strategies and sub-optimal heuristics) that guide searchers to specific locations.

As a concrete example, a consumer might consider visiting two websites when searching for the lowest price on a microwave oven. For each website, she has different expectations in the form of price distributions. At one website, she expects that the price will be moderate and will vary little. The prices at the other website are higher, but occasionally there is an attractive discount. Once a consumer visits a website, her price expectations are replaced by the actual price of the oven she found at that website, and that source's price is fixed for the remainder of the search. The shopper visits the website with the highest RV first. If the value obtained at that website is higher than the RV for the second website, she stops searching and purchases from the first website. Otherwise, she moves on to the second website. If there were a third website with a sufficiently high RV, she would inspect that one after looking at the second website. Ultimately the consumer purchases at the website that offered her the best deal on the oven found up to that point. To do so, she can return to any of the websites she has already visited (i.e., recall) without accruing additional search costs.

Several studies have applied the Weitzman Model in the analysis of consumer marketing data. Kim et al. (2010) used the model to identify heterogeneous consumer tastes and search costs using data from online camcorder sales. Honka and Chintagunta (2017) used it to model shopping for auto insurance. Ursu (2018) used data from a field experiment conducted by a major travel website and showed that rankings in search results impact search costs, but not expectations. And Moorthy et al. (1997) used the model to test the impact of product category expertise using survey data. But, very few studies have tested the basic assumptions of the Weitzman Model or have examined its implications in controlled laboratory experiments (cf., Gabaix et al., 2006).

1.3. Value-Based Search Heuristics

While the optimal models provide important insights into factors that consumers ought to attend to as they inspect options, they also require a lot of computation. It is unlikely that consumers, even in high stakes situations, will solve the equations that define the reservation value. Instead, consumers are likely to simplify decisions by approximating the value of each location based on a set of available cues. We call these simplifying strategies “Value-Based” search heuristics because they use approximate values to guide navigation and stopping. As we will see later in the introduction, consumers may also rely on cues unrelated to value to guide navigation as well. For example, in a realistic physical or electronic "store," consumers will often rely on proximity to guide search inspections. And, consumers often rely on social information from experts or peers to find desirable products. More generally, it is likely that as computational complexity increases, consumers will increasingly rely on non-optimal value-based heuristics, or even on non-value-based cues, as they search.

Past research into non-optimal stopping rules suggests two classes of value-based heuristics: fixed threshold heuristics and variable threshold heuristics. Consumers using both classes of heuristics set a value threshold and stop when an item they find exceeds this threshold. Consumers following the optimal strategy use the reservation value as a threshold and stop when they find an item that exceeds the reservation value. In a similar way, consumers using a fixed threshold heuristic determine a threshold without precisely calculating the reservation value. For example, a consumer might use the expected value of the distribution instead of the reservation value as a fixed value threshold.

Consumers using variable threshold heuristics also rely on a threshold, but these thresholds can change as a search unfolds. For example, a consumer might reduce a threshold and become less selective as a search continues. While variable thresholds are not consistent with the Basic Search Model, more complex optimal search models that incorporate learning (Rothschild, 1974) or finite search horizons (Benhabib & Bull, 1983; Cox & Oaxaca, 1989) propose shifting thresholds.

1.3.1. Fixed Threshold Value Heuristics

Broadly speaking, two families of fixed-threshold heuristics have been proposed. One class of heuristics compares the *value of each inspected item* to a fixed threshold. The other class compares the *overall earnings* from a search to a threshold.

Several studies have found that participants in behavioral experiments use fixed-value threshold heuristics. For example, Hey found that 41% (1982) and 74% (1987) of participants behaved consistently with a fixed-value threshold; Schunk and Winter (2009) found that 77% of participants behaved consistently with a fixed-value strategy. And when participants were asked to write down their strategies, Sonnemans (1998) found that 22% of them described a fixed-value heuristic.

While many participants behave consistently with fixed-value threshold heuristics, there is substantial variation in how these thresholds are calculated. Moon and Martin (1990) found that roughly half of participants behaved as though they stop once they find an item with a value that is at least 0.75 standard deviations above the mean. In addition, some fixed-value heuristics rely on statistical parameters (e.g., mean or median) of partial information acquired prior to search (c.f., Martin & Moon, 1992).

Fixed-earnings threshold heuristics describe thresholds in terms of the level of earnings to which a searcher aspires (Butler & Loomes, 1997; Kogut, 1990; Sonnemans, 1998). These heuristics are called “satisficing” heuristics (Simon, 1955, 1956). Generally, these thresholds are not based on the distribution of option values. Instead, the heuristics balance the value of each option against cumulated costs. For example, a consumer using a simple fixed-earnings threshold heuristic might aspire to earn more than $1 at the end of search. She would subtract the current cumulative search cost from the value of each new option and stop if that quantity exceeds $1. Schunk and Winter (2009) found that 37% of their participants seemed to follow this type of fixed-earning heuristic. They do not, however, comment on how their participants may have arrived at their thresholds.

Note that the use of cumulative costs in these heuristics contradicts the economic logic that implies a searcher ought to only consider the costs associated with inspecting the next option. Kogut (1990) suggested that many participants who stop their search earlier than the optimal model's prescription may do so because they track the cumulative cost of search as each new option is considered (see also Friedman et al., 2007). Kogut conjectured that some participants have a rule to stop once their earnings become negative (equivalent to an earnings stopping threshold of 0). Such a participant would be described as *avoiding* a level of earnings that is too low, rather than *aspiring to* a level of earnings that is sufficiently high.

Earnings thresholds are often used as a component of even more complex heuristics. When participants were allowed to explicitly formulate a stopping rule from a menu of options, Sonnemans (1998) found that nearly 60% of participants included an earnings threshold as one part of a hybrid strategy.[1] These hybrid heuristics avoid some difficulties associated with relying exclusively on fixed-earnings heuristics. For example, heuristics that terminate only when an earnings threshold is exceeded risk stopping too late or never (Schunk, 2009). Additional conditions combined with an (earnings) aspiration level threshold ensure that such heuristics terminate.

One practical conjoined condition is for the search to end once the searcher has expended a pool of funds earmarked for paying search costs. If the concept of search cost is broadened to include time costs, there are many plausible strategies that "stop" when a time limit is exceeded. In contrast, for heuristics that terminate once earnings fall *below* an earnings threshold (e.g.,

[1] Due to the structure of this experiment, all rules that used earnings as a termination rule used an "aspirational" earnings threshold that terminated once the level of earnings exceeded the threshold. Participants were not permitted to use thresholds that caused heuristics to terminate once earnings fell below the threshold.

zero), additional conditions make it more likely that a heuristic will terminate above that threshold.

Participants sometimes end searches after sampling a fixed number of items only if they haven't found something with a value that would have ended search earlier (Sonnemans, 1998). Further, several laboratory sequential search studies without navigation have found that some consumers end searches after sampling a fixed number of items regardless of what they find (Houser & Winter, 2004; Moon & Martin, 1990). This kind of strategy does not necessarily contradict predictions of rational models. For example, simultaneous search models assume that people fix the number of items they want to look at before commencing a search (Stigler, 1961). However, it is possible that consumers also have fixed time strategies that do not consider value.

1.3.2. Variable Threshold Value Heuristics

Variable threshold heuristics change the thresholds they use for making termination decisions as the search proceeds (Martin & Moon, 1992; Moon & Martin, 1990; Sonnemans, 1998; Zwick et al., 2003). These changes are usually based on the values of options encountered earlier in the search (Butler & Loomes, 1997; Hey, 1982, 1987; Martin & Moon, 1992; Moon & Martin, 1990; Zwick et al., 2003). In other cases, heuristics reduce their thresholds as a searcher conducts more inspections, regardless of the values of options inspected along the way (Martin & Moon, 1992; Moon & Martin, 1990; Zwick et al., 2003).

Some variable threshold heuristics change over the course of a single search by adjusting thresholds based on the range, minimum, or maximum of values encountered during that search. Hey (1982) proposed a heuristic that terminates when the accumulated cost of search exceeds a fixed percentage of the highest value option encountered so far. This heuristic approaches the performance of a full-information optimal model, and it outperformed several other heuristics when the distribution of outcomes was not known prior to commencing search (Moon & Martin, 1990). Similar heuristics describe search behavior in many experiments. Schunk and Winter (2009) found that 70% of their participants behaved consistently with having a declining value threshold.[2] And, Lee (2006) found that a variable threshold heuristic explained his data in a simplified, fixed horizon search task.

Butler and Loomes (1997) have proposed a plausible, non-optimal search heuristic (although their proposal has been ignored by most researchers). They described a Level of Aspiration Model in which searchers set an initial stopping threshold based on earnings at the beginning of search. Even if the searcher is poorly informed about the distribution of possible values, she is hypothesized to rely on ill-defined prior beliefs to set an aspiration level. As she inspects more options, she adjusts the level to reflect a weighted average of all her past aspiration levels, incorporating the new information from inspected options. The searcher stops searching when she finds an option that exceeds her current aspiration level, or when her aspiration level adjusts below the earnings possible with an already inspected option.

We like this proposal because there is considerable evidence from psychological experiments that a serial averaging rule (like Butler & Loomes' proposal) describes many everyday and economic judgments and inferences (Anderson, 1981, 2014; Furnham & Boo, 2011). Furthermore, the general form of the heuristic calculation mimics the calculation of rational reservation values based on partial information, but in a computationally efficient, psychologically plausible manner. Butler and Loomes (1997) report an experiment that provides

[2] Schunk and Winter frame this results as consistency with a finite horizon optimal model with a risk aversion parameter, but it can, just as easily, be framed as a variable threshold heuristic.

support for their serial-averaging, aspiration-level model from subtle option order sequence effects on stopping.

It is important to note that the distinction between fixed- and variable-threshold heuristics is not always clear. In some cases, a threshold that is fixed in earnings is variable in value. For example, a simple fixed-earning heuristic might terminate according to the following condition:

$$v_{max} - nc < \$0$$

Where v_{max} is the value of the best option inspected so far, n is the number of inspections conducted so far, and c is the search cost in dollars. This heuristic will terminate when the earnings from search fall below \$0. Rearranging the terms of this equation makes it clear that this heuristic is also a variable value threshold. In this case, the value of the best item is compared to the threshold nc which increases with the number of searches.

$$v_{max} < nc$$

1.3.3. Value-Based Navigation Heuristics

Very few studies have attempted to test the basic assumptions of the Weitzman Model or have examined its implications in controlled laboratory experiments. One exception is Gabaix et al. (2006), that examined navigation and proposed a heuristic in which consumers order their search as though they are rational but treat each inspection as their last. This heuristic implies that people navigate based on the expected value (rather than the optimal reservation value) of locations. This model describes participants' behavior more effectively than the optimal Weitzman model in a simple task with 3 locations and binary outcomes.

Urbany (1986) examined the role that different types of expectations play in search. He found that having more specific expectations about prices at individual locations, as opposed to general expectations about a market, influenced search order and reduced search duration. But Urbany did not test implications of the model for specific patterns of individual search behavior.

Beyond Gabaix et al., there have been few studies that have attempted to identify heuristics that consumers actually use to navigate. Given the computational complexity and large amount of information necessary to calculate RVs for all possible products, however, it seems likely that consumers often rely on heuristics to simplify their searches.

1.4. Spatial Search Strategies

In navigation, spatial position can be a key criterion in deciding what option to inspect next. For example, a consumer may choose to visit the closest store next rather than visiting a more distant store with the possibility of greater gains. Or she may choose to learn more about the product at the top of the screen on a website rather than looking at all the items on the screen and clicking the one with the highest reservation value. In this section we will focus on these spatial search strategies.

In the standard economic models of search with navigation, these strategies can be modeled by applying different search costs for different locations. In the physical world, for example, more distant stores impose greater search costs. If differences in search costs are large enough, they will impact search order such that a consumer will prefer to visit a close location before a farther one, most of the time. Similarly, a tendency to learn more about the top item before proceeding progressively farther down the screen on an electronic device may involve cognitive costs.

The online case is particularly interesting because the different search costs between items on different parts of the screen seem trivial. The literature on working memory and cognitive offloading, however, suggests that there are cognitive costs associated with this type of search that can be reduced by taking advantage of the spatial layout of a website.

Including undefined variable search costs in tests of the optimal model, however, produces a challenge for the researcher in the sense that those search costs become “free parameters,” that can be adjusted to improve the fit of the model. Another “adjustable” parameter in the optimal model is associated with the searcher’s risk attitudes. For example, some theorists have proposed that variations in risk-aversion can explain search durations that are apparently too short given a simple application of the Basic Model with a risk-neutral assumption.

Adjusting the optimal model using variable costs and risk parameters analyzes behavior at a different level than we on which we focus. Fitting these parameters to minimize statistical error allows for effective prediction but does not provide much insight into the psychology of searchers. Our aim is to create a principled cognitive account of *why* search costs vary between locations and *how* consumers use available cues in the environment to navigate. By understanding search in this way, we gain more insight into how consumers search, and we develop a set of principles that allows us to better predict consumer search behavior in novel situations.

With respect to spatially-driven search, at minimum we can identify conditions under which there is a clear shift away from the risk neutral Weitzman Model with fixed costs and an obvious reliance on spatial layout to guide navigation. Given the modest deductive power of the rational model approach, when central parameters defining costs and values are weakly constrained by theory, there will always be room for a defense of the optimal model. Our intention is to provide clear behavioral results to inform a sensible discussion of the relative theoretical merits of optimal versus heuristic models as descriptive of actual behavior.

1.4.1. The Concept of Working Memory

Working memory is a buffer that holds information which consumers can transform and process (Baddeley, 2001). People likely have multiple working memory systems, specializing in areas like visuospatial or auditory information (Baddeley, 2001). Importantly, working memory is limited, often requiring a significant amount of effort to encode (Logie, 1995) and maintain (Baddeley, 2001; Baddeley & Hitch, 1974; Jansma et al., 2007) information in a working memory store. These limitations are central to many of the mental operations associated with human reasoning (Baddeley & Hitch, 1974; Logie, 2011). The use of different strategies in multi-attribute choice problems can be attributed to these limitations (Johnson & Payne, 1985; Payne et al., 1993). For example, consider choices involving a series of pairwise judgments, either against a threshold (e.g., elimination by aspects; Tversky, 1972) or against other items (e.g., the majorty of confirming dimensions heuristic; Russo & Dosher, 1983). These pairwise judgments require less working memory capacity than full comparisons, since they only require a decision maker to hold information about two items in memory (Shah & Oppenheimer, 2008).

1.4.2. The Impact of Working Memory on Search

Working memory also likely influences navigation in search. Product displays presented in lines (vertical or horizontal), grids, or clusters can all help to alleviate working memory load and simplify search. As an illustration of how a linear display can influence navigation, let's begin with a shopping situation based on the Weitzman model. Imagine a consumer standing before a shelf arrayed horizontally with 12 different varieties of honey. The consumer, from experience, has varied expectation about each jar. He needs to spend a moment reading the label to determine exactly how much he will like a given jar, but can form a reservation value just by glancing at it. The consumer wants to buy the best honey possible without expending too much effort.

How might a fully rational consumer with no working memory or computational constraints pick a product? First, even though the consumer has expectations about each jar of honey, until he peruses the whole shelf, he will not know which twelve jars of honey are available or where each jar of honey resides on the shelf. Next, he would store several pieces of information in memory. He must remember what honey is available, what expectations he has for the honey, and where each jar of honey is located on the shelf. Finally, he can then proceed to calculate the reservation value for each jar of honey and proceed with the Weitzman algorithm for search with navigation by looking at the jars of honey in order of their reservation values. See the left panel of Figure 1 for a schematic of this process.

Without infinite and effortless working memory, a consumer may instead rely on pairwise comparisons, essentially conducting a roughly structured tournament to identify the "winning" product. For example, a consumer could randomly select two jars of honey and compare their reservation values. Then the consumer could retain the highest reservation value and randomly select another, previously unexamined jar of honey for comparison. One notable feature of this situation is that the consumer is storing information about reservation values *in the environment* and only accessing them when needed. Rather than storing all 12 reservation values in memory and selecting the highest one, the consumer only extracts data from the environment (by glancing at one of the jars) when he needs it.

Accessing information in this just-in-time manner allows a consumer to go from storing 12 reservation values to storing 2: the maximum reservation value encountered so far and the reservation value of the jar that he's currently inspecting. After conducting pairwise comparisons, the consumer returns to the jar with the highest reservation value and inspects it to see whether its value is high enough to warrant ending the search.

1.4.3. The Concept of Cognitive Offloading

Cognitive offloading occurs when people use their environments to help them remember and think in the face of cognitive constraints (see Risko & Gilbert, 2016 for a review). For example, people choosing among many options read information off of items only as they need it to make specific comparisons (Ballard et al., 1997; Droll & Hayhoe, 2007). This use of the environment to support memory and inference (Land, 2014; O'Regan, 1992) is illustrated in the example above where the consumer only determined the reservation value of an item when it was needed for a specific pairwise comparison, quickly discarding information about items with inferior reservation values.

More broadly, principles of cognitive offloading suggest that consumers are adaptively deciding whether to expend effort in *storing* more information about products in working memory or finding it later in the environment, as needed (Ballard et al., 1997; Gray et al., 2006; Schonpflug, 1988). One trivial example of this phenomenon occurs when people use writing as a way of reducing the load on working memory (Risko & Dunn, 2015).

Other research suggests that people also reduce the effort of *processing* information by using or altering their environments (Clark & Chalmers, 1998; Hutchins, 2010; Risko & Gilbert, 2016). For example, people will spontaneously rotate their heads when presented with images and text in non-standard orientations (Risko et al., 2014). By rotating their heads, they can substitute the cognitively costly task of mental rotation with the less effortful muscular task of moving their heads (Kirsh & Maglio, 1994). More broadly, consumers may reduce cognitive effort by moving distant items closer together (Kirsh, 2010), changing the orientation of items

(Kirsh & Maglio, 1994), and sorting items (Kirsh, 1995). These strategies often appear in concrete examples of consumer search and choice.

1.4.4. The Impact of Cognitive Offloading on Search

One feature of the illustrative search for a jar of honey is that the consumer *randomly* looks at jars for the pairwise comparison. But intuition and behavioral research suggest that people would look sequentially across the products, from left to right (or up and down). Indeed, both humans and capuchin monkeys tend to search sequentially along rows and columns when the environment permits it (Smith & De Lillo, 2022). But why is scanning horizontally in this case the intuitive thing to do?

To effectively execute the strategy described in the example, the consumer needs to store two additional types of information: the location of the highest reservation value item and the locations of all items examined so far. The location of the highest reservation value item so far is needed because the consumer must return to that item at the end of the search. The location of all items examined so far must be stored to prevent a consumer from accidentally revisiting previously inspected items and wasting effort[3]. With a small number of products, this strategy makes sense. A person might remember the locations of four or five items that he has already inspected. With a larger number of items, remembering the location of previously inspected items is likely to become more difficult, resulting in more revisiting errors. See the right panel of Figure 1.4.1 for a schematic of this process.

Scanning across a horizontal array uses the environment to reduce working memory loads. As the consumer scans from left to right across the shelf of honey, he knows that he has

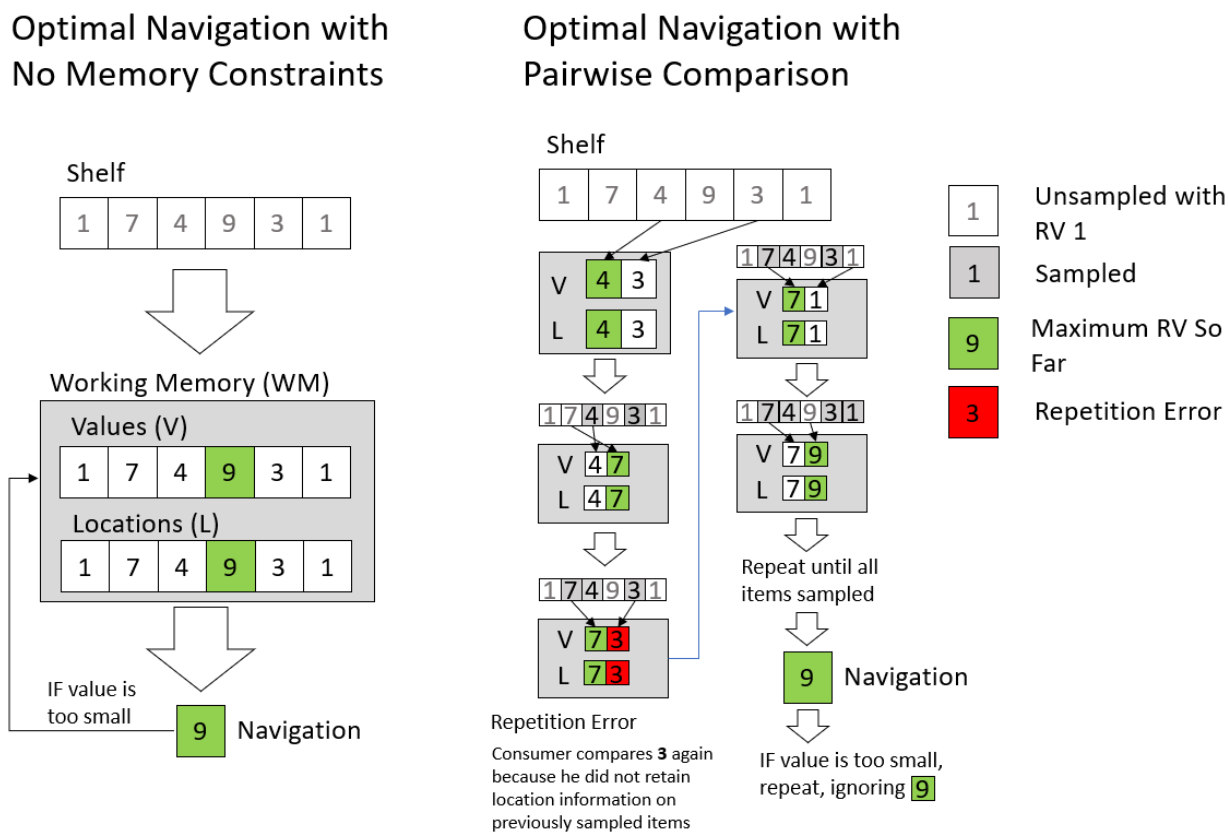
[3] Implicitly, when a consumers remember the locations items they have visited, they also remember that they have visited that item.

already looked at everything to the left of the item he is looking at and has not looked at anything to the right of that item. That is, the consumer reduces the memory load of the task by using the position of his gaze rather than his working memory to track what he has already looked at. See the left panel of Figure 2 for a schematic of this process.

By using his gaze like this, he can make his pairwise comparisons without having to remember the locations of what he has already looked at. Now, he can conduct this search while storing only the reservation value and location of the highest valued item found so far and the reservation value of the current item. A similar mechanism has been proposed to explain why both human adults and capuchin monkeys tend to search in lines in certain environments (De Lillo et al., 1998, 2014; Smith & De Lillo, 2022). And, in a sequential search task where options were arrayed vertically but could not be differentiated (therefore not allowing for navigation), participants whose searches proceeded from top to bottom had better performance than those who searched using other patterns (Caplin et al., 2011).

If our hypothetical consumer were to use this strategy, he would scan the entire shelf each time he makes a navigation or stopping decision. For example, once he determined the value of the honey jar with the highest reservation value, he would repeat the process while excluding the item he already looked at. Note, however, that even in this case, the consumer would have to track which of the jars of honey he already inspected, which could become taxing after inspecting a few products. Ultimately, this strategy would still lead to choices consistent with the navigation decisions in an optimal strategy.

Figure 1.4.1. Comparison of Search With No Working Memory Constraints (left) with a Simple Pairwise Navigation Model that Extracts Reservation Values (RVs) Randomly (right)



So far, we have assumed that people have limited and costly working memory but infinite computational ability. In our example, once a person has information about a location, he can discern the reservation value without cost. But anyone who has tried to integrate over probability density functions in their head knows that mental computations are costly. And, as described earlier, there is substantial evidence that people in sequential search situations simplify computations to make stopping decisions (e.g., Butler & Loomes, 1997; Caplin et al., 2011; Hey, 1982; Sonnemans, 1998).

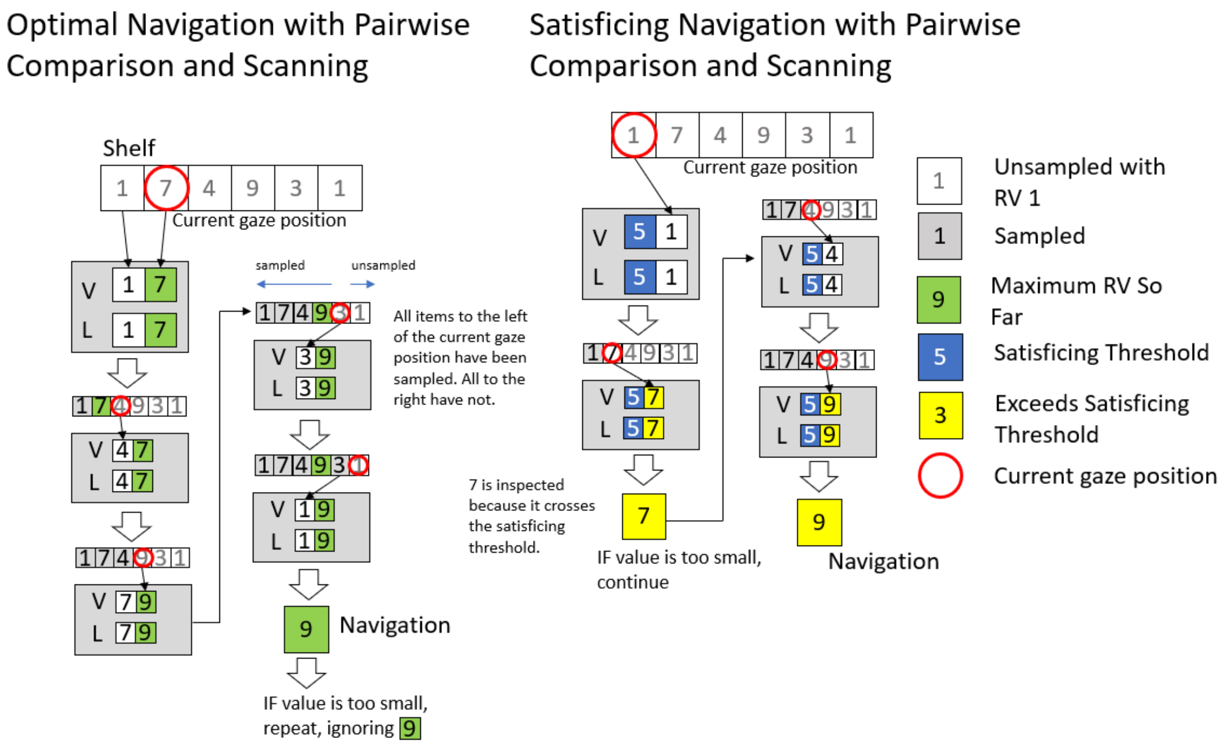
The optimal strategy for sequential search with navigation is significantly more computationally intensive than for that in search without navigation. A reservation value must be computed for each location to decide where to go. Furthermore, consumers using a just-in-time strategy would have to repeatedly compute reservation values for each item as they make pairwise comparisons for each navigation and stopping decision. The computational burden of doing this, even when using a simplifying heuristic to approximate reservation value, can be high.

One way to simplify navigation decisions dramatically is to combine the just-in-time left-to-right strategy with satisficing. Traditionally, a consumer using a satisficing heuristic will set a threshold that is “good enough” and pick the first item that crosses that threshold (Simon, 1955, 1956). As described earlier, there is evidence that people satisfice in sequential search when deciding when to stop (Caplin et al., 2011; Kogut, 1990; Schunk & Winter, 2009; Sonnemans, 1998). In navigation, a person can satisfice by inspecting items when they exceed a certain threshold. Unlike the case with the optimal navigation strategy, the order of presentation matters in this satisficing navigation strategy. Whereas a consumer abiding by the optimal model will always look at the highest reservation value item first, a person using a satisficing navigation strategy may not if a sufficiently attractive, but non-optimal, location is presented to them early in a search.

When we combine satisficing with the just-in-time left-to-right strategy, we get a new pattern of behavior which we call the Scanning Search Strategy. In the Scanning Strategy, a person will scan across an array of products only once and examine carefully all items that are promising enough to warrant inspection. See the right panel of Figure 1.4.2 for a schematic of this process.

If our consumer shopping for honey were to use this strategy, he would go from left to right and inspect any item that seemed promising starting on the left of the shelf. For example, even if the second honey from the left isn’t the product with the higher reservation value, our consumer would inspect it if its reservation value were sufficiently high. A consumer relying on this strategy would have a set of inspections that moves only in one direction. In our example with honey, each subsequently inspected jar of honey will be to the right of the previous jar.

Figure 1.4.2. Optimal Pairwise Navigation with Scanning (Left) and Satisficing Pairwise Navigation with Scanning (Right)



In this section, we have highlighted how constraints on human working memory and computational abilities can combine with an ability to use the environment to support working memory to produce a new search strategy. We have walked through the logic a specific Scanning Strategy, to illustrate how these conditions might give rise to search strategies that are influenced by the spatial configuration of products. Similar explorations would uncover other spatial heuristics in consumer environments.

One key implication of this type of strategy is that complexity of a search situation and the layout of the environment jointly determine whether consumers rely primarily on value or spatial cues to navigate. For example, when a display has only a few items, a consumer may navigate based *on value* because the cognitive cost of doing so is relatively low. But when a display has many items, consumers will switch to navigating based *on spatial cues*. Similarly, even when there are many items, if information displays make it difficult to make comparisons or easily track what items have already been inspected, consumers will rely less on spatial strategies. In this way, consumers engage in a value-based versus spatially-based strategy tradeoff as they search in order to maximize outcomes while minimizing cognitive effort.

1.5. The Present Research

In the present research, we examine consumer search, focusing primarily on how consumers *navigate*. In Chapter 2, we examine how consumers navigate and stop when placed in an experimental search environment where they have differing expectations about products. We compare participants' navigation and stopping choices to both the optimal model and to simple heuristic models to better understand what information consumers use and how they are using it.

In Chapter 3, we examine the impact that environment factors have on search strategies. We examine how increasing the number of choices impacts navigation patterns, independent of the information typically integrated into formal models. Furthermore, we look at how spatial layouts of products can change the information used to navigate and stop. In these studies, we compare participants' navigation choices to both value-based and spatially-based strategies.

In Chapter 4, we examine learning in two studies in which participants repeatedly search in the same environment and receive global feedback about their success. We examine how strategy use changes as participants gain experience.

Across these studies, we seek to answer key questions about how consumers search. First, how do consumers' search behaviors compare to the predictions of optimal models? Throughout our studies, we compare participant behaviors to the optimal model and identify conditions under which the optimal model provides a good description. Second, when consumers are not searching optimally, what are they doing instead? In Chapter 2, we focus mainly on value-based heuristics that consumers use to simplify search tasks. In Chapter 3, we introduce spatially-based strategies and find that they describe aspects of search in moderately complex search environments. Finally, when are consumers likely to use non-optimal strategies? In Chapter 3, we examine how the complexity of the situation impacts the use of optimal and non-optimal strategies. In Chapter 4, we focus on how learning influences search strategies when searchers receive repeated experience performing similar search tasks.

2. Chapter 2: Value-Based Navigation

In Chapter 2, we examine Value-Based Search. In Study 1, we look at stopping rules. These stopping rules are important because the Reservation Values used to decide when to stop in the optimal model are also the basis of navigation in that model. In Studies 2 and 3, we look directly at navigation by having participants search among five locations. In Study 2, we present information about likely outcomes at each location numerically; in Study 3 we present the information graphically.

2.1. Study 1: Estimating Reservation Values

In Study 1, we examined how participants' stated RVs varied based on the distributions of values for a single source. Essentially, we asked participants to report their stopping rules, in the form of numerical thresholds to terminate search.[1] Importantly, in optimal models of search, these stopping rules also form the basis for navigation, as people decide their search ordering based on RVs. Our objective is to take a first look at stopping strategies, under conditions where we force participants to think in terms of criteria to terminate search. In Studies 2 and 3, we examine whether the rules used to think about these stopping criteria also influence navigation.

2.1.1. Method

Participants. We recruited 100 adults (65% male) on Prolific Academic; no data was excluded from our analyses. The experimental sessions averaged approximately 15 minutes in duration. Participants were paid a participation fee of \$1.82, and they earned bonuses contingent on the optimality of their responses, ranging from \$1.78 to \$1.31.

Design. We manipulated the maximum and the minimum of uniform distributions of

[1] For the remainder of this article, we refer to the optimal reservation values derived from the Weitzman model as "RVs" and the estimates of RVs provided by participants as "stated RVs."

payoff values and the search costs (detailed stimulus parameters are presented in Appendix A, Table A1). We employed a 3 (Minimum Value: Low (value = 5) vs. Medium (25) vs High (45)) X 3 (Maximum Value: Low (55) vs. Medium (75) vs. High (95)) X 3 (Search Costs: Low (1) vs. Medium (5) vs. High (9)) within-subjects design.

Procedure. Participants were instructed to imagine that there were two boxes with prizes in them. Participants were shown the bounds of the uniform distributions from which the prizes in each box would be drawn and the cost of opening the second box. They had already opened one box and learned the prize value and were asked if they wanted to inspect the second box. If they elected to open both boxes, they kept the larger prize and paid the search cost.

Each of our 27 trials involved a single question that we interpret as asking for the participant's stated RV: "If I find a prize worth less than or equal to [value] in the first box, I want to open the second box that contains a prize worth some value from 1 through 100." Recall that the RV is the value of a hypothetical product a consumers could find during a search which would make the expected gain from additional searches negative. This procedure simulates the decisions made according to a RV rule in the Basic Search Model by asking them to state this hypothetical value.

We did not provide participants with any feedback until the end of the experiment to minimize learning during the session. Finally, we aggregated the points across all rounds and multiplied them by 0.001 to arrive at a performance bonus in dollars.

Appendix Table 6.1.1 shows the stimuli, the RVs predicted by the normative, risk-neutral Basic Search Model. Stimuli were blocked according to cost, such that each participant saw all nine boxes for each cost level in one block. Within blocks, boxes were presented in random order, and order was balanced so that each random sequence and its reverse were presented (to different participants). Prior to these 27 trials, all participants completed two practice trials which we discarded without analysis.

2.1.2. Results and Analysis

Responses to the Minimum, Maximum, and Search Cost. We depict the relationship between stated RVs and the minimum, maximum and search costs graphically in Figure 2.1.1. Stimuli and average responses are presented numerically in Appendix Table 6.1.1. We regressed participants' stated RVs onto the minimum, maximum, and search costs, using cluster-robust standard errors to account for the presence of multiple observations per participant (Table 2.1.1). We note two responses to our manipulation. First, participants decreased stated RVs as search costs increased. The coefficient on Cost in Model (1) shows that this relationship is statistically significant, $b = -1.086$, $t(2695) = 5.872$, $p < 0.001$. However, a comparison with the gray curves in Figure 1, shows that participants did not respond to increases in search costs as strongly as the normative model; note the narrower gap between the black and gray (dashed) lines in the panels on the right side of Figure 2.1.1 compared to those on the left.

Second, participants increased their stated RVs in response to increases in both the minima ($b = 0.337$, $t(2696) = 12.118$, $p < 0.001$) and the maxima ($b = 0.506$, $t(2696) = 21.680$, $p < 0.001$). However, changes in stated RVs as maxima increased were not as large as those implied by the rational model. In Figure 2.1.1, this relationship is visible in the flatter black lines compared with the optimal gray lines.

One prediction of the rational model is that consumers ought to weight the upper end of a distribution more than the lower end. To test if this is the case, we compared Model (1) to a restricted model where the minimum and maximum have the same coefficient, presented as Model (2). When we compare the constrained and unconstrained models using a Wald test, we find that the minimum and maximum have different coefficients suggesting that participants weight the maximum (b = 0.506) more than the minimum (b = 0.337), consistent with the implications of normative model.

Table 2.1.1. Regression of stated RVs on the minimum, maximum, and search costs in Study 1

	(1) Full	(2) Restricted
Minimum	0.34***	
	(0.03)	
Maximum	0.51***	
	(0.02)	
Min + Max		0.42***
		(0.01)
Cost	−1.09***	−1.09***
	(0.18)	(0.18)
Order	−2.37	−2.37
	(1.90)	(1.90)
Obs.	2700	
Wald Test	$F(1,99)=13.6$	$p < 0.001$
R^2	0.32	0.31
BIC	22440.61	22476.61

Note. ***:p<0.001; **:p<0.01; *:p<0.05. In the full model, minimum and maximum are allowed to have separate coefficients. In the restricted model, we constrain the coefficients on minimum and maximum so that they must be the same. The coefficient on "Min + Max" in model 2 is the constrained estimate. This regression uses cluster robust standard errors to account for the within-subjects design.

Responses to Variance of the Distribution of Values. Consumers ought to have higher stated RVs for options with greater variation, holding everything else constant. After controlling for Cost and Expected Value (EV), the relationship between Range and stated RVs remains positive, b = 0.085, $t(2695) = 3.691$, $p < 0.001$ (Model(1) in Table 2.1.2). In Model (2) of Table 2.1.2, we added the interaction between Range and Cost to Model (1). The coefficient on this interaction is negative, $b = -0.020$, $t(2695) = 4.735$, $p < 0.001$. It suggests that for each unit increase in cost, the impact of an additional point of range is reduced by 0.02 points. A model comparison between a model with and without the interaction shows that the two models are different, $F(1, 99) = 22.5$, $p < 0.001$, supporting the conclusion that cost moderates the impact of variance (range), as expected in the optimal model.

Figure 2.1.1. Average stated RVs (reservation values) by Box Minimum, Box Maximum, and Search Costs in Study 1 (Horizontal black lines are box ranges; light gray lines are risk-neutral normative model predictions) in Study 1.

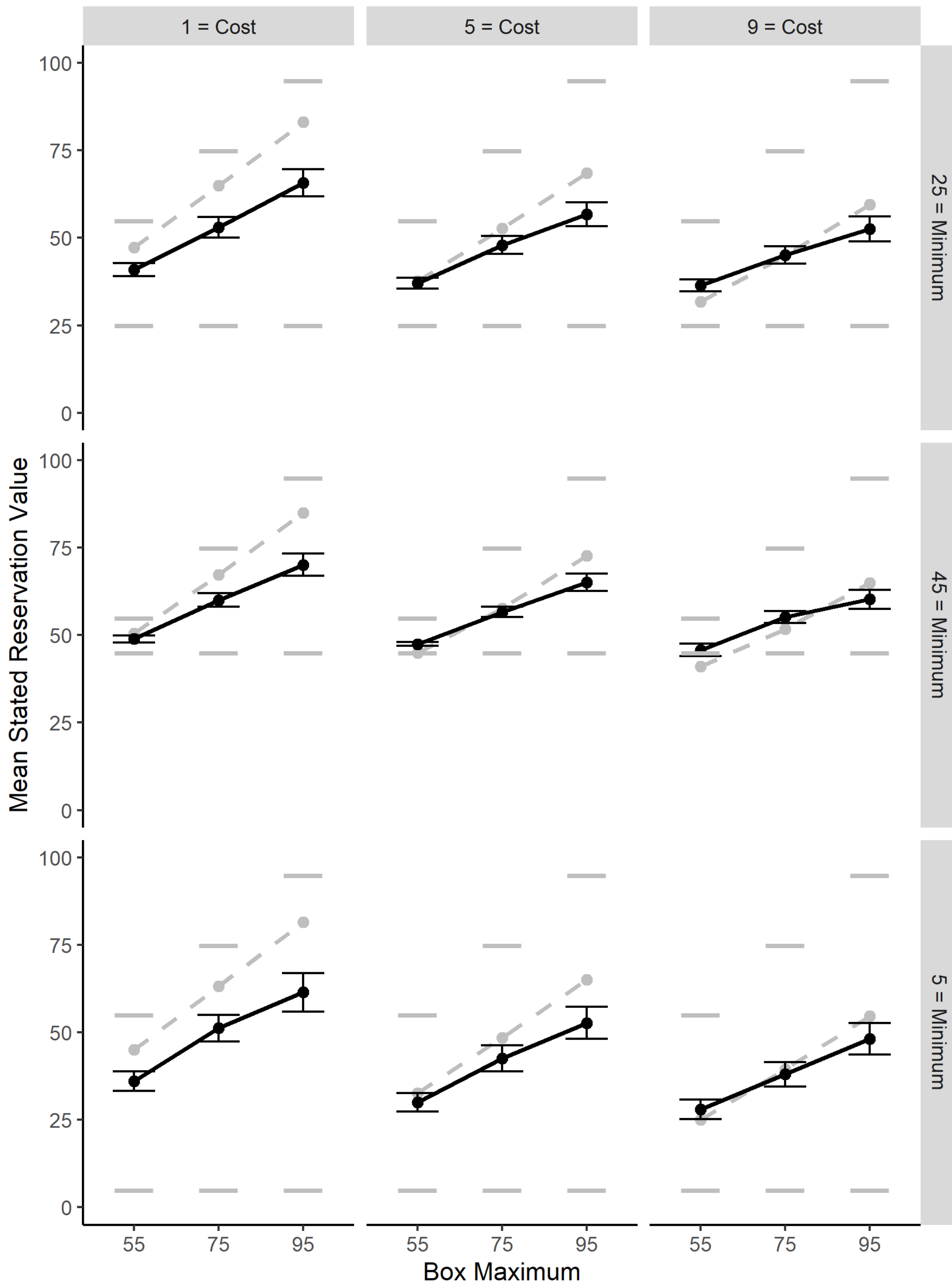


Table 2.1.2. Regression of Stated RVs onto Ranges, EVs (expected values), and Search Costs in Study 1

	(1) No Interaction	(2) Interaction
Range	0.08***	0.08***
	(0.02)	(0.02)
EV	0.84***	0.84***
	(0.02)	(0.02)
Range X Cost		−0.02***
		(0.00)
Cost	−1.09***	−1.09***
	(0.18)	(0.18)
Order	−2.37	−2.37
	(1.90)	(1.90)
Observations	2700	
Wald Test	$F(1,99)=22.5$	$p < 0.001$
R^2	0.32	0.33
BIC	22440.61	22423.31

Note: ***:p<0.001; **:p<0.01; *:p<0.05. This regression uses cluster robust standard errors to account for the within-subjects design.

Comparison to Risk Neutral Rational Model. Each of the 27 boxes has a normative RV calculated from its minimum, maximum, and search cost (see Appendix Table 6.1.1). Many participants perform close to normatively (mean RMSE = 16.17, SD = 6.62, MDN = 15.01). As a benchmark, we calculate the RMSE of the EV = 11.54. The mean participant RMSE is significantly higher than that of the simple EV strategy, $t(99) = 6.999$, $p < 0.001$; 24% of participants have an RMSE lower than the RMSE for the EV strategy.

Next, we regress participants' stated RVs onto normative RVs with cluster-robust standard errors to account for multiple observations per participant. This model shows that stated RVs are correlated with normative RVs, $b = 0.648$, $t(2698) = 20.990$, $p < 0.001$. However, compared with our best fitting model (Model(2), Table 2), this model has a lower R^2 (R^2 = 0.331 versus $R^2 = 0.294$) and a higher BIC (best BIC = 22,423 versus RV BIC = 22,535). Again, we can compare this regression to one where stated RVs are regressed onto EVs. The EV is also correlated with participant responses, ($b = 0.843$, $t(2698) = 37.008$, $p < 0.001$), but the model based on the RV has a higher R^2 (RV $R^2 = 0.294$ versus EV $R^2 = 0.273$) and a lower BIC (RV BIC = 22,535.55 versus EV BIC = 22,614.95).

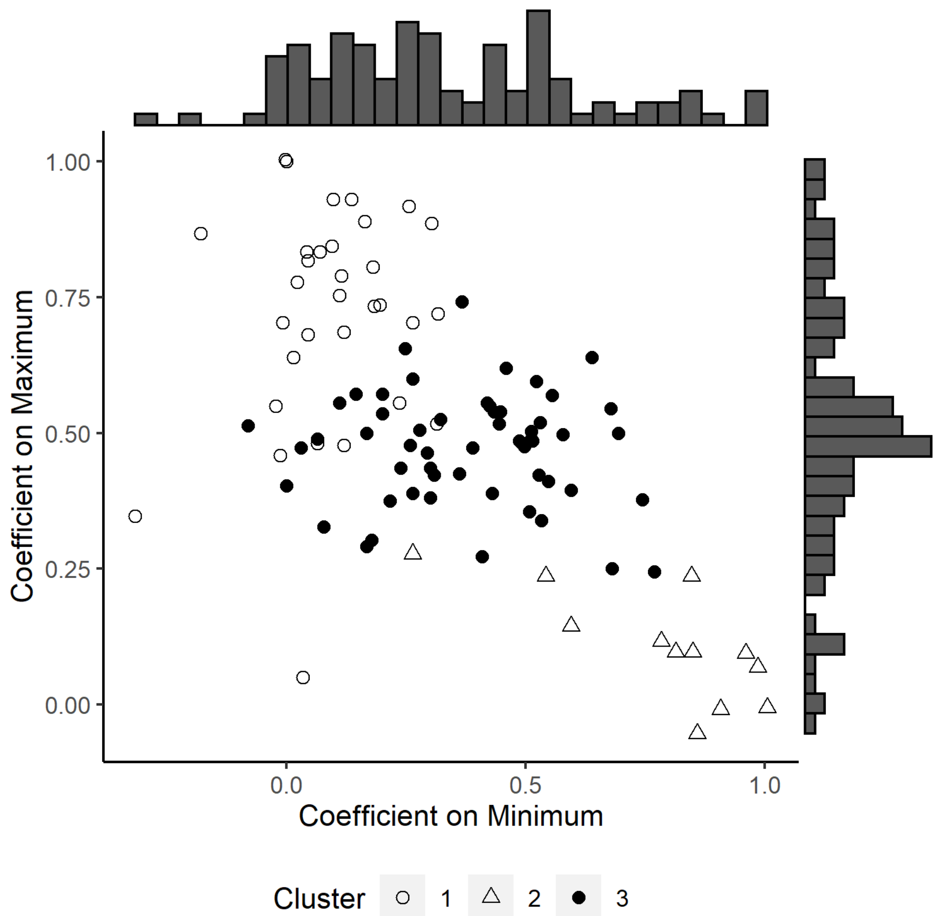
Individual Differences in Strategies. If we look underneath the aggregate average performance statistics, we find evidence of reliable individual differences in strategies to estimate RVs. We applied three analytic methods to sort participants into "strategy types." First, we conceptualized strategic heterogeneity in terms of differences in individual weights on the two values defining the payoffs available from each source (the maxima and minima of the "boxes" in our experimental materials). Figure 2.1.2 is a scatterplot with points representing participants' regression weights on the Minima and Maxima of the boxes for which they judged RVs.

We followed up this descriptive analysis with a cluster analysis of participants' stated RVs using the partition around medoids (PAM) algorithm (Kaufman and Rousseeuw, 1990). A three-cluster solution provides insights into the Minima-Maxima weight plot (see Figure 2.1.2). Participants in Cluster 1 (N = 32) were sensitive to box maxima and search costs and relatively insensitive to box minima. These participants are best described by the RV strategy. Participants in Cluster 2 (N = 12) were sensitive to box minima and relatively insensitive to box maxima. Surprisingly, participants in this cluster increased stated RVs in response to increases in search costs. Perhaps these participants could be described as exhibiting a "loss aversion" reaction to search costs and over-weighting potential values on the low end of the distributions. Participants in Cluster 3 (N = 56) were moderately sensitive to search costs and equally sensitive to the minima and maxima of distributions, they are best described as following an EV-based stopping rule. We graphically present means by cluster for each stimulus in Appendix Figure 6.1.1.

Finally, we took a more conceptually driven approach and proposed seven "heuristic strategies" based on our understanding of plausible solutions to the experimental task (see Table 2.1.3). For half of participants, the best-fitting heuristic was based on the EV (EV and EV-Cost), while for 28% of participants the normative RV heuristic fit best. For the sample overall, the two EV-based heuristics have the lowest RMSEs, followed closely by the RV heuristic. There is also a small sub-set of participants (approximately 20%) whose behavior is best described by strategies that focused on the Maximum (the MAX strategy) or the Minimum (the MIN strategy) values. The maximum (minimum) strategy values boxes (and therefore navigates) according to the maximum (minimum). Note this breakdown maps onto our prior empirically-driven methods of identifying strategic sub-types (cf. Figure 2.1.2).

When we look at mean RMSEs only for participants for whom each heuristic fit best, we see that they are much lower than they are for the sample overall, implying that there is reliable heterogeneity in strategies used by participants. For example, for a small number of participants, the MAX-Cost or MIN+Cost heuristics describe their behavior better than estimation rules based on the EV or RV. If we examine the RMSEs for each heuristic among participants for whom the heuristic fits best, the EV, EV-Cost, MAX-Cost, and MIN+Cost rules fit nearly perfectly for small sub-sets of participants. Notably, no participants had very low RMSEs for the RV strategy. Our interpretation is that a substantial number of participants were thinking in terms of the RV, expected gain principle, but that their implementation of the cognitively demanding strategy was only approximate.

Figure 2.1.2. Scatter plot and histograms of participants by individual regression weights on the Minima and Maxima of the values defining the boxes for which they estimated RVs in Study 1.



Note. Symbols (o, •, Δ) indicate sub-types of participants following different search strategies identified by a cluster analysis (see text for discussion). A histogram of coefficients on the minimum are above the scatterplot and one for coefficients on the maximum are to the right.

Table 2.1.3. Heuristic fit of Stated RVs in Study 1

			Best		Total	
Heuristic	Width	Best Percent	Mean	SD	Mean	SD
RV	Normative Reservation Value	28%	11.2	4.0	16.2	6.6
EV	Expected Value	23%	9.9	4.8	15.0	5.5
EV - Cost	Expected Value - Cost	27%	10.8	5.4	15.1	5.7
Max - Cost	Maximum - Cost	10%	7.8	4.2	25.6	10.2
Min + Cost	Minimum + Cost	12%	9.7	4.8	27.4	9.7
Max	Maximum	0%			30.0	10.2
Min	Minimum	0%			30.5	9.9

Note. The Best columns show the means and standard deviations among participants for whom each heuristic had the lowest RMSE. The sample columns show the means and standard deviations for the whole sample.

2.1.3. Discussion

In Study 1, we asked participants to state RVs while varying search costs and the distribution of possible values contained in boxes. These RVs are central to both navigation and stopping in the Weitzman model. On average our participants conformed to predictions derived from the optimal search model: They decreased stated RVs as search costs increased and increased stated RVs as box minima and box maxima increased. Furthermore, they placed more weight on changes to the maximum (vs. minimum). They also increased stated RVs as variance increased. Their most distinctive departure from the optimal model was to under-search when search costs were low. But, they were less sensitive to costs and maxima compared to the exact implications of the optimal model.

When we looked underneath global averages, we saw reliable heterogeneity in responses between participants. Separate regressions for each participant revealed significant heterogeneity in the weights placed on the maximum and minimum. The largest group of participants had estimates that were best described by heuristics based on EVs. Another slightly smaller group of participants had estimates best described by the normative RV principle. Finally, a few participants' estimates were best described by heuristics based on only the minimum or only the maximum. In addition, while most participants (65%) had best-fitting strategies that involved searching less as search costs increased, 23% were best described by heuristics that did not incorporate cost at all. And, surprisingly, 12% of participants followed heuristics that led to searching more when costs increased. This unexpected behavior was corroborated by a cluster analysis, where Cluster 2 exhibited higher stated RVs in response to increasing search costs.

Taken in combination, these three analyses suggest that our data capture qualitatively different strategies. Participants who we typed as adhering to RV-based strategies weighted the maximum heavily; those adhering to EV-based heuristics weighted the maximum and minimum equally. And, the small group of participants who searched more as costs increased and weighted the minimum heavily seem to be relying on an anomalous non-optimal heuristic. Thus, even within this simple experimental search task, there is considerable heterogeneity in participant strategies with substantial numbers of participants systematically departing from the rational model.

2.2. Study 2: Search with Navigation with Numerical Stimuli

In Study 1, we found that most stated RVs were consistent with strategies using the optimal RV or the EV. In Study 2 we shift the focus to observing navigation among five uncertain options to see whether navigation behavior can be described based on strategies using the optimal RV or EV. To do so, we created a task that reflects the conditions specified for the Weitzman search model. Participants could inspect up to five options ("prize boxes") and paid a fixed search cost (five points) to open each box.

2.2.1. Method

Participants. We recruited 100 adults (46% Male, 52% Female; Average Age 42 years) to complete an experimental search task in a 15-minute session. Participants were paid a $1.71 fixed participation fee and earned an additional average of $0.55 in bonuses.

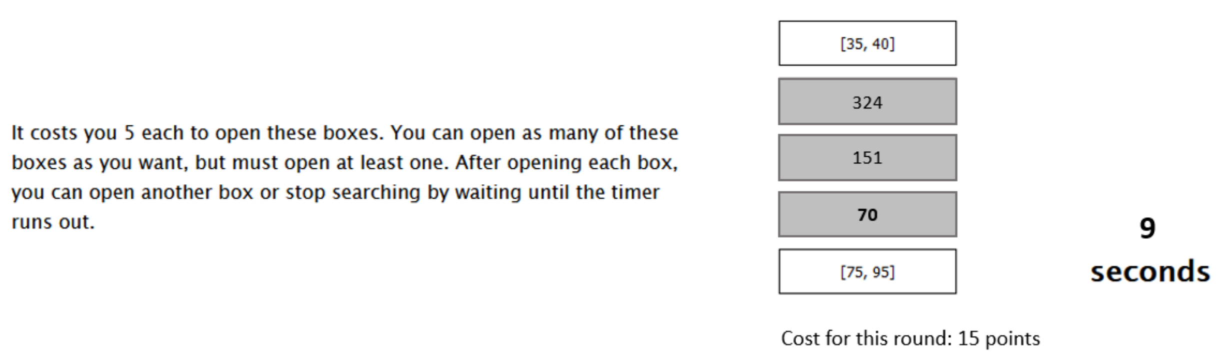
Design. The study design was a 6 (unique sets of five prize Boxes) X 2 (Repetitions) within-subjects design. Each participant completed 12 searches, with five boxes in each search round. We designed six sets of boxes to distinguish between different cues and strategies that participants might use to navigate based on the results of Study 1. In each of our sets, search orders for the rational RV strategy were distinct from those for the EV heuristic. We present our stimulus sets, as well as the navigation order implied by the RV and EV strategies in Appendix

Table 6.2.1.

Procedure. We told participants that they were looking for a high value within five boxes that contained point values that would be converted into a monetary payoff at the end of the experiment. Participants paid five points each to open one or more boxes on each search. Each box was labeled with the bounds of the uniform distribution from which a value would be drawn. When participants clicked on a box, the uniform distribution would be replaced with a randomly drawn prize value. The accumulated search costs were displayed at the bottom of the screen. Figure 2.2.1 shows the interface for a participant who has opened 3 boxes.

Participants had 20 seconds to complete each of 12 five-box searches. Once participants terminated search, they waited until the timer ran out and then moved to the next round. For each search, a participant's point total was the maximum value that they found minus their search costs. For example, if a participant opened three boxes and found a maximum of 324 points, she would receive 309 points at the end of the round (324 points – [3 x 5], given a 5-point search cost for each of 3 inspections). To ensure that participants understood the task, they answered five comprehension questions, and completed 2 practice searches. At the end of the session, we randomly selected one search and paid participants 0.5% of the points they earned in that round in dollars. For example, if a participant won 324 points, she would receive a $1.62 bonus.

Figure 2.2.1. Interface display shown to participants in Study 2



2.2.2. Results and Analysis

Descriptive Statistics. Overall, participants inspected 2.32 options on average (SD=1.22). Participants' mean per-round winnings were 100.33 (SD = 22.32).

For each search, we drew prizes for both the opened and unopened boxes. Using these numbers, we can generate counterfactual earnings for how much the optimal Weitzman Model *would* have earned in the same situation that the participant faced. By dividing a participant's realized earnings by these counterfactual earnings, we estimate participants' percentage of optimal earnings. Average earnings were 96% of the normative model.

Fully Optimal Searches[2]. 23% of searches perfectly matched the Weitzman Model's (RV-consistent) navigation and stopping prescriptions. The median participant exactly matched the rational strategy in two searches out of 12, only two participants matched the RV strategy more than 80% of the time (10 out of 12 searches), and 32% of participants never matched the rational model.

Navigation. Our stimuli were designed to distinguish between participants navigating with the optimal RV strategy and with EV heuristics. Participants' first choices (Figure 2.2.2, Panel A) were consistent with the optimal RV strategy 41% of the time, and they were consistent with the EV heuristic 32% of the time. However, first choices with the EV heuristic often overlap with a strategy that navigates towards the highest minimum value (the MIN strategy). We were only able to separate the MIN strategy from the EV heuristic in Stimulus Sets 1 and 6. Similarly, first choices with the RV strategy always overlap with a strategy that navigates based on the highest maximum (the MAX strategy). When we examine participants' entire paths

[2] All searches potentially had two normative paths which could result in different levels of earnings. In cases where earnings between normative paths diverged, we used the average of the earnings.

(Figure 2.2.2, Panel B), we find a similar pattern. 33% of the paths fit the RV strategy and 21% fit the EV heuristic.

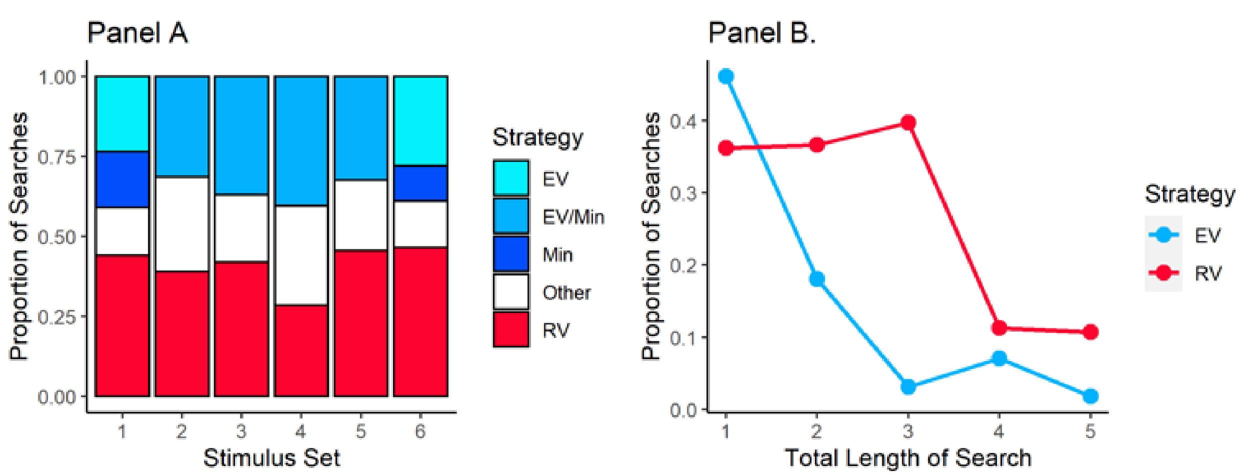
We also break down strategy use by search length to see whether different strategies are associated with different search lengths. Searches with a single inspection are more likely to be consistent with the EV heuristic (46%) than the RV strategy (36%); however, in longer searches, the RV strategy becomes more common than the EV heuristic. To compare search length to navigation strategies, we regressed total length of search on dummies for the RV, EV[3], and other strategies, setting the RV dummy as the reference level. We include fixed effects for the set and cluster standard errors at the participant level to account for multiple observation per participant. We run separate regressions for both full paths and first choices.

Overall, full paths consistent with the EV heuristic are 0.62 inspections shorter than those using the RV strategy, $t(1192)=7.445$, $p < 0.001$. Those using strategies other than EV and RV have searches that are 0.84 inspections longer than those using the RV strategy, $t(1192)=11.957$, $p < 0.001$. When we classify strategies based on first choices, we again find that participants who select the EV first choice have searches that are 0.42 inspections shorter than those using the RV strategy, $t(1192)=5.224$, $p<0.001$ while those using other strategies have searches that involve 0.26 more inspections, $t(1192)=3.077$, $p =0.002$. These analyses imply that participants following the EV heuristic tend to engage in shorter searches than those who follow the RV strategy, consistent with the RV strategy relying more on the upper part of a distribution of possible values, compared to the EV heuristic.

[3] RV in this dummy includes both paths that are purely consistent with the RV navigation strategy and those consistent with both the RV and Max strategies. Similarly, the EV dummy includes both paths that are purely consistent with EV as well as those consistent with both EV and Min Strategies.

In our first analysis (above), if a person mistakenly selects a non-RV box on the first inspection, their entire path is dropped from the RV strategy count. To account for this, we also analyze individual decisions by looking at at whether the decision is consistent with a heuristic given the locations that are still available to inspect. We call these *contingent* choices. We selected the 1718 decisions (64.3%) in which the RV and EV consistent choices were distinct, and participants made choices consistent with one of those strategies. We regressed a dummy variable coded as 0.5 when participants made the RV choice (and -0.5 for the EV choice) onto a null model with cluster robust standard errors to account for multiple observations per participant. Overall, 59.6% of these decisions were consistent with the RV strategy. Participants are significantly likelier to make RV consistent choices than EV consistent ones, $b = 0.096$, $t(1717) = 3.133$, $p=0.002$.

Figure 2.2.2. Navigation Heuristics Consistent with Participant First Choices (Panel A) and by Total Length of Search for Full Paths (Panel B) in Study 2.

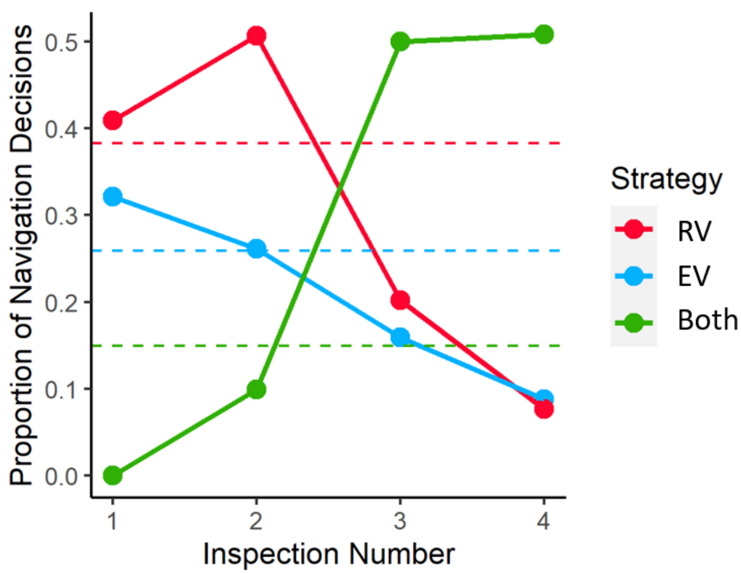


Note. In Panel A we classify searches based only on first choices. In Stimulus Sets 1 and 6 we could distinguish between the EV and Minimum Strategies. In Sets 3 through 5, we could not, and therefore use the label EV/Min. The RV heuristic cannot be distinguished from the Max heuristic in any sets. In Panel B we classify searches based on the entire search path. In Panel B we include all searches that could be classified as EV or RV including those that could also be classified as being consistent with other strategies. Importantly the EV and RV strategies never overlap.

A logistic regression on an indicator coded 1 for an RV choice with cluster robust standard errors also concludes that over half of observations are consistent with RV, $b = 0.389$, $z = 3.056$, p=0.002. We also plotted the frequency of all 2672 decisions separated by consistency with RV and/or EV strategies in Figure 2.2.3. Overall, 53.2% of inspections were consistent with an RV strategy while 40.8 % were consistent with an EV strategy, and (among these decisions) 14.9% were consistent with both. In Figure 2.2.3, we can see that RV and EV strategies are clearly distinguished in earlier inspections but not in later ones.

These numbers reflect the tendency for participants using the RV strategy to inspect more items than those using the EV heuristic. To correct for this tendency, we also average contingent choices by search and then calculate the average percentage of inspections consistent with different strategies *within each search.* By this measure, 49% of inspections in each search were consistent with the RV strategy while 41% were consistent with the EV strategy. Among these inspections, 10% were consistent with both the EV and RV strategy.

Figure 2.2.3. Proportion of Contingent Navigation Decisions Consistent with RV and EV Strategies by Inspection Number in Study 2



Note. The dashed lines represent the mean proportions consistent with the RV, EV, and both strategies. These three categories are mutually exclusive.

Stopping. Of the 2672 decisions[4] made by participants, overall, 59% of stop/continue decisions were to continue search. In 649 (24%) of these decisions, the RV strategy continues searching while the EV heuristic stops, and in these cases, participants continue 65% of the time[5]. When both the RV and EV strategies continue after 847 (32%) of inspections, participants continue 88% of the time. (There were no decisions where the EV heuristic continued while the RV strategy stopped.) Finally, when neither the RV nor EV strategies continued, participants continued 36% of the time. Overall. these results suggest that participants make the stop/continue decision according to the RV strategy about two-thirds of the time.

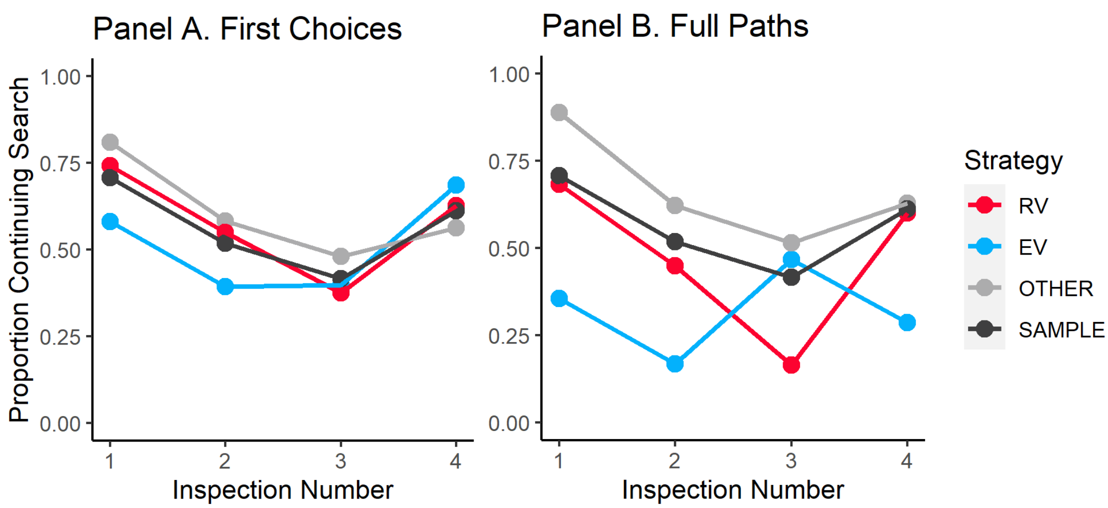
In Figure 2.2.4, we present the percentage of continuation decisions for each inspection separated by strategy. The tendency to continue after opening the fourth box is best interpreted as participants' motivation to see all the options, as only one outcome is still unknown.

Outcomes. To assess outcomes, we compare profits, costs, and maximum values with those obtained by the optimal model. In each case, we construct an index by dividing the participant's value by the optimal model's value. For example, for costs, we would divide the cost accrued by a participant in a search by the cost accrued by the optimal model. When this index is greater than 1 the participant has *higher* costs than the optimal model, when it is less than 1 the participant has *lower* costs. As noted above overall earnings were high, averaging 96% of the optimal model.

[4] This number excludes final choices, after the 5th inspection, where participants must end their search.

[5] We use an EV minus search cost heuristic; since all of our boxes have the same search costs, including search costs has no impact on navigation but could impact stopping.

Figure 2.2.4. Stop/Continue Decisions by Inspection and Strategy Classified based on First Choices (Panel A) and Full Paths in Study 2 (Panel B)



In Table 2.2.1 we regress these indices onto binary variables indicating the use of the RV, EV, or Other Strategies. We control for stimulus set and presentation order and use clustered standard errors to account for multiple observations per participant. We do this both for navigation classified with first choices and with full paths.

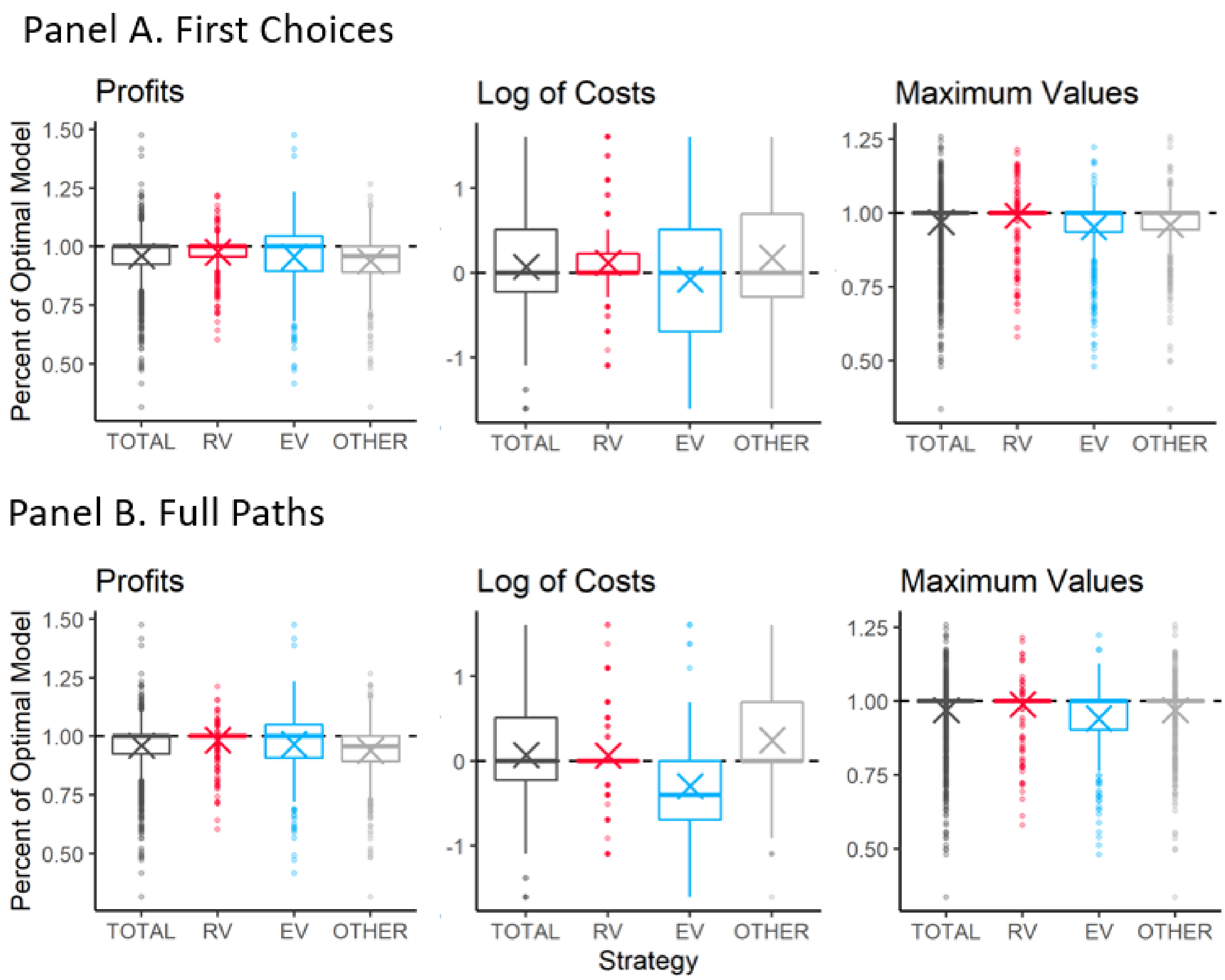
Table 2.2.1. Regression of Winnings and Outcomes as a Percent of Optimal Search Outcomes on Strategy in Study 2

	First Choices			Full Paths		
	Winnings (1)	Costs (2)	Max Values (3)	Winnings (4)	Costs (5)	Max Values (6)
RV strategy	0.021**	0.086	0.04***	0.017+	0.257***	0.049**
	(0.007)	(0.079)	(0.008)	(0.009)	(0.061)	(0.009)
Other strategy	-0.018†	0.315***	0.008	-0.025**	0.648***	0.029**
	(0.009)	(0.084)	(0.010)	(0.010)	(0.075)	(0.010)
Obs	1200					
R2	0.033	0.035	0.042	0.042	0.100	0.039
BIC	-1790.09	3243.83	-2102.78	-1801.27	3160.51	-2098.94

Note. *** p < 0.001; ** p<0.01; * p<0.05; †<0.10

In aggregate, participants tend to profit less than is optimal, both because they tend to search more than the optimal model (and therefore accrue greater search costs) and end their searches with less valuable items (Figure 2.2.5). When we separate outcomes by strategy, searches using the RV strategy earn more compared to other strategies but earn less than the optimal model. They tend to stop with options nearly as valuable as those found by the optimal model but accrue more in search costs. In contrast, those following the EV heuristic perform better than all other participants who are not following the RV strategy. Compared with searches using the RV strategy, those using the EV heuristic end up with less valuable items but accrue less in search costs.

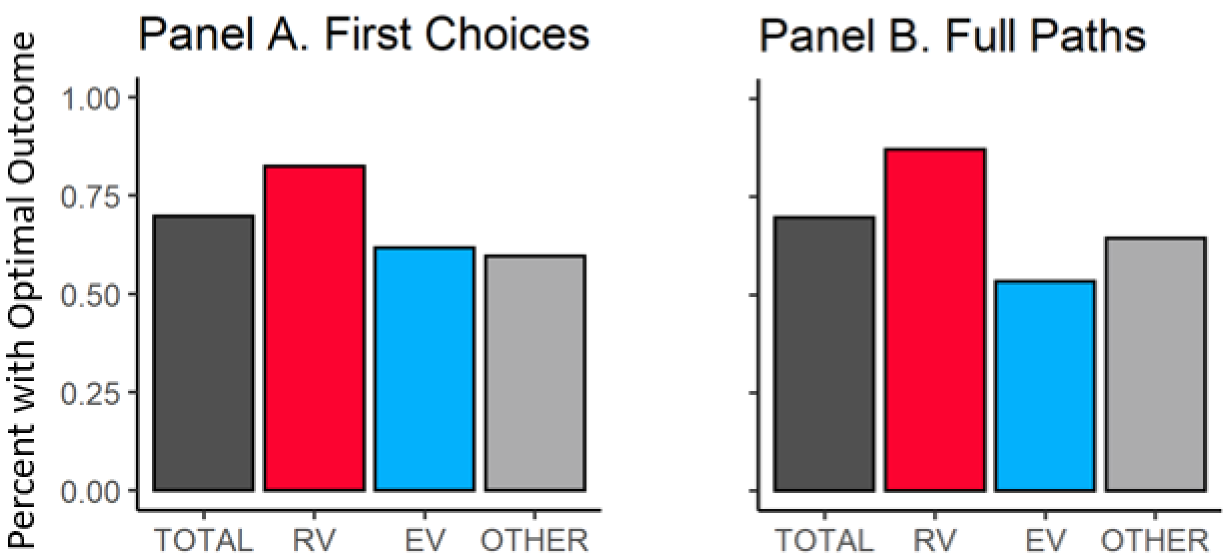
Figure 2.2.5. Participant Percent of Optimal Model Profits, Costs, and Maximum Values for Strategies Classified by First Choices (Panel A) and Full Paths (Panel B) in Study 2



Note. Xs indicate means. The performance of the optimal model is at 1 for profits and maximum values and 0 for the log of costs.

Overall, participants end their searches with the same item as the optimal model 70% of the time (Figure 2.2.6). When we classify by full paths, searches using the RV strategy are more likely to choose the same item as the optimal model (87%) compared with searches using the EV heuristic (53%). Searches that do not follow the RV or EV strategies end up with the same item as the optimal model 64% of the time. When strategies are classified by first choices, we find similar results (RV=82%, EV=62%, OTHER=59%).

Figure 2.2.6. Percent of Searches with the Same Outcome As The Optimal Model For Strategies Classified by First Choices (Panel A) and Full Paths (Panel B) in Study 2



2.2.3. Discussion

In Study 2, we examined navigation and stopping decisions when participants had five locations from which to choose. Looking at individual searches, about half of the navigation decisions and two-thirds of the stop/continue decisions were consistent with the RV strategy. The EV strategy was a close second, consistent with approximately 40% of the navigation decisions and one-third of the stopping decisions. These patterns are consistent with the implications from Study 1's findings based on stated reservation values.

In terms of outcomes, both strategies achieved at high levels of profits, 95% of the optimal model's performance. Participants following the RV navigation strategy failed to profit as much as the fully optimal model because, even though they were likely to conclude a search with the same option as the optimal model, they tended to over-search and accrue excessive search costs. Those following the EV strategy searched less and were less likely to end their search with the same option as the optimal model. Although participants' earnings did not differ much based on the strategy they used, participants using different strategies often ended up with different products. For marketers, this suggests that consumers using different search strategies may end up purchasing different products or purchasing products at different locations even if navigation strategies have only a modest impact on their welfare.

2.3. Study 3: Search with Navigation with Graphical Stimuli

Study 3 replicates the design and methods from Study 2, with two changes: the stimuli were presented in a graphical, rather than numerical format, and the participant is instructed to play the role of a consumer searching for an automobile to purchase from five brands.

2.3.1. Method

Participants. We recruited 208 participants from the Prolific online website and paid a \$2.50 fixed participation fee for a session lasting 16.7 minutes on average; participants earned an average of \$0.54 additional in performance-contingent bonuses. We discarded 7 participants who did not complete one or more parts of the task, and report results from the remaining 201 participants (42.8% Male, 57.2% Female; Average Age 39 years).

Design. Each participant completed 12 searches, each composed of five options, replicating the plan from Study 2.

Procedure. We told participants that they were shopping for a new car and that in each search, they were considering buying one of five brands of cars. In each search, participants saw a graphic depicting the user ratings by 30 people of each brand of car. They were told that to learn how much *they* would like the car, they would have to take the car for a test drive (inspection), and they could inspect up to five cars in each search (Figure 2.3.1). When they tested a car, the graphic for a brand would be replaced by their personal numerical rating for that car (on a scale from 0 to 200). This numerical rating was randomly drawn for each test drive from the same uniform distribution represented in the graphic. Participants were charged a five-point search cost that represented the effort, time, and money required to test drive a car.

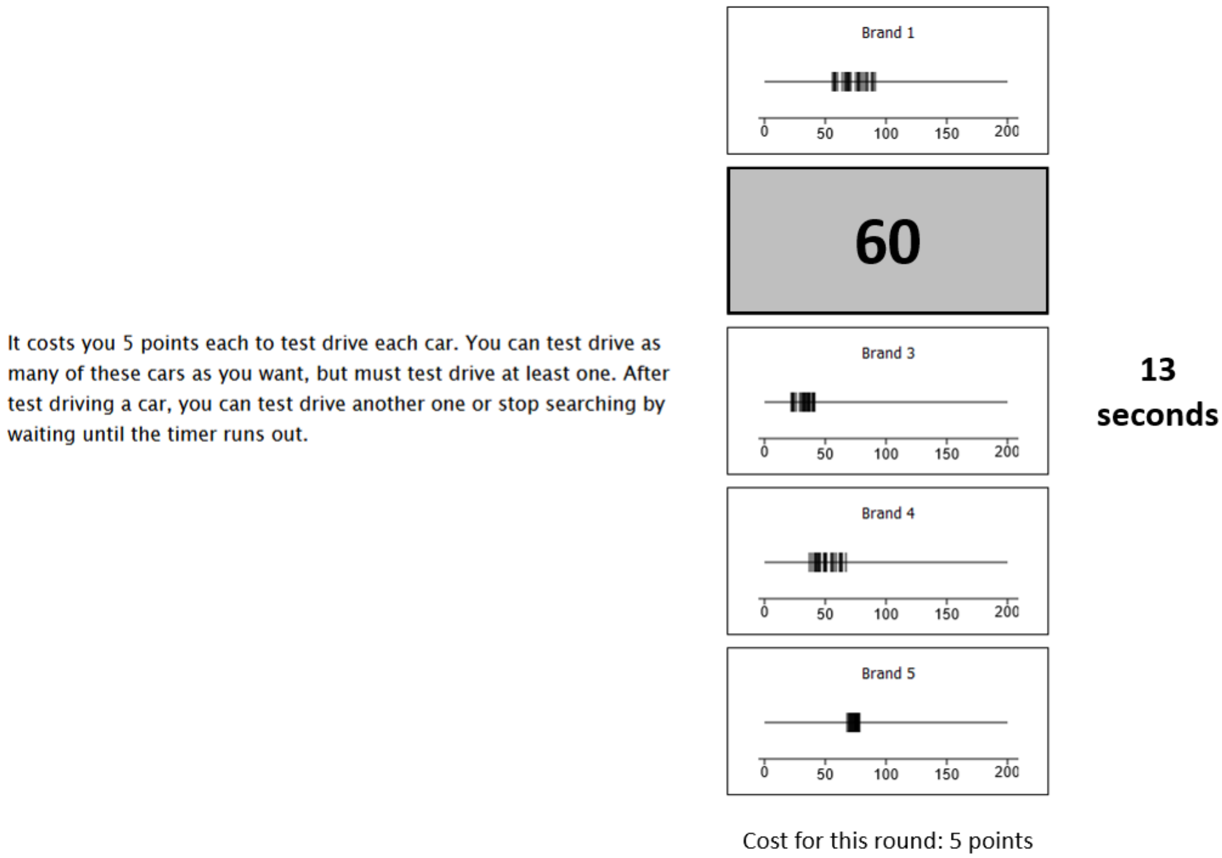
As in Study 2, participants had 20 seconds to complete each of the 12 searches. For each round, a participant's point total was the maximum rating that they found in their test drives during a search minus their total accrued search costs. Performance bonuses were paid as in Study 2. To make sure that participants understood our task, we had them answer seven comprehension questions, and complete two practice trials.

2.3.2. Results and Analysis

Descriptive Statistics. Overall, participants inspected 2.13 options on average (SD=1.03). In total, only 14% of searches perfectly matched the rational RV model's navigation and stopping prescriptions versus 23% in Study 2. The median participant matched the rational strategy in 1 search out of 12; 33% of participants never matched the rational model.

Participants' mean per-search value obtained was 99.05 points (SD = 22.68). As in Study 2, we generated values for every option and so could simulate the performance of each strategy's earnings, and average participant earnings were 95% of the optimal model's achievement.

Figure 2.3.1. Stimulus Display in Study 3



Note. Each box depicts 30 user ratings for each brand on a scale from 0 to 200. In this search, the participant has taken one test drive revealing a value of 60 points.

Navigation. Participants' first choices (Figure 2.3.2, Panel A.) were consistent with the RV strategy 33% of the time and consistent with the EV heuristic 38.6% of the time. When we examine participants' entire paths, we find a similar pattern. Overall paths were equally likely to be consistent with the RV and EV strategies (26.6% RV versus. 26.1% EV, or 52.7% together). Searches with a single inspection are more likely to be consistent with the EV heuristic (52.1%) than the RV strategy (28.4%). In longer searches, however, the EV heuristic becomes less common than the RV strategy (Figure 2.3.2, Panel B).

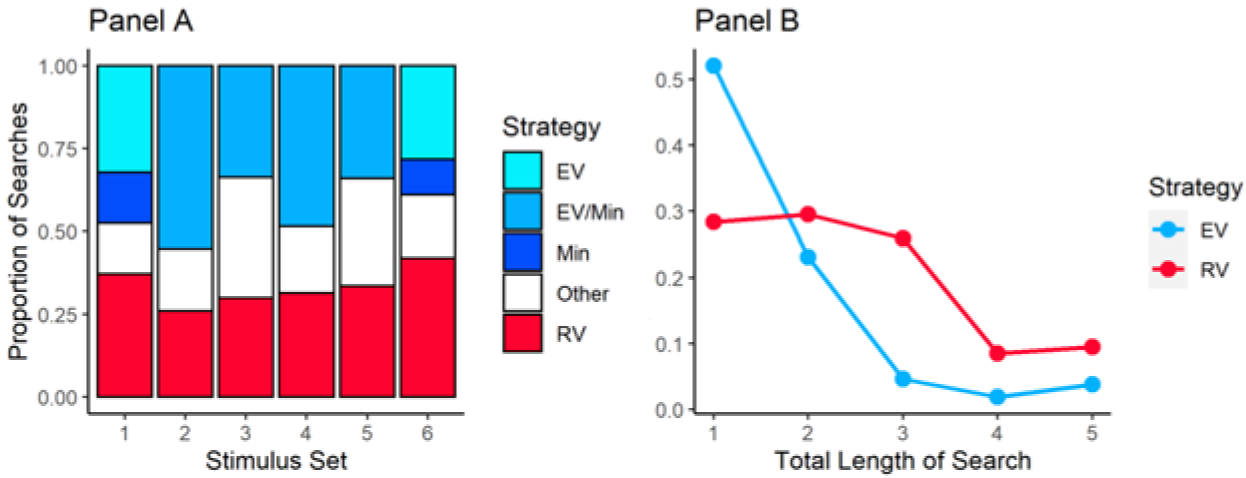
To relate search length to navigation strategies further, we regress the total length of search on dummies for the RV, EV[6], and other strategies, setting the RV dummy as the reference level. We fitted separate models for full paths and first choices, including fixed effects for the set and set presentation order and cluster standard errors at the participant level to account for multiple observation per participant. Overall, full paths consistent with the EV heuristic are 0.50 inspections shorter than those following the RV strategy, $t(2381)=8.283$, $p < 0.001$. Those using strategies other than EV and RV have searches that are 0.64 inspections longer than those using the RV strategy, $t(2381)=10.019$, $p < 0.001$. When we classify strategies based on first choices, we again find that participants who select the EV first choice have searches that are 0.30 inspections shorter than those using the RV strategy, $t(2381)=4.250$, $p<0.001$ while those using other strategies have searches that involve 0.26 more inspections, $t(2381)=1.654$, $p=0.098$. These results are consistent with our findings from Study 2.

We can also consider navigation decisions one at a time (contingent navigation decisions) and ask whether they adhere to the EV or RV strategies given the remaining options, regardless of past choices. For the 5010 navigation decisions, overall, 44.1% of contingent decisions are consistent with the RV strategies whereas 45.2% are consistent with the EV strategy. Among these choices, 11.3% of decisions were consistent with both the RV and EV strategies (see Figure 2.3.3). When we break down these contingent navigation decisions by inspection number, we see that for the first and fourth test drive, participants are more likely to navigate based on the EV heuristic while for the second and third test drive, they are likelier to follow the RV strategy.

[6] RV in this dummy includes both paths that are purely consistent with the RV navigation strategy and those consistent with both the RV and Max strategies. Similarly, the EV dummy includes both paths that are purely consistent with EV as well as those consistent with both EV and Min Strategies.

(The number of decisions consistent with both strategies increases in later rounds as the number of remaining choices declines.)

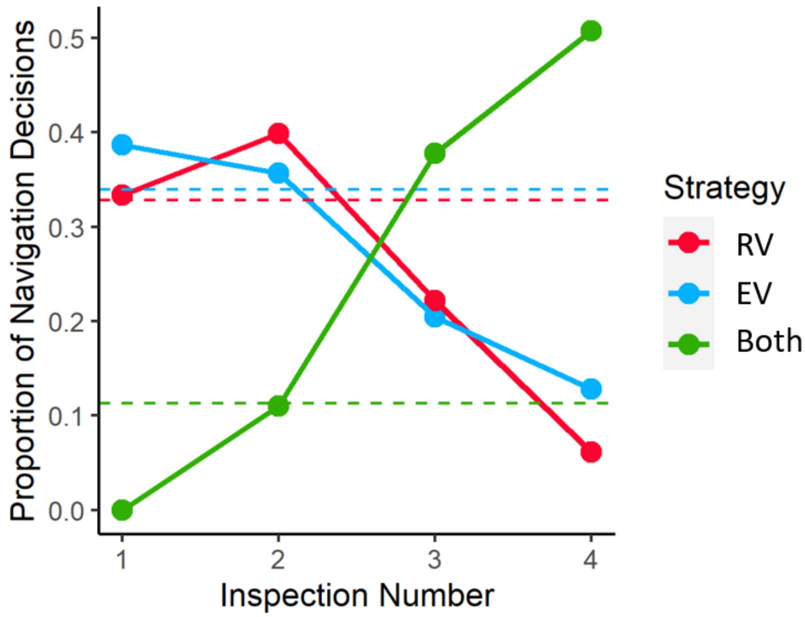
Figure 2.3.2. Navigation Heuristics Consistent with Participant First Choices (Panel A) and by Total Length of Search for Full Paths (Panel B) in Study 3



Note. In Panel A we classify searches based only on first choices. In Stimulus Sets 1 and 6 we could distinguish between the EV and Minimum Strategies. In Sets 3 through 5, we could not, and therefore use the label EV/Min. The RV heuristic cannot be distinguished from the Max heuristic in any sets. In Panel B we classify searches based on the entire search path. In Panel B we include all searches that could be classified as EV or RV including those that could also be classified as being consistent with other strategies. Importantly the EV and RV strategies never overlap.

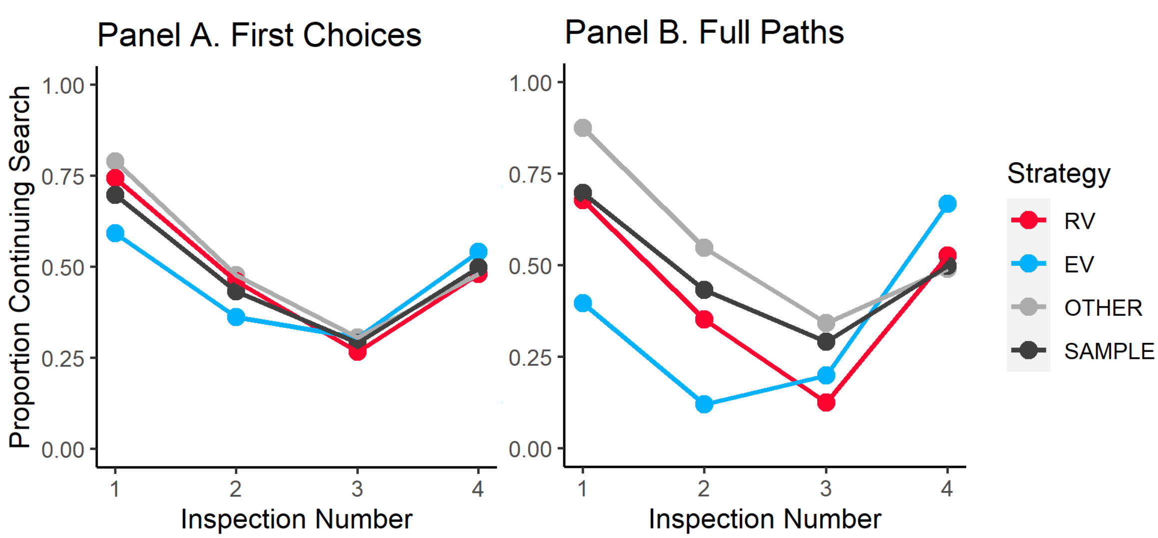
We considered the 3342 decisions (66.7%) in which the RV and EV consistent choices were distinct, and participants made choices consistent with one of those strategies. We regressed a dummy variable coded as 0.5 when participants made the RV choice (and -0.5 for the EV choice) onto a null model with cluster robust standard errors to account for multiple observations per participant. Unlike in Study 2, where participants were more likely to make RV consistent choices, participants in Study 3 are equally likely to make EV and RV consistent choices, b = -0.009, $t(3341) = 3.133$, p=0.658. A logistic regression on an indicator coded 1 for an RV choice with cluster robust standard errors also indicates that participants are equally likely to make EV and RV consistent navigation choices, b = -0.034, z=0.443, p=0.658.

Figure 2.3.3. Proportion of Contingent Navigation Decisions Consistent with RV and EV Strategies by Inspection Number in Study 3



Note. The dashed lines represent the mean proportion consistent with the RV, EV, and both strategies.

Figure 2.3.4. Continuation Decisions by Round and Strategy based on First Choices (Panel A) and Full Paths (Panel B) in Study 3



Stopping. As in Study 2, we examined stopping overall as well as by navigation strategies. Of the 5010 stop/continue decisions[7] made by participants, 54.2% continued search, slightly less than in Study 2. In 1410 (28.1%) of these decisions, the RV strategy continues searching while the EV heuristic stops, and in these cases, participants continue 56.0% of the time. When both the RV and EV strategies continue after 1408 (28.1%) of inspections, participants continue 80.5% of the time. When neither the RV nor EV strategies continued, participants continued 36.2% of the time. Overall, these results suggest that participants follow the RV strategy slightly more than half the time. We present these data in Figure 2.3.4.

Outcomes. Finally, we compare the outcomes of using different strategies. As noted above overall earnings were high, averaging 95% of the optimal model. We repeat all analyses that we conducted for Study 2: participant profits, search costs, and maximum compared to the optimal model and how often participants' final choices match those of the optimal model. Overall, we find similar results to those for Study 2. We present regression results for profits, costs, and maximum values in Table 2.3.1 and present these results graphically in Figure 2.3.5. We present differences in chosen items between participants and the optimal model in Figure 2.3.6.

We again find that in aggregate, participants tend to have lower profits than optimal, both because they tend to search more than the optimal model (and accrue greater search costs) and end their searches with slightly less valuable items. When we separate outcomes by strategy use, we again find that searches using the RV strategy end up with items nearly as valuable as the optimal model does but with higher search costs. In contrast, searches using the EV heuristic end

[7] This number excludes final choices, after the 5th inspection, where participants must end their search.

up with lower value items while accruing lower search costs. Unlike in Study 2 where the RV strategy was associated with higher profits than the EV strategy, in Study 3 there is no difference in profits between strategies.

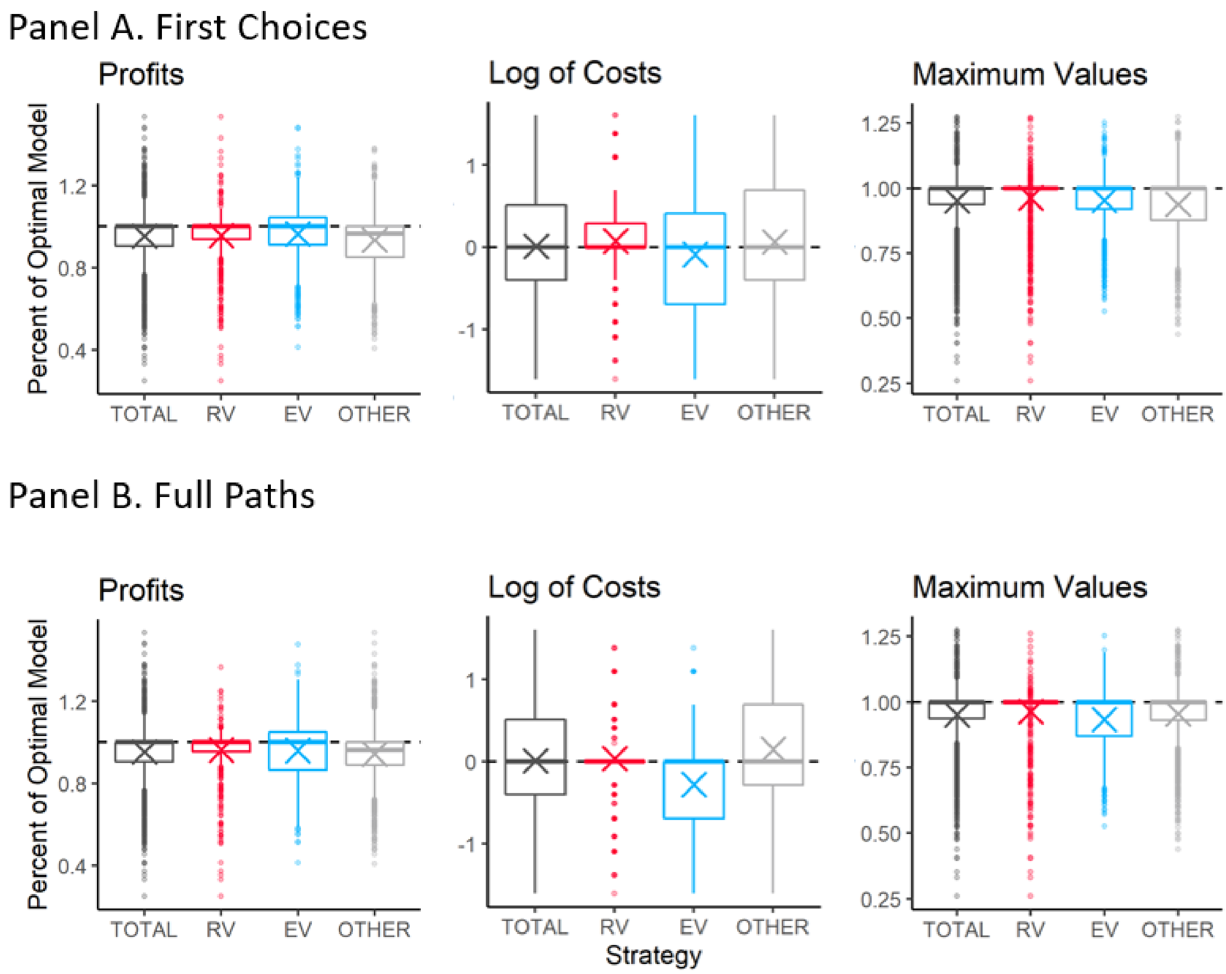
Overall, participants end their search with the same item as the optimal model 61.8% of the time, slightly less than in Study 2. Importantly, however, when we classify by full paths, searches following the RV strategy are more likely to choose the same item as the optimal model (79.8%) compared with searches using the EV heuristic (48.3%). Searches that don't use the RV or EV strategy end up with the same item 59.2% of the time. When strategies are classified by first choices, we find similar results (RV=75.2%, EV=58.3%, OTHER=50.6%).

Table 2.3.1. Regression of Winnings and Outcomes as a Percent of Optimal Search Outcomes on Strategy in Study 3

	First Choices			Full Paths		
	Winnings (1)	Costs (2)	Max Values (3)	Winnings (4)	Costs (5)	Max Values (6)
RV strategy	-0.007	0.103*	0.011†	0.004	0.268***	0.032***
	(0.007)	(0.049)	(0.007)	(0.009)	(0.046)	(0.008)
Other strategy	-0.030***	0.187***	-0.013†	-0.016*	0.498***	0.023***
	(0.008)	(0.051)	(0.007)	(0.007)	(0.041)	(0.007)
Obs	2400					
R2	0.014	0.017	0.013	0.010	0.073	0.017
BIC	-2723.82	5856.98	-3355.48	-2715.12	5716.25	-3364.86

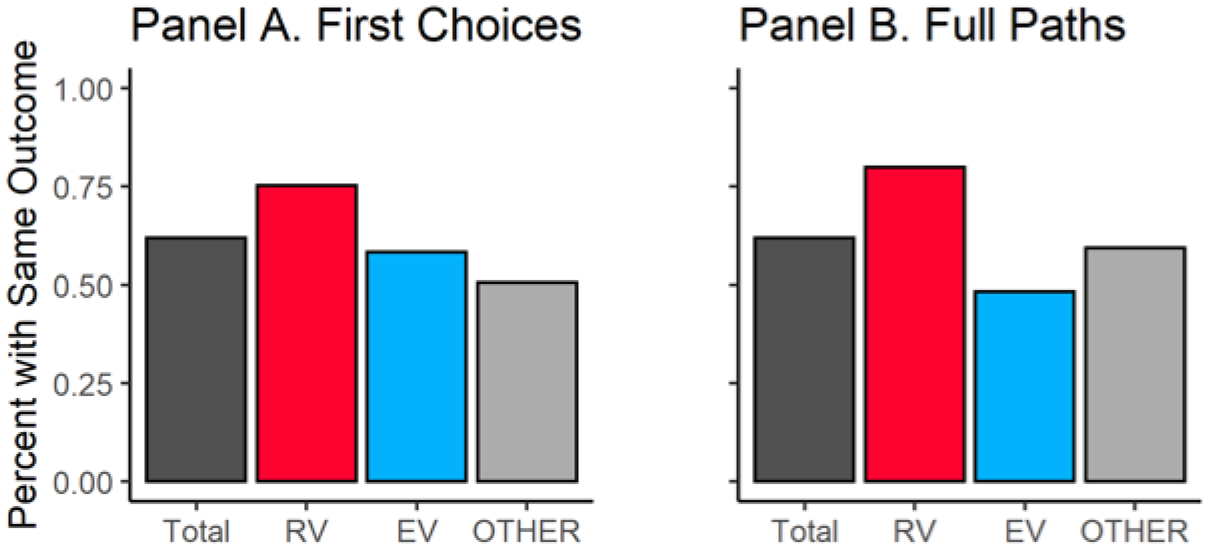
Note. *** p < 0.001; ** p<0.01; * p<0.05; †<0.10

Figure 2.3.5. Participant Percent of Optimal Model Profits, Costs, and Maximum Values for Strategies Classified by First Choices (Panel A) and Full Paths (Panel B) in Study 3



Note. Xs indicate means. The performance of the optimal model is at 1 for profits and maximum values and 0 for the log of costs.

Figure 2.3.6. Percent of Searches With Same Outcome As The Optimal Model For Strategies Classified by First Choices (Panel A) and Full Paths (Panel B) in Study 3



2.3.3. ***Discussion***

Study 3 replicated many of the results from Study 2, with a consumer search cover story and with a graphical (rather than numerical) format for the search task information.

As in Study 2 only a few participants were perfectly consistent across all searches with either of the dominant RV and EV strategies. But, looking at individual searches, about one-third of the navigation decisions and one-half of the stop/continue decisions were consistent with the RV strategy. Participants were slightly more likely to follow the EV heuristic, consistent with approximately 40% of the navigation decisions and one-half of the stop/continue decisions. Compared with Study 2, participants were more likely to navigate based on the EV strategy.

In terms of outcomes, both strategies achieved at high levels, 95% of the optimal model's performance. Again, the RV strategy was more likely to conclude with the best option but tended to over-search and accrue excessive search costs. The EV strategy searched less and was less likely to find the best option. These patterns replicate Study 2's findings.

2.4. General Discussion

In a world where thousands of retailers and millions of products vie for a consumer's attention, it is important to understand how consumers search. To date, most scientific attention has been focused on when consumers stop searching, with much less research on selection or navigation. The present research shifts the focus from stopping to navigation. Our conclusion is that an extension of the basic economic model for stopping based on estimates of expected gain (the Weitzman Model) provides a good account for about one-half of participants' navigation decisions. A sensible, but sub-optimal strategy based on expected value is a close second; and a small sub-set of searchers follow anomalous strategies. We see our results as conceptual replications of the findings from the two most similar prior experimental studies (Gabaix, et al., 2006; Urbany, 1986; we interpret our EV heuristic as similar to Gabaix, et al.'s "Direct Cognition Algorithm").

The most important implications of our results support the general conclusion that participants performing an analogue to a consumer search task exhibit adaptive search strategies that are approximately rational. We also introduce a new experimental paradigm for research on goal-directed navigation in a controlled consumer search task.

3. Chapter 3. The Value-Spatial Navigation Tradeoff

In Chapter 2, we looked at how people search in simple environments. In general, people tended to use value as a cue, although not always optimally. But as environments become more complex, we expect that people will increasingly rely on heuristics to navigate as they search. And, as using value-based strategies for searching require increasing amounts of working memory and becomes more difficult in general, we expect that consumers will ease the burden on their working memory by relying more on spatial navigation heuristics. In Chapter 2, we manipulate the environment by increasing the number of choices and altering the spatial layout of choices on the screen. Increasing the number of choices increases working memory load for using value-based heuristics. And changing the layout of locations on the screen can make value-based navigation more difficult.

3.1. Study 1: The Impact of the Number of Boxes

In this study, we examine the impact that the number of items presented to participants has on their navigation decisions. We hypothesize that as the number of potential locations to search increases, participants will increasingly navigate using spatial strategies rather than value-based strategies. The purpose of this study was to understand how participants' strategies differ based on the size of the set of products that they can navigate to.

3.1.1. Method

Participants. We recruited 150 adults (55.3% Male, 44.7% Female; Average Age 40 years) to complete an experimental search task in an 18-minute session. Participants were paid a $2.70 fixed participation fee and earned an additional average of $1.50 in bonuses.

Design. The study design was a 6 (within; unique search sets of five brands) X 2 (within; Repetitions) X 3 (between; number of items in the search set– 6, 9, and 12) design. After completing two practice searches, each participant completed 12 searches. Across the different between-subjects conditions, we varied the number of items by starting with the 6-item stimulus set and adding unattractive items to create the 9 and 12 item sets. This is analogous to adding additional low-value products to a set of products. The order of presentation of the sets of items and the items themselves were randomized. We present our stimulus sets in Appendix Table 6.3.1.

Procedure. We told participants that they were shopping for a new car and that in each search, they were considering buying one of several brands of cars. In each search, participants saw a graphics depicting the user ratings by 30 people of each brand of car. The boxes were vertically centered on the screen. Participants were told that to learn how much *they* would like the car, they would have to take the car for a test drive (inspection), and they could inspect as many brands as were available. When they tested a car, the graphic for a brand would be replaced by their personal numerical rating for that car (on a scale from 0 to 200). This numerical rating was randomly drawn for each test drive from the same uniform distribution represented in the graphic. Participants were charged a one-point search cost that represented the effort, time, and money required to test drive a car.

Participants had 25 seconds to complete each of the 12 searches. For each search, a participant's point total was the maximum rating that they found in their test drives during a search minus their total accrued search costs. Performance bonuses were 1% of the highest value found in one randomly drawn search. To make sure that participants understood our task, we had them answer seven comprehension questions, and complete two practice trials.

3.1.2. Analysis

Number of Test Drives. Overall, participants test drove 2.69 brands (SD = 1.81[1]). However, participants test drove more cars on average when more cars were available (6 Brands: M = 2.40 SD =1.25; 9 Brands: M = 2.63, SD = 1.49; 12 Brands: M = 3.05 SD =2.41).

We run a regression of the number of test drives on the number of brands available controlling for stimulus and whether it was the first or second time a stimulus set was displayed. For this and all future regressions, we use clustered standard errors to account for multiple observations per participant. Making an additional brand available leads to an additional 0.11 test drives, $t(1792) = 2.02$, $p = 0.0434$. However, when we remove two participants who test drove all brands in all searches, the increase is no longer statistically significant, $b = 0.0461$, $t(1768) = 1.43$, $p = 0.150$.

Even when we account for extreme values, however, we see that participants search less in the last 6 searches than the first 6 searches, $b = -0.18$, $t(1768) = 2.47$, $p = 0.013$. When we regress the number of test drives onto the number of brands available for the first and last 6 searches separately, participants search longer when they are offered more brands in the last 6 searches, $b = 0.075$, $t(881) = 1.980$, $p = 0.048$ (M_6=2.27 M_9=2.46 M_{12}=2.72), but not in the first 6 searches, $b = 0.02$, $t(881) = 0.904$, $p = 0.366$ (M_6=2.53 M_9=2.80 M_{12}=2.64).

As a final analysis, we run a regression that includes an interaction between the number of available brands and whether a search was in the first 6 or second 6 searches. This analysis allows us to see whether the number of available brands impacts search lengths differently between these two groups of searches. There is a marginally significant interaction, $b = -0.06$,

[1] Standard Deviations do not account for multiple observations per participant. 95% confidence intervals on graphs also do not account for repeated observations.

$t(1767) = 1.91$, $p = 0.057$, indicating that there may be different impacts of search set size on early versus late searches.

Profits. For each search, we calculate profits by subtracting accrued search costs from the value of the best car found in that round. More available brands does not lead to significantly lower profits, $b = -0.22$, $t(1768) = -1.29$, $p = 0.197$ (M_6=148.04, M_9=147.62 M_{12}=146.73), although there is a trend in that direction. However, participants do profit more in the last 6 searches compared with the first 6 searches, $b = 1.90$, $t(1768) = 3.09$, $p = 0.002$, suggesting that participants learn to be more optimal with practice.

Value of First Choices. Overall, we find no differences in how often participants make the optimal first choice when search set size varies. In all three conditions, participants make the optimal first choice approximately 25% of the time. Overall, participants make first choices consistent with the EV strategy 49.8% of the time. In the 6 brand condition, participants test drive the highest EV brand 59.2% of the time. They do so 49.0% percent of the time in the 9 brand condition and 41.3% of the time in the 12 brand condition. A regression controlling for stimulus test set and whether a search was among the first or last 6 searches shows a 2.9 percentage point decrease in the chance of visiting the highest EV brand first for each brand added, $t(1792) = 3.24$, $p = 0.001$.

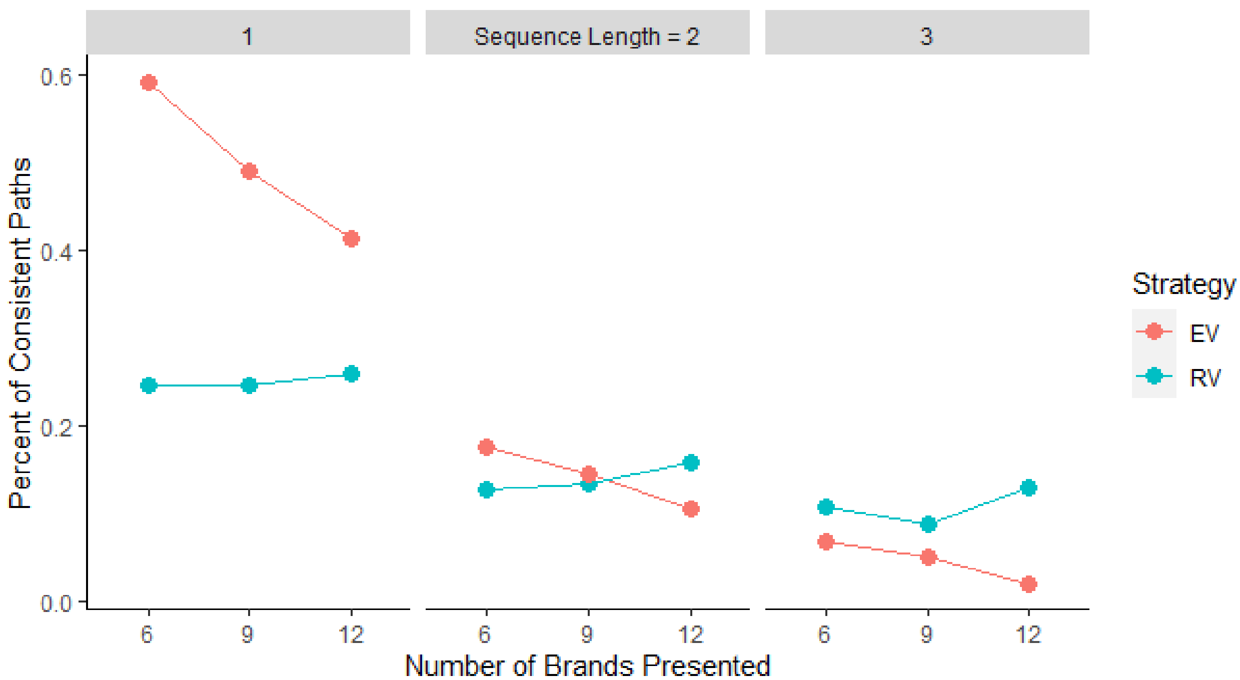
One explanation for these results is that the optimal strategy was relatively easy to follow as the number of brands increase while the expected value strategy was not easier to follow in larger search sets. In our stimulus set, the optimal first choice always had the highest maximum. Anyone following this strategy could take advantage of the visual salience of the maximum in our graphics. On the other hand, the brand with the highest expected value was not as obvious, particularly as the number of brands increased.

Value-Based Sequences. Next, we examine how often the first 2 and first 3 navigation decisions conform to the RV and EV strategies. Overall, the first two inspections of searches that are 2 or more inspections long are fully consistent (i.e., all inspections are consistent with the heuristic) with the RV strategy 14.0% of the time (compared with 25% of first choices). For the EV strategy 14.1% of searches are consistent (compared with 49.8% of first choices). When we look at adherence to strategies for the first 3 inspections of searches that are 3 or more inspections long, 10.8% of searches adhere to an RV strategy while 4.4% adhere to an EV strategy. The precipitous decline in the EV strategy relative to the RV strategy may, once again, have to do with the salience of cues. Even for the second or third inspection, the maximum (which corresponds to the RV strategy) is easy to detect thanks to the visual display. In contrast, the EV is relatively difficult to detect, as it must be inferred based on the distributions presented in the set.

When we separate adherence to strategy by the number of brands presented, we see trends like those we saw when looking only at first choices. We present these data graphically in Figure 3.1.1. The percentage of searches that follow the RV strategy for the first two searches remains relatively constant as the number of brands presented increases (rising from 12.6% for 6 brands to 15.7% for 12 brands. See middle panel of Figure 3.1.1). A linear probability model regressing whether a search started with 2 RV strategy consistent inspections, controlling for stimulus shows no difference between conditions, $t(1353) = 0.811$, p=0.418. Searches following the EV strategy decline from 17.6% for 6 brands to 10.6% for 12 brands. A model like the one above for EV strategy consistent sequences shows a significant decline as the number of brands increases, $t(1353) = -2.469$, $p = 0.014$.

We see similar trends, albeit less pronounced, when we look at the first 3 choices (Presented in the right panel of Figure 3.1.1). There is, again, no difference between conditions in the presence of 3 inspection sequences that follow the RV strategy, $t(868) = 0.534$, $p = 0.593$. But as the number of brands presented increases, the number of 3 inspection sequences following the EV strategy declines, $t(868) = -2.873$, $p = 0.004$.

Figure 3.1.1. Percentage of Sequences of 1, 2, and 3 Strategy Consistent Inspections by Number of Brands Presented in Study 1



Note. Percentages are of the number of participants who have search lengths at least as long as the sequence. For example, for sequences of 2, percentages are of all searches that last two or more inspections that conform to a strategy in their first 2 inspections.

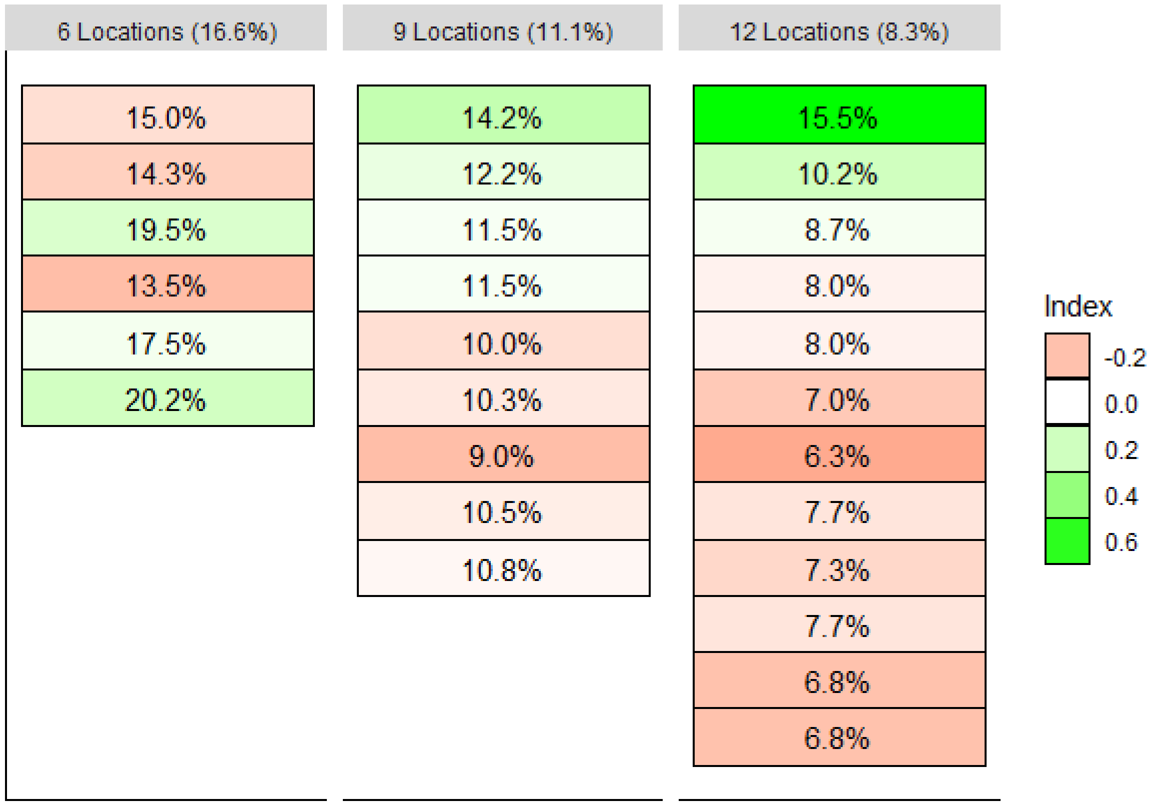
Spatial Search Patterns. Next, we look at whether participants are searching from top to bottom, as opposed to using a purely value-based strategy. As a rough analysis, we look at the correlation between the vertical position of a choice and the inspection number within the search. If participants are searching from top to bottom, we expect that low inspection numbers (early in a search) will be near the top of the screen (a vertical position close to 1) whereas a high inspection number will be closer to the bottom of the screen (with a higher position number) implying a positive correlation between inspection number and vertical position of a choice.

For each participant, we calculate a spearman correlation coefficient between the inspection number and the vertical position[2]. Overall, the average correlation when we presented 12 brands was r=0.132 while it was r=0.063 for 9 brands and r=-0.036 for 6 brands. When we regress the correlation coefficients onto a continuous variable with the number of brands presented in each condition, there is a positive relationship between the number of brands presented and the correlation coefficient, $t(146)=3.157$, $p=0.002$. These results suggest that, in general, as the number of boxes presented increases, participants tend to search from top to bottom more often.

Unsurprisingly, given the increasing number of boxes, the average position of first choices is higher (i.e., further down the screen) when there are more brands available when we control for learning and stimulus set, $b = 0.361$, $t(1792) = 8.550$, $p < 0.001$. Next, we re-center each choice as a distance from the center. For example, in the 6-brand condition, we calculate the vertical position of each choice from 3.5. After this re-centering, a negative number means that a choice is below the middle while a positive number means that it is above the middle. As the number of brands increases, the average first position is increasingly above the center of the screen, $b = 0.139$, $t(1792) = 3.289$, $p = 0.001$. While in the 6 box condition, average first choices were 0.14 boxes below the midpoint of the screen, in the 9 box condition it was 0.25 boxes above the midpoint and 0.69 boxes above in the 12 box condition. This result suggests that participants focus more on brands closer to the top when the number of available brands increases.

[2] Two participants were omitted because they only inspected 1 brand each in all of their searches

Figure 3.1.2. First Inspections at Each Vertical Position for Each Condition in Study 1



Note. To account for the different number of boxes in each condition, we created an index by dividing the percentage of choices for each vertical location by the percentage we would expect by chance for each position in each condition (noted in brackets next to condition names). We then took the log of this number such that numbers greater than 0 indicate more choices than expected in a condition (depicted in green) and those below zero indicate fewer (depicted in red).

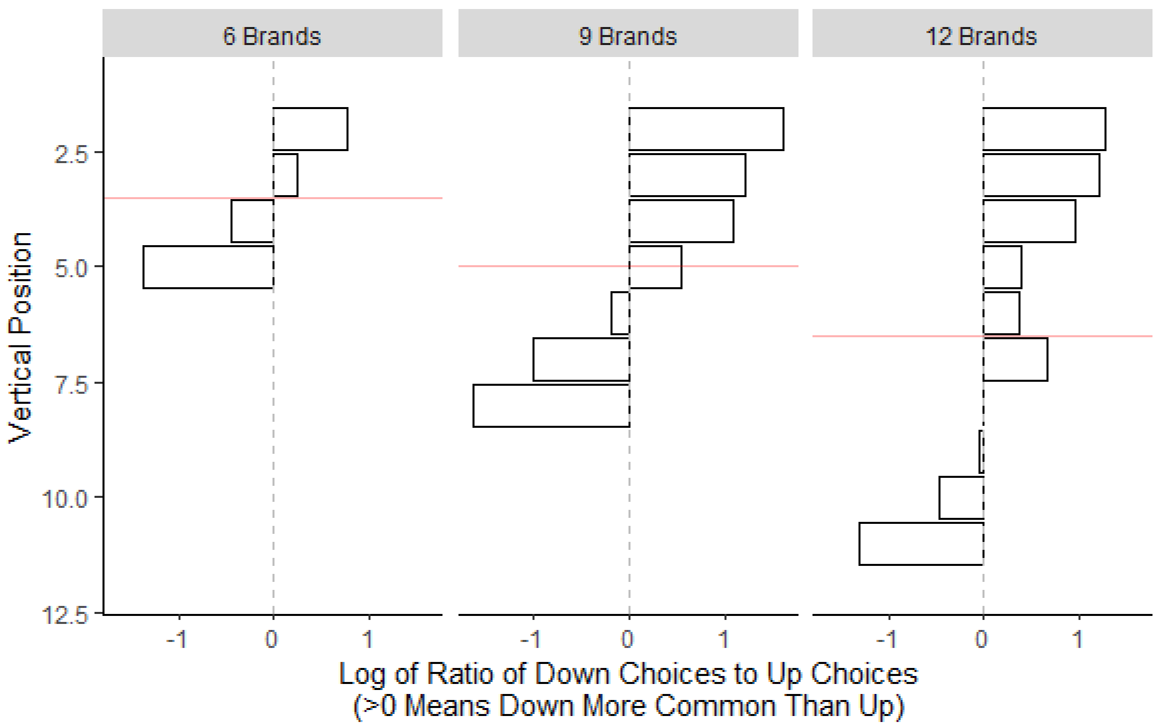
We also look at the vertical position of first choices made by participants. In Figure 3.1.2 we present the percentage of searches that started at each vertical position by condition. A visual inspection suggests that in the 6-box condition, participants tend to start in a variety of locations while in the 9- and 12-box conditions, they tend to start at the top. This trend is particularly evident in the 12-box condition where nearly twice as many searches start at the top of the screen than by random chance. We aggregate the percentage of first choices at each location for each participant and compare to the random model using t-tests and find that participants are significantly likelier to select the top location a in the 12-box condition and marginally likelier to do so in the 9-box condition. And they are significantly less likely to select the bottom location in the 6-box condition. Note that these tests are not independent of one another and therefore significance levels are only approximate. We present these results in full in Appendix Table 6.3.2.

Direction of Subsequent Choices. Our previous analyses have focused on first choices. Here, we look at whether participants are more likely to move up or down as they search. Participants navigating based on a value-based principal like RV or EV should not systematically choose a brand that is either above or below a previously inspected brand after controlling for the number of brands above or below the previous choice. On the other hand, if participants are using the screen layout as a guide to navigate, there ought to be a bias towards selecting an item that is above or below. It's worth noting that since we do not control for the length of search in our experimental design, these results should be approached with caution.

We run a linear probability model of whether a subsequent choice was above or below the current choice (1 = below) on the number of brands presented controlling for the vertical position of the current choice, whether a search was in the first or second half of the experiment, and the number of choices made in the current search. We control for the vertical position because we expect that the number of brands above and below the current item will impact choices. If there are two boxes above and nine boxes below the current choice, there are likely to be more attractive choices below the current choice, regardless of strategy. Finally, we removed observations that did not have a subsequent inspection. We also removed choices where a participant had already inspected all of the locations above or below the current one because participants had not choice on the direction in which to move in those cases. We present the log of the ratio of up versus down choices in Figure 3.1.3. In Appendix Table 6.3.1 we present the same data only for second searches and observe a similar, if attenuated, pattern. In Table 3.1.1 we present the results of regression analyses. In Appendix Table 6.3.3 we repeat this analysis with a logistic regression. Results are similar to those presented in Table 3.1.1.

Participants are more likely to choose an item below (vs above) the current one when there are more brands available. Controlling for other variables, each increase in the number of boxes increases the probability of choosing a box below it next by 5.7 percentage points, $t(2305) = 9.175$, $p < 0.001$.

Figure 3.1.3. Log of Ratio of Up vs Down Choices by Condition and Vertical Position in Study 1



Note. The red line indicates the middle of the screen. The dashed grey line indicates an equal probability of moving up or down. The top and bottom positions were excluded from the plot as participants could only make subsequent choices in one direction. All datapoints where participants had only an up or a down choice were also excluded.

As a robustness check, we examine how sensitive participants in each of our conditions are to the percentage of brands remaining below the most recently inspected brand. We do this because, in a model that picks randomly, the probability of moving down is determined by the percentage of uninspected boxes remaining below the current one. To our previous regression, we add controls for the percent of uninspected locations below the current one. We find that controlling for these factors, the number of boxes does not make a difference, b=-0.020, t(1904) = 2.010, p=0.237. Based on this analysis, there is limited evidence that participants have a bias to move downward. We present these models in Appendix Table 6.3.4.

Table 3.1.1. Regression of Down (vs Up) Moves onto Number of Boxes Presented

	Basic Model	With % Remaining Below
Number of Boxes	0.055***	-0.020
	(0.006)	(0.017)
Vertical Position	-0.079***	0.065*
	(0.007)	(0.032)
Second Half	-0.010	-0.005
	(0.022)	(0.021)
Choice Number	0.002	-0.016
	(0.014)	(0.013)
% Remaining Below		1.393***
		(0.294)
Intercept	0.479***	-0.247
	(0.064)	(0.191)
Clustered Ses	X	X
Search FE	X	X
Obs	1915	

Note. *=p<0.05 | **=p<0.01 | ***=p<0.001. Models were run on a truncated data set which excluded choices where participants lacked a choice.

We also examine choices at the individual level by averaging a variable coded -1 for a subsequent inspection below the most recently chosen item and 0 for one above. We take the average only of the first 5 inspections to hold the number of possible inspections equal across our conditions. Using all inspections provides similar results. We also do not include observations where a participant can only move in one direction.

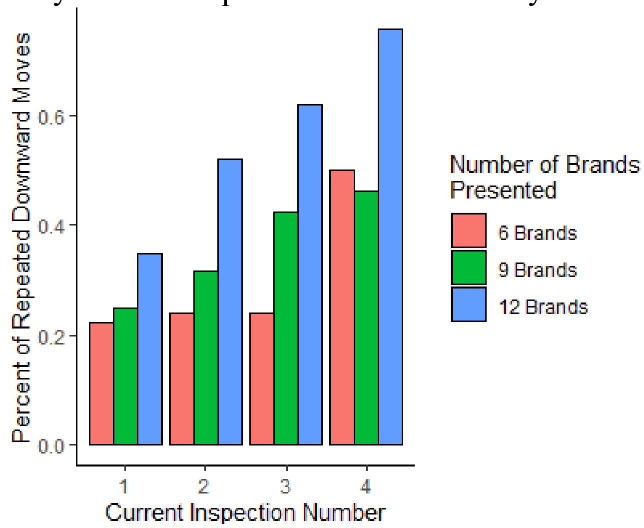
Overall, participants in the 12 brand condition are more likely to move down in their first 5 inspections compared with participants in the 6 brand condition, b = -0.151, t(140) = 3.085, p =

0.002, as are those in the 9 brand condition, b=-0.124, t(140)=2.523, p = 0.013 (M_6=-0.465, M_9=-0.589, M_{12}=-0.616). The 9 brand condition does not differ significantly from the 12 brand condition, b = -0.027, t(140) = 0.556, p = 0.579. Finally, treating the condition variable as continuous shows that moving to a condition with 3 more boxes (e.g., moving from 6 to 9) is associated with a 2.5 percentage point increase in the average percentage of the times that participants move down (versus up), t(141)=3.061, p=0.002.

Next, we look at how often participants' subsequent two inspections are below the current position. This is another measure of the degree to which participants next inspections tend to be below the current one. To make the data in our 3 conditions as similar as possible, we look only at the first 4 brands that participants inspected. This is because participants in the 6 brand condition could only inspect 6 brands, meaning that we could only look at two inspections ahead through the fourth inspection. We also excluded searches that involved either the top or the bottom brand on the screen. These restrictions left only 989 of our original 5756 observations (17.2%).

Participants in the 12 brand condition have subsequent inspections that move downward twice in a row 45.7% of the time while those in the 9 brand condition do so 29.7% of the time and those in the 6 brand condition do so only 23.2% of the time. These results should be interpreted with caution because there is more space to move downward as the number of brands increases. When we separate these data by the current inspection number (i.e., whether a person had just completed their first, second, third, or fourth inspection) there are similar trends, and the pattern is clearer after later inspections. This trend suggests that longer searches are associated with a stronger tendency to move downward. It is uncertain whether longer searches select for people who tend to search top-to-bottom. We present these results in Figure 3.1.4.

Figure 3.1.4. Percentage of Two Inspection Sequences that are Both Downward from the Most Recently Inspected Brand by Current Inspection Number in Study 1



Finally, with this constrained dataset, we regress whether the next two inspections were below the current one or not onto the number of brands presented, controlling for stimulus set. We run these regressions separately for each current inspection number because the sequential nature of moving downward twice introduces dependencies between sequential observations. We account for multiple observations per participant using clustered standard errors. Across all 4 regressions (one for each inspection number) we find that the probability of moving downward twice in a row is greater when there are more brands presented, all p's < 0.04. This result suggests that participants are more likely to inspect brands sequentially moving downward according to their position on the page when there are more brands presented. We present these regressions in Appendix Table 6.3.5.

3.1.3. Discussion

In Study 1, participants were more likely start their search at the top of the screen when there were more options available. Furthermore, there was some evidence that participants were more likely to search in a downward direction when there were more options available. When there were only 6 locations, there was little relationship between the vertical position of an item and whether it was inspected. On the other hand, when there were 12 locations, participants were likely to inspect an item below the one that they had just inspected. One interpretation of these results is that participants are increasingly relying on spatial heuristics to ease their memory load as the number of items increases. Rather than keeping track of the values from visiting a large number of locations, or repeatedly looking for the next highest value, participants may be looking at the locations guided by spatial location and inspecting them to determine if an RV or "satisficing" criterion has been exceeded. With that said, these results should be interpreted with caution. Our experimental design did not control for the length of search. And we cannot cleanly distinguish between the locations that participants choose and whether they have a bias to move downward. For example, one likely possibility is that sequential downward movements are a function of participants tending to start by inspecting locations higher up on the screen.

3.2. Study 2: The Impact of Display Features

In Study 1, we examined the impact of the number of possible search locations on search strategies. In Study 2, we hold the number of locations constant but vary the spatial layout of the locations on the screen and the format in which information about the locations is displayed. To vary the information format we present data about the possible range of outcomes numerically (as the bounds of a uniform distribution) in one condition while in the other we present them graphically. To vary screen layout we present locations stacked vertically (as in previous studies) in one condition, while in the other we present them in a circle.

We vary the screen and information formats to manipulate the difficulty of accessing cues to value. Overall, we expect that participants will be most likely to use value-based strategies (like the optimal RV strategy or an EV strategy) when it is easy to do so and use spatial strategies when it is relatively difficult. For example, a simple strategy that approximates the optimal strategy with low effort (which is the case in this study) is to inspect options in descending order of the maximum values possible at each location. In the graphical-linear condition, participants can look for the location with a rating furthest to the right to approximate the optimal model. This comparison is relatively easy because a participant can use the rapid visual processing to quickly find the best location. On the other hand, in the graphical-circular condition, locations are horizontally offset from one another making this type of rapid visual comparison difficult. Thus, we expect, that participants in the graphical-linear condition will search more optimally than those in the graphical-circular condition. And we expect that those in the graphical-circular condition will be more likely to rely on spatial strategies like inspecting attractive options in a clockwise direction without examining the entire set of locations.

In the numerical condition, participants can hold a numerical estimate of value in memory and conduct a series of comparisons between the values in locations and the values in working memory. While this memory-based comparison is more difficult than the visual comparisons in the graphical-linear conditions, it should be less sensitive to the screen format. In contrast to the graphical conditions, it is not more difficult to make these numerical comparisons in circular conditions than in the linear conditions. We therefore do not expect differences in search patterns between numerical conditions. In general, we expect that (i) participants will employ value-based strategies the most and spatial strategies the least in the graphical-linear condition, (ii) participants employ value-based strategies the least and spatial strategies the most in the graphical-circular condition, and (iii) that the numerical conditions will fall in between the two graphical conditions and not differ based on whether the locations are presented in a circle or a vertical line.

3.2.1. Method

Participants. We recruited 422 adults (54.3% Male, 45.0% Female; Average Age 40 years, see Table 3.2.1 for treatment assignments) to complete an experimental search task in an 18-minute session. Participants were paid a \$2.25 fixed participation fee and earned an additional average of \$1.47 in bonuses.

Table 3.2.1. Number of Participants in Each Treatment in Study 2

	Graphical	Numerical	Total
Linear	105	99	204
Circular	107	111	218
Total	212	210	422

Design. The study design was a 6 (within; unique sets of five prize Boxes) X 2 (within; Repetitions) X 2 (between; display format: numerical vs graphical) X 2 (between; spatial layout: vertical vs circular) design. After completing two practice searches, each participant completed 12 searches. Across between-subjects conditions, we varied the numerical format of the information for the locations. In the numerical condition, participants were presented information about the range of possible outcomes at a location as numbers that defined the endpoints of a uniform distribution. In the graphical condition, we presented participants with a sample of 30 draws from the same uniform distributions we used in Study 3 of Chapter 2. We also varied the spatial layout of the boxes between subjects such that in the linear condition, locations were displayed vertically (as in Study 1 of this chapter) while in the circular condition they were laid out in a circle. In Figure 3.2.1A and B we present examples of these conditions. Stimulus sets for the study were the same as those used in the 12-box condition of Study 1 of this chapter (See Appendix Table 6.3.1).

Procedure. We told participants that they were picking one movie from choice sets of twelve movies. Participants saw either a graphic depicting the ratings of 30 people who had seen the movie (Graphical condition) or the numerical bounds of a uniform distribution (Numerical condition). They were told that to learn how much they would like the movie, they could spend some time learning about the movie, for example by watching a trailer. They had to inspect at least one movie and could inspect as many as they wanted to. When they inspected a movie, the information depicting others' ratings for the brand would be replaced by their personal numerical rating (on a scale from 0 to 200). As in previous studies this numerical rating was randomly drawn each time a participant chose to learn more about a movie from the same uniform distribution represented in the location. Participants were charged a search cost of 1 point that represented the effort, time, and money required to inspect each movie.

Participants had 25 seconds to complete each of the searches. In a question asked at the end of the study, 96% of participants stated that they had enough time or too much time to complete each search. Across all conditions, 93% or more stated that they had enough time or too much time. For each search, a participant's point total was the maximum rating that they found in their search minus their total accrued search costs. We told participants that we would randomly select one of the 12 searches and pay them 1% of their profits in USD (Maximum Value – Search Costs) from that randomly selected search.

3.2.2. Analysis

Global Descriptive Statistics. Overall, participants inspected 2.95 options on average ($SD_{participant}$=1.66). The mean per-search value obtained (without accounting for cost) was 149.54 ($SD_{participant}$=6.01), although unsurprisingly there were significant differences between stimulus sets due to the different distributions used to generate values between them. The average profit (Maximum Found – Total Search Cost) was 146.59 ($SD_{participant}$=5.96).

Table 3.2.2. Average Duration (A), Maximum Value (B), and Profit (C) in Study 2

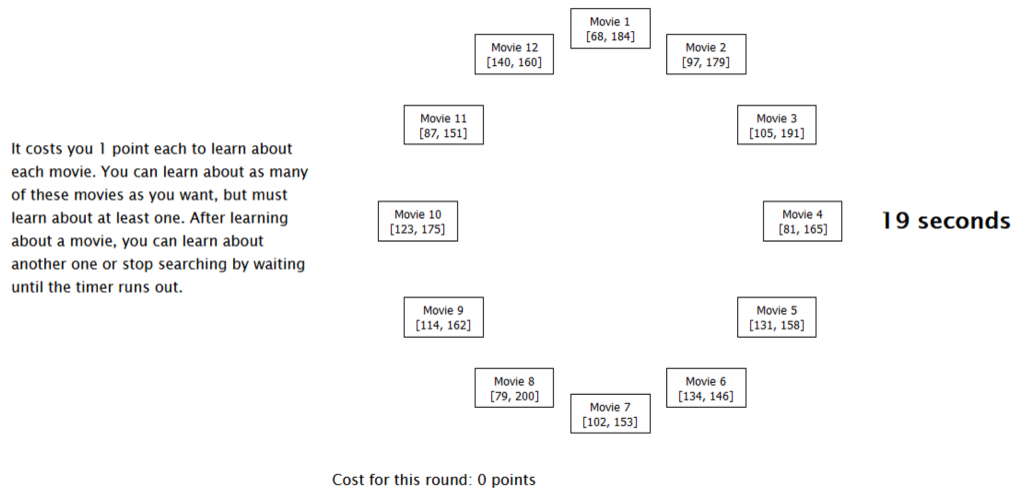
	(A) Duration					
	Graphical		**Numerical**		**Total**	
	M	SD	M	SD	M	SD
Linear	2.84	1.59	2.84	1.09	2.84	1.37
Circular	3.17	2.11	2.95	1.64	3.05	1.88
Total	3.01	1.87	2.90	1.41	2.95	1.66
	(B) Max Value					
	Graphical		**Numerical**		**Total**	
	M	SD	M	SD	M	SD
Linear	149.99	5.25	149.86	6.05	149.93	5.64
Circular	148.25	6.20	150.08	6.37	149.18	6.34
Total	149.11	5.80	149.98	6.21	149.50	6.02
	(C) Profit					
	Graphical		**Numerical**		**Total**	
	M	SD	M	SD	M	SD
Linear	147.15	5.28	147.02	6.11	147.09	5.68
Circular	145.08	5.69	147.13	6.49	146.12	6.18
Total	146.10	5.58	147.08	6.29	146.59	5.96

Note. Standard deviations are at the respondent level

For duration, maximum value, and profit, we ran regressions with each of these variables as the dependent variable and the graphical/numerical condition, circular/linear condition, and their interaction as independent variables. We controlled for the stimulus set, whether it was the first or second time a set was shown, and the order in which the stimulus was presented. We used cluster robust standard errors to account for the within-subjects design. We present full regression tables for these results in Appendix Table 6.4.1 and only provide highlights here.

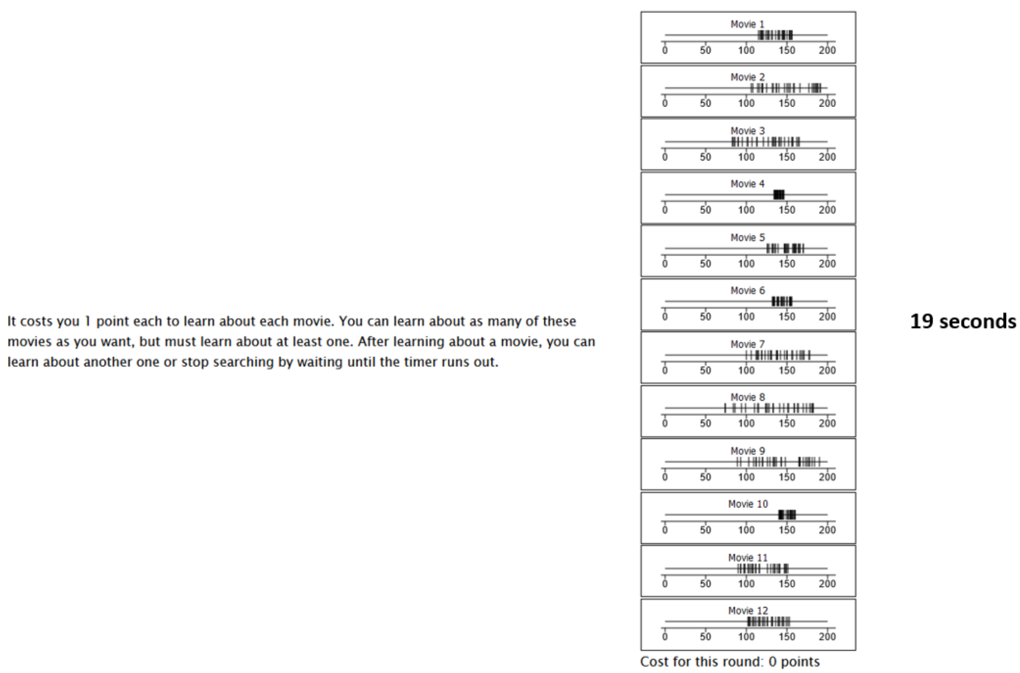
Figure 3.2.1. Examples of the (A) Numerical-Circular and (B) Graphical-Linear Layout from Study 2

Panel A



Note. When we index positions in the circular condition, 12 O'clock is indexed as 1 and 11 O'clock is indexed as 12.

Panel B



There was no difference between conditions in search duration. Among maximum values found within a search, there was a marginally significant interaction between the graphical/numerical and circular/linear conditions, $t(5053)=1.682$, $p=0.093$. Further inspection of this interaction shows that among the circular conditions, searches in the graphical condition ended with lower maximum values than those in the numerical condition, $b=-1.83$, $t(5053)=2.152$, $p=0.031$ ($M_{Graphical} = 148.25$, $M_{Numerical} = 150.08$). In addition, within the graphical condition, participants in the circular conditions ended with lower maximum values than those in the numerical conditions, $b=1.74$, $t(5053)=2.21$, $p=0.027$ ($M_{Linear} = 149.99$, $M_{Circular} = 148.25$).

Among profits within a search, there was a marginally significant interaction between the graphical/numerical and circular/linear conditions, $t(5053)=1.90$, $p=0.057$. Similar to the maximum value regression, an inspection of the interaction suggests that among the circular conditions, searches in the graphical condition ended with lower profits than those in the numerical condition, $b=-2.05$, $t(5053)=2.494$, $p=0.012$ ($M_{Graphical} = 145.08$, $M_{Numerical} = 147.13$). And within the graphical condition, participants in the circular condition ended with lower profits than those in the linear conditions, $b=2.07$, $t(5053)=2.76$, $p=0.006$ ($M_{Linear} = 147.15$, $M_{Circular} = 145.08$). Overall, maximum values and profits are lower in the circular/graphical condition than they are in the other 3 conditions suggesting that searching effectively based on value may be more difficult in this condition than in the other three.

First Choices. In this section, we examine first choices made in searches. These first choices provide a rough view of the strategies that participants were using as they navigated between locations. We first look at how often participants inspected locations first that were consistent with following an RV (i.e., having the highest RV) or an EV (i.e., having the highest EV) strategy. In Table 3.2.3, we provide the percentage of participants consistent with each strategy. Note that our stimulus sets were designed to distinguish between RV and EV strategy users, so based on first choices, participants could be classified as either RV, EV, or other. Overall, participants visited the RV location first 30.9% of the time while they visited the EV location first 33.8% of the time. Participants started with the maximum RV or EV location nearly 65% of the time.

Table 3.2.3. Percent of Participants Making RV or EV Choices First in Study 2

RV	Graphical	Numerical	Total
Linear	30.8%	31.5%	31.1%
Circular	28.3%	33.0%	30.7%
Total	29.5%	32.3%	30.9%

EV	Graphical	Numerical	Total
Linear	39.8%	31.8%	35.9%
Circular	33.1%	30.5%	31.8%
Total	36.4%	31.1%	33.8%

To analyze differences in first choices across conditions, we created two indicator variables: whether a participant selected the high RV location first and whether a participant selected the high EV location first. We regressed each of these binary variables onto the graphical/numerical condition, circular/linear condition, and their interaction as independent variables. We controlled for the stimulus set, whether it was the first or second time a set was shown, and the order in which the set was presented. We used cluster robust standard errors to account for the within-subjects design. We ran both logistic regressions and linear probability models. Since both models produced qualitatively similar results, we discuss highlights based on the linear probability models here. We provide both sets of models in Appendix Table 6.4.2.

There are no significant differences between conditions in the probability of inspecting the high RV location first. One trend worth noting is that among the circular conditions, participants tend to select the high RV choice more in the numerical condition than in the graphical condition, b = -0.048, $t(5053)$ = 1.296, p=0.195. When looking at visiting the high EV option first, there was no interaction between the Graphical/Numerical and Circular/Linear conditions, b = 0.053, $t(5053)$=0.997, p=0.319. However, among the graphical conditions, there was a greater probability of inspecting the high EV location in the linear condition than in the numerical condition, *b* = 0.067, *t*(5053) = 1.826, *p*=0.068. And among linear conditions, participants were significantly more likely to visit the high EV location first when information was presented graphically rather than numerically, *b*=0.079, *t*(5053)=2.095, *p*<0.036.

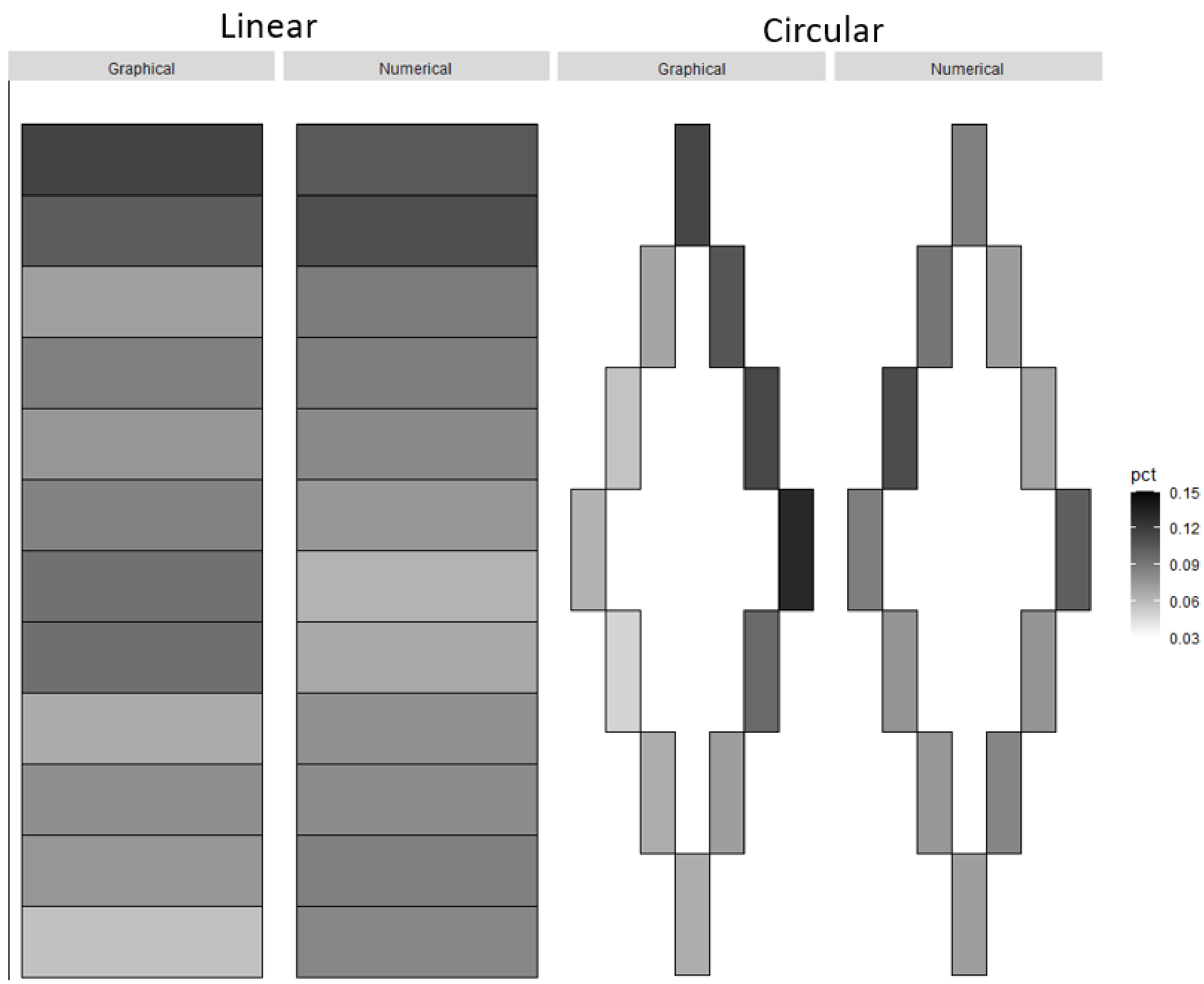
Next, we look at whether participants differ in their use of spatial strategies across conditions. We examine how often participants make choices based on the position of items on the screen rather than value-based cues (i.e., RV or EV). One simple way of understanding the importance of spatial position in search is by looking at the position of first searches. For example, in the linear conditions, if participants search from top to bottom, they are more likely to select items near the top of the screen early in their search. Similarly, in the circular conditions, if participants search clockwise from the top of the screen (12 O'clock), they are more likely to select items in the upper right quadrant of the circle of locations. In

Figure 3.2.2, we present the spatial position of first choices for each of our conditions. A visual inspection shows that there is a strong tendency for searches in the Graphical-Circular condition to be clustered spatially. In addition, the top two locations in the linear conditions are often selected first.

To examine this trend, we run two analyses. First, we look at how likely the selection of each box was compared to chance (8.34%) using proportion tests. In the Graphical-Circular condition, participants were more likely than chance to select one of the first 4 boxes (starting at the top and moving clockwise) and less likely than chance to inspect locations on the left side of the circle. There were no discernable patterns in deviations from chance for the Graphical-Numerical condition. In both linear conditions, the top two locations are more likely to be inspected than chance, and in the linear-graphical condition, the boxes near the bottom of the screen are less likely than chance to be selected first. We present these results graphically in

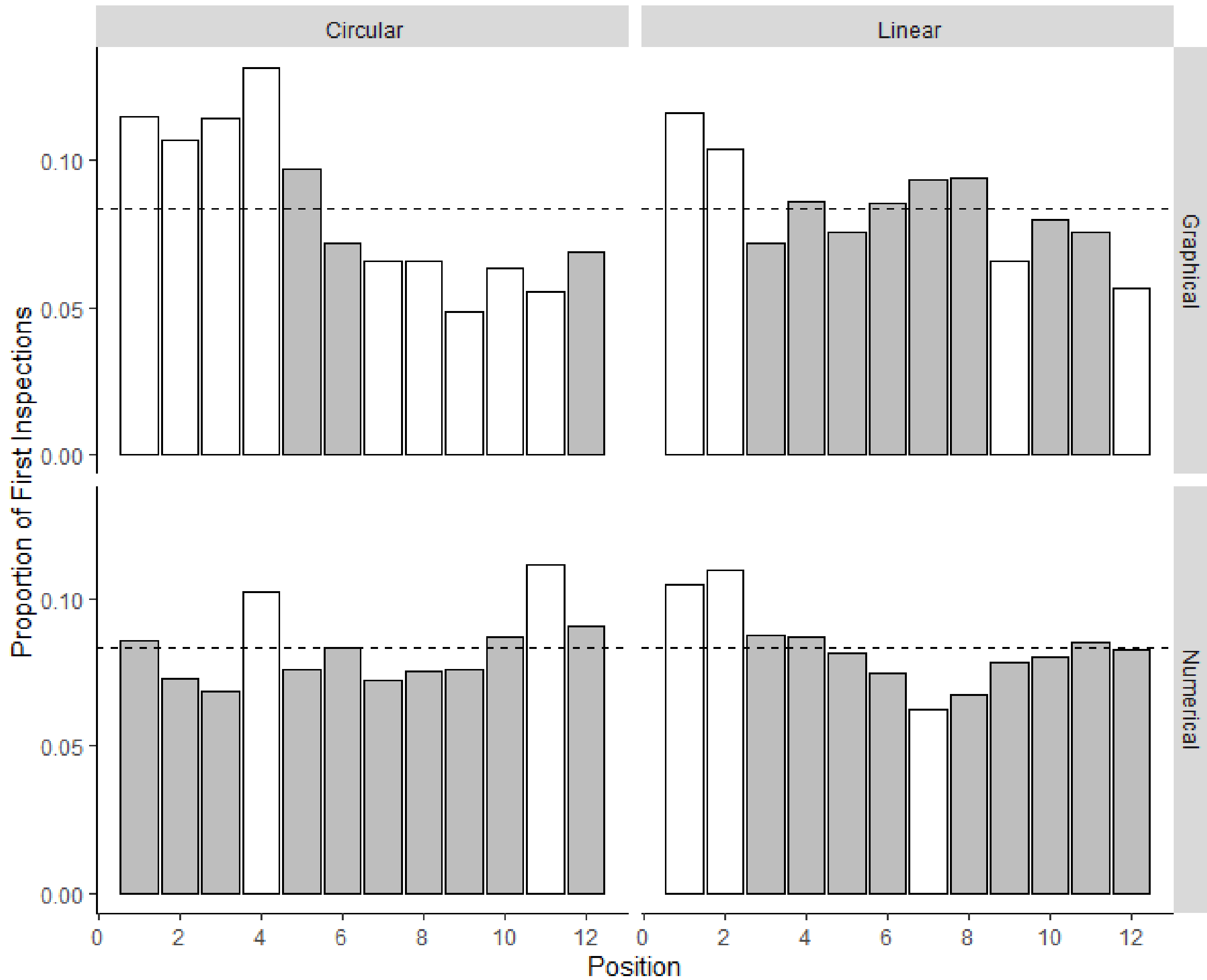
Figure 3.2.3. We present a table of this data in Appendix Table 6.4.3. We also run this analysis at the individual level. We look at the average of the percentage of first choices at each location for each participant and compare to the random, equal probability model (8.34%) using t-tests. We find very similar results, which we present in Table 6.4.4.

Figure 3.2.2. Spatial Position of First Choices in Study 2



Note. The proportion of first choices is indicated by colors with darker colors indicating higher proportions.

Figure 3.2.3. Proportion of Participant First Choices by Screen Position and Condition in Study 2



Note. White bars indicate proportions that are significantly different from random. Random is indicated by the dashed horizontal line.

Next, we compare across conditions. We regressed a dummy variable coded as 1 when the location of a first choice was one of the top 3 in the linear condition or in the upper right quadrant in the circular conditions[3] (12 O'clock to 2 O'clock). We regressed this variable onto the graphical/numerical condition, circular/linear condition, and their interaction as independent variables. We controlled for the same set of variables as in previous analyses and used cluster

[3] For convenience we refer to these 3 boxes as the "top" boxes in both the linear and circular conditions.

robust standard errors. As with previous analyses we ran these regressions as both logistic regressions and linear probability models. Since both models produced qualitatively similar results, we discuss the linear probability models here. We provide both sets of models in Appendix Table 6.4.5.

We find a significant interaction between Graphical-Numerical condition and the Circular-Linear condition, $t(5053)=3.801$, $p<0.001$. When information is presented numerically, participants are less likely to inspect one of the top locations in the circular condition than the linear condition, $b = -0.075$, $t(5053)=3.662$, $p<0.001$. When data is presented graphically, participants are marginally more likely to select one of the top boxes in the circular condition than the linear condition, $b=0.044$, $t(5053)=1.874$, $p=0.061$. When data is presented in a circular format, participants are more likely to inspect a top box first when data is presented graphically rather than numerically, $b=-0.108$, $t(5053)=4.985$, $p<0.001$. When data is presented linearly, there is no difference between graphical and numerical conditions, $b=0.011$, $t(5053)=0.483$, $p=0.629$. This interaction reflects the patterns in Figure 3.2.2 and Figure 3.2.3. Searches in the Circular-Graphical conditions are most likely to start with one of the top boxes and searches in the Circular-Numerical conditions are least likely to do so.

All Choices. In this section we look at all the items inspected by participants (as opposed to only first inspections). It is worth noting that since participants choose how many locations to inspect (rather than being controlled by the experimenter), the results in this section should be approached with caution.

Table 3.2.4. Average RV (Panel A) and EV (Panel B) of locations inspected in Study 2

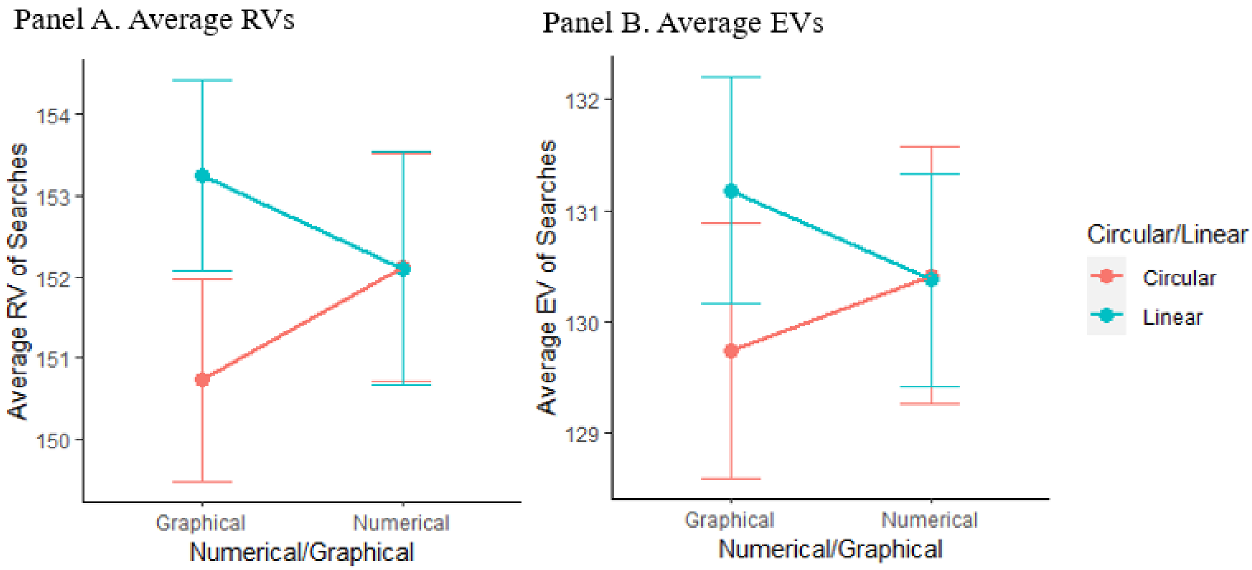
	Reservation Value					
	Graphical		**Numerical**		**Total**	
	M	SD	M	SD	M	SD
Linear	153.26	6.14	152.10	7.29	152.70	6.73
Circular	150.72	6.58	152.12	7.57	151.43	7.12
Total	151.98	6.48	152.11	7.42	152.04	6.96

	Expected Value					
	Graphical		**Numerical**		**Total**	
	M	SD	M	SD	M	SD
Linear	131.19	5.37	130.37	4.89	130.79	5.15
Circular	129.74	6.09	130.42	6.26	130.08	6.17
Total	130.46	5.78	130.40	5.65	130.43	5.71

Note. Standard deviations are at the respondent level

To examine participants' inspected locations, we look at the average RV and the average EV of the locations they visited. We present these averages by condition in Table 3.2.4 and Figure 3.2.4. Overall, participants in the graphical-linear condition have the highest average RVs and EVs while those in the graphical-circular condition have the lowest. There is no difference between circular and linear formats in the numerical conditions. To better understand these trends, we run regressions with the average RV and EV as dependent variables and the graphical-numerical and linear-circular conditions and their interaction as independent variables. We control for all factors that we controlled for in previous regressions in this chapter. In addition, we control for the length of the search since the average RV and EV will converge towards the average of the 12 available locations as searches grow longer. We use cluster robust standard errors to account for the within-subjects design. We include highlights of the regressions here and full tables in Appendix Table 6.4.6.

Figure 3.2.4. Average RV (Panel A) and EV (Panel B) of locations inspected in Study 2



Note. Error bars are 95% confidence intervals at the individual level

Regressing average RV onto our conditions yields a marginally significant interaction, $t(5052)=1.770$, $p = 0.077$. When we decompose this interaction, among the graphical conditions, searches in the linear condition have a higher average RV than those in the circular condition, $b=1.84$, $t(5052)=2.46$, $p=0.014$. There were no other significant differences between conditions. There is no interaction and no simple effects when we regress average EV onto our conditions.

Sequential Inspections. Finally, we examine sequential choices to understand when searches proceed in a top to bottom or clockwise manner. Due to the differences in format between the linear and circular conditions, we analyze these conditions separately. In the linear condition, we look at each inspection after the first and code it as being either above or below the previously inspected location. We cannot code our data in the same way in the circular condition because the circular condition does not have a salient top or bottom. For example, in the linear condition, if a person selected the bottom location (indexed as 12) and then the top location (indexed as 1), we would code this as moving upward in a subsequent inspection because the index of the first inspection is greater than the index of the second inspection. In contrast. If a participant inspects the option one unit counterclockwise from the top (indexed as 12) and then inspects the option at the top (indexed as 1) we would want to code this as a clockwise move even though the index of first inspection is greater than the index of the second inspection. To account for the circular layout we coded subsequent moves as the clockwise or counterclockwise based on the shortest number of moves between inspections. So, if a second inspection was 3 units clockwise and 9 unites counterclockwise of a first inspection, we would classify it as clockwise. Note that when a subsequent move is 6 units in either direction, we cannot classify a move as either clockwise or counterclockwise (We exclude these data points from our analyses). It is also worth noting that inferences based on these analyses should be approached with caution. Our experimental designs do not control for the length of search.

Linear Conditions. To better understand whether participants are biased to move up or down for each position, we compare their choices to a random model. This model selects one random location from the remaining locations with equal probability. A random model serves as a good comparison in this case because we randomized the position of the boxes for each participant. This randomization means that, with enough participants, the likelihood of moving up and down for the optimal model will converge on that of the random model.

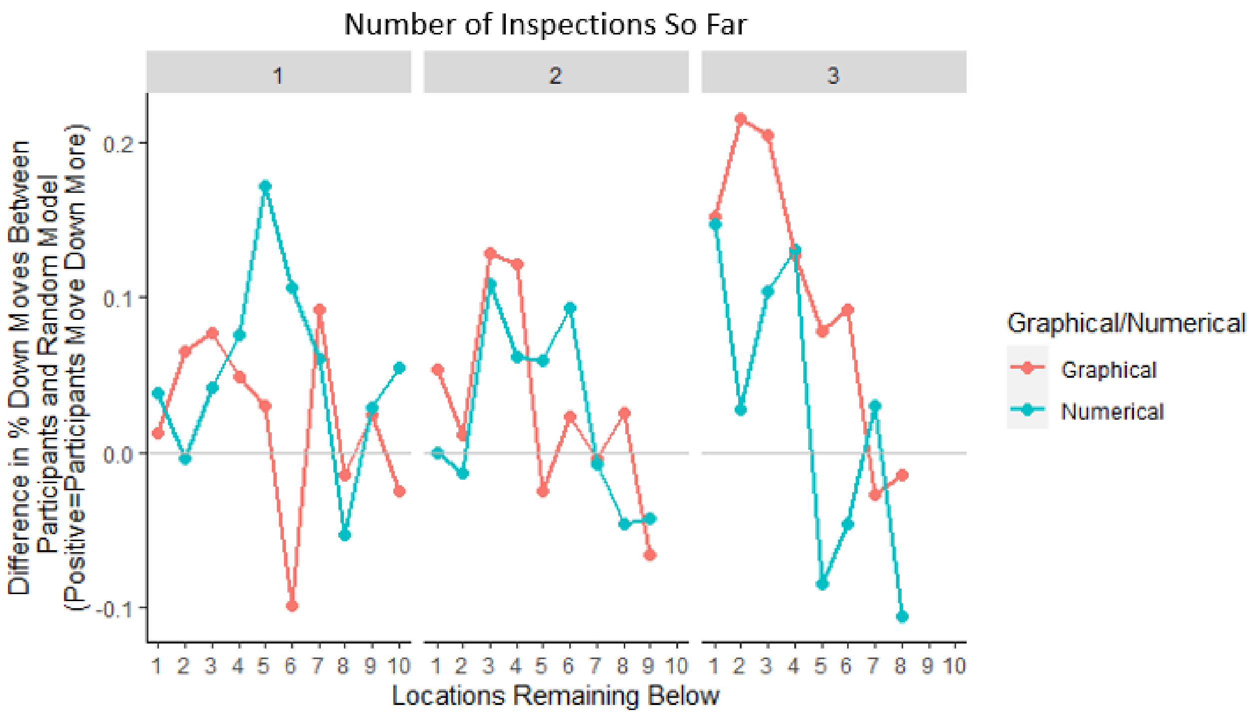
Based on this random model, we know that the probability of the random model moving up or down given the locations that have already been searched is proportional to the number of uninspected boxes above or below the current location. For example, if there are 12 locations, if the random model inspects the 9^{th} location from the top followed by the 5^{th} location, we know that the probability of moving down, given that the model inspects another location is 60% because there are 6 locations remaining below the most recently inspected location. In Figure

3.2.5, we depict the difference between the percent of participants' downward moves and the percent of downward moves by the random model. We do this separately for each number of locations left below the most recently inspected location, the numerical and graphical conditions, and the number of locations already inspected. Due to limitations in our data, we only look at cases where participants have already inspected one, two, or three locations. The panel for 3 inspections should be interpreted with caution because of small sample sizes.

First, to see whether there are any differences between graphical and numerical conditions, we run a regression of up and down moves (down = 1) on condition controlling for stimulus set, set presentation order, vertical position, the percent of locations below the current one and the inspection number. We use clustered standard errors to account for the within-subjects design. We run both a linear probability model and a logistic regression but report the results from the linear probability model here because the results are similar. We remove observations where participants are constrained to move in one direction. We find no difference between conditions, $b = 0.009$, $t(3442)=0.476$, $p = 0.634$. We present both models in Appendix Table 6.4.7.

A visual inspection of the linear conditions Figure 3.2.5 suggests that participants have slight bias towards moving downward in subsequent searches for positions, particularly when there are relatively few locations remaining below. To test this, we compare the proportion of participant who move up vs. down to the random model for each number of locations remaining above and below the most recently selected location. We run a regression of whether a participant moved up or down on condition each combination of remaining locations above and below. We only do this through the transition from the 3rd to the 4th inspections because later inspections have very few observations. We recentered the dependent variable on the random model proportion so that an intercept different from 0 would indicate more up or down moves than predicted. We use cluster robust standard errors to account for our within-subjects design. This analysis should be treated with caution because the regressions are not independent from one another.

Figure 3.2.5. Difference between Percent of Participants' Downward Moves and Percent of Random Model Downward Moves



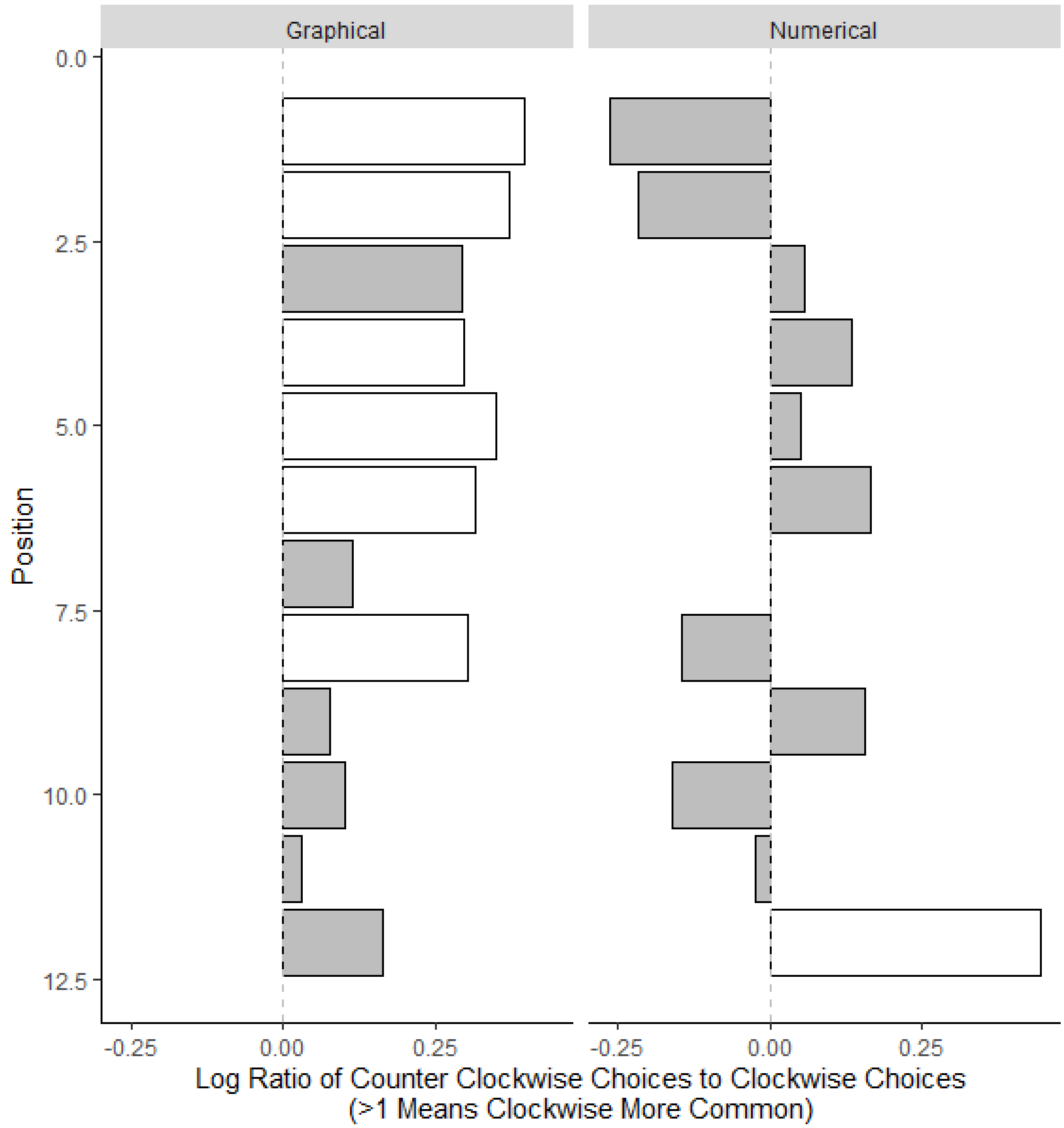
Note. We exclude cases where remaining locations are either all above or all below the most recently inspected locations because participants can only move in one direction.

Despite the visual trends, we find little evidence that participants are biased towards moving downward. It is possible that we were unable to detect a trend after participants had completed 3 inspections because of small sample size. It is also notable that, even though the intercepts individually were not significantly different from 0, 39 out of 54 coefficients (72%) were positive, which is significantly more than we would expect if there were no bias to move upward or downward, $X^2(1) = 9.796$, p=0.002. We present these models in in Appendix Table 6.4.8.

Circular Conditions. In Figure 3.2.6. Counterclockwise vs Clockwise Sequential Choices in Circular Conditions by Screen Position in Study 2 we depict the log ratio of clockwise versus counterclockwise subsequent inspections in the circular conditions for each screen position. A positive bar for a position indicates that participants are more likely to look at a box clockwise from (vs. counterclockwise from) the current position in the circular conditions. Based on a visual inspection of Figure 3.2.6, among the circular conditions, there is a clear bias toward inspecting options clockwise from the current position in the graphical condition while there is no trend in the numerical condition.

To examine the circular conditions, we first regress an indicator for whether a participant moves clockwise or counterclockwise onto the graphical/numerical condition. We controlled for the stimulus set, stimulus presentation order, the position on the screen, and the search number. We use clustered standard errors to account for the within-subjects design. There is a marginally significant difference between conditions, b = -0.051, t(4985)=-1.925, p=0.054, such that participants in the graphical condition are more likely to make sequential clockwise moves than those in the numerical condition. Unlike in the linear conditions, there is no impact of search number on subsequent inspections, so we do not examine it further. We present both the linear probability model and the logistic regression in Appendix Table 6.4.9.

Figure 3.2.6. Counterclockwise vs Clockwise Sequential Choices in Circular Conditions by Screen Position in Study 2



Note. A position of 1 is 12 O'clock on the circle. White bars indicate that significantly more inspections in that position were clockwise (vs. counterclockwise).

Next, we create a contrast code where clockwise moves are coded as 0.5 and counterclockwise moves are coded as -0.5. We regress this contrast code on the graphical/numerical condition for each screen position separately. If participants are more likely to move clockwise for a position, the intercept of the regression will be positive. Overall, in the graphical condition, there is a positive intercept for many positions in the first three-quarters of the circle, suggesting that participants tend to move clockwise. In the numerical condition, participants are only more likely to move clockwise than counterclockwise in the 12th position (11 O'Clock) only. Statistically significant positions are shaded in white in Figure 3.2.6. We present these intercepts in Appendix Table 6.4.10. Importantly, however, the graphical and numerical conditions only differed significantly from one another in the 1st, 2nd, and 8th positions. When we run these regressions on only the first inspection, we find similar, but attenuated results. Overall, these regressions suggest that participants in the Graphical-Circular condition are likely to search sequentially clockwise whereas those in the Numerical-Circular condition are not.

3.2.3. Discussion

In Study 2, we find that participants change their search strategies depending on the difficulty of accessing cues to value. Participants in the graphical-circular condition had the lowest average profits, had the lowest average RVs and EVs among the items they inspected, and used spatial strategies most often. In addition, we find very little difference between the linear and circular layouts when values were presented numerically but large differences between them in the graphical conditions. And we find suggestive evidence that participants are more likely to use value-based (vs spatial strategies) in the graphical-circular conditions than in the two numerical conditions. These results suggest that participants are altering their navigation strategies based on the ease of accessing cues to value. In cases where cues to value are difficult to access (graphical-circular) they are more likely to rely on spatial strategies than when they are easy to access (graphical-linear).

3.3. General Discussion

In Chapter 3, we looked at how consumers use the spatial layout of a search environment to navigate as complexity increases. In Study 1, participants were more likely to use a spatial strategy of searching from top-to-bottom when the number of options increased. In Study 2, participants were more likely to use spatial strategies (top-to-bottom or clockwise) when information used to evaluate locations was difficult to access and compare.

These studies suggest that participants switch between value-based and proximity-based strategies depending on the associated costs and benefits. Participants switch to proximity-based strategies when the cognitive costs of adhering to a value-based strategy increase relative to their benefits. In this way, consumers adaptively switch between strategies in order to find high-value items without incurring exorbitant physical and cognitive search costs.

4. Chapter 4. Learning

4.1. Study 1. Learning with 5 Boxes

In Study 1, we examine how learning impacts search. While using the same Graphical 5-box paradigm and automobile scenario that we used in Chapter 2, we have participants conduct 30 searches. Unlike previous studies, we do not change the distributions from which participants draw values. Note that since the random draws from inspected boxes changed, participants did receive varying feedback from inspections of the same boxes. In addition, we randomize the order of presentation of boxes in each search.

4.1.1. Method

Participants. We recruited 150 adults (66% female, 32% male) on Prolific Academic; no data was excluded from our analysis. The experimental sessions averaged approximately 22 minutes. Participants were paid a participation fee of $3.50 and an average bonus of $1.09.

Design. Participants completed 30 searches using the same set of 5-boxes. Participants were randomly assigned to one of 3 possible stimulus sets as they entered the study (see stimulus set information in Table 4.1.1). These stimulus sets differed in both search costs (1, 5, and 9) and in the distributions for boxes. It is important to note that because we vary both search costs and box distributions simultaneously, we cannot identify the source of any differences between conditions. For the remainder of this study, for convenience we refer to these 3 conditions by their costs (C1, C5, and C9). Draws from the boxes, the vertical order of the boxes, and the draws used to graphically depict the possible outcomes for each box were randomly generated for each search.

Procedure. We told participants that they were shopping for a new car and that in each search, they were considering buying one of five brands of cars. As in previous studies, in each search, participants saw a graphic depicting the user ratings by 30 people of each brand of car. They were told that to learn how much they would like the car, they would have to take the car for a test drive (inspection), and they could inspect up to five cars in each search. When they tested a car, the graphic for a brand would be replaced by their personal numerical rating for that car (on a scale from 0 to 200). This numerical rating was randomly drawn for each test drive from the same uniform distribution represented in the graphic. Participants were charged a 1, 5, or 9 point search cost that represented the effort, time, and money required to test drive a car.

Table 4.1.1. Stimulus Sets for Study 1

			Box Range			
Condition Name	Cost	Box Name	Lower Bound	Upper Bound	RV	EV
C1	1	A	24	142	126.6	83.0
		B	95	110	104.5	102.5
		C	90	137	127.3	113.5
		D	28	157	140.9	92.5
		E	40	115	102.8	77.5
C5	5	A	39	133	102.3	86.0
		B	79	111	93.1	95.0
		C	72	130	105.9	101.0
		D	40	146	113.4	93.0
		E	44	118	90.8	81.0
C9	9	A	52	164	119.1	108.0
		B	98	125	103.0	111.5
		C	91	151	118.1	121.0
		D	63	168	124.5	115.5
		E	25	172	120.6	98.5

As in our previous studies, participants had 20 seconds to complete each of the 30 searches. For each round, a participant's point total was the maximum rating that they found in their test drives during a search minus their total accrued search costs. To encourage experimentation, we told participants that we would draw their bonus from the final 5 searches.

We paid participants 1% of their profits in USD (Maximum Value – Search Costs) from the randomly selected round.

4.1.2. Results and Analysis

Global Descriptive Statistics. Overall, participants inspected 2.25 options on average ($SD_{participant}$=0.80). When we separate our 3 conditions, we find that higher search costs are associated with shorter search durations. The mean per-search value obtained (without accounting for cost) was 120.72 ($SD_{participant}$=10.5). Unsurprisingly there were significant differences between stimulus sets due to the different distributions used to generate values between them. The average profit (Maximum Found – Total Search Cost) was 110.2 ($SD_{participant}$=9.12). Finally, when we compare the participant profits to the profits made by the optimal model, participants earn 96% of the optimal model profits. We break down these statistics by condition in Table 4.1.2.

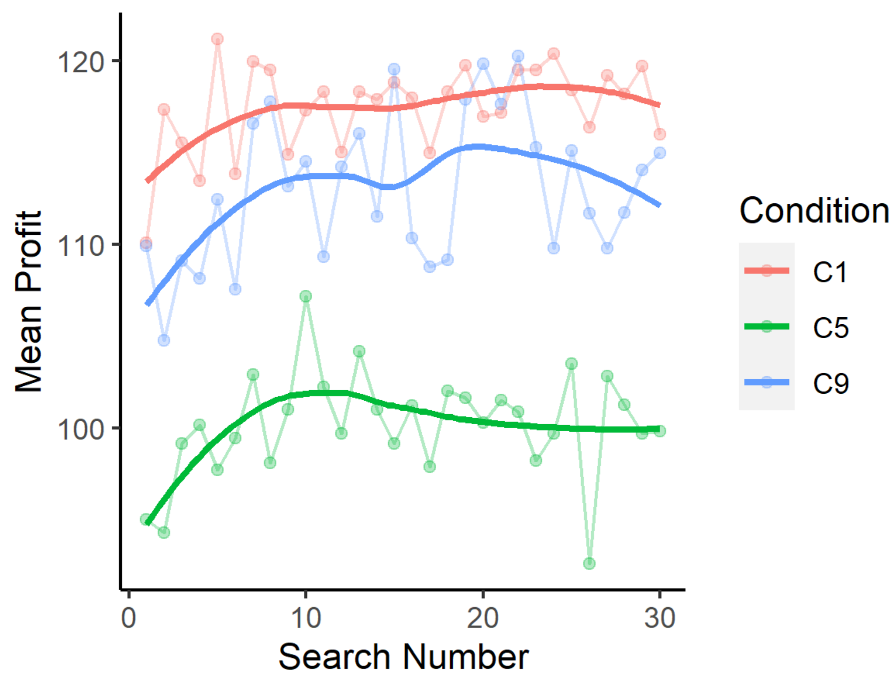
Table 4.1.2. Global Descriptive Statistics for Searches in Study 1

	Duration		Max Found		Profits		% Optimal Profit	
	M	SD	M	SD	M	SD	M	SD
C1	2.55	0.91	120	2.55	117.45	5.99	0.96	0.05
C5	2.16	0.57	110.95	2.16	100.14	4.19	0.95	0.03
C9	2.02	0.81	131.24	2.02	113.02	5.89	0.96	0.05
Total	2.25	0.80	120.72	2.25	110.2	9.12	0.96	0.04

Note. Standard deviations are at the participant level as opposed to at the individual search level

Profits. As a first step, we plotted the profit at each search number for each of our stimulus sets (see Figure 4.1.1). Note that raw profits should not be compared across conditions because the conditions utilized different stimulus sets and had different search costs. A visual inspection shows the mean profit increased in the early searches before leveling off. We fit two models to this data. The first was a linear model regressing profit onto search number (for example, for the 15^{th} search, we would regress the profit of the 15^{th} search onto 15), controlling for stimulus set and with cluster robust standard errors. Given that learning appears to slow in later searches, the second model used the natural logarithm of search number.

Figure 4.1.1. Mean Profits by Search Number

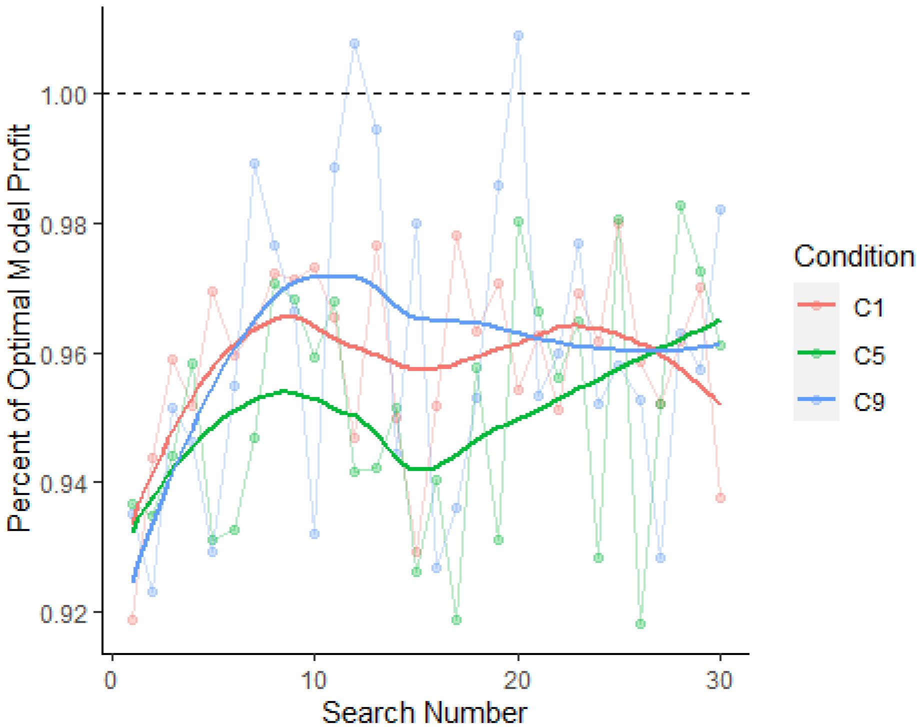


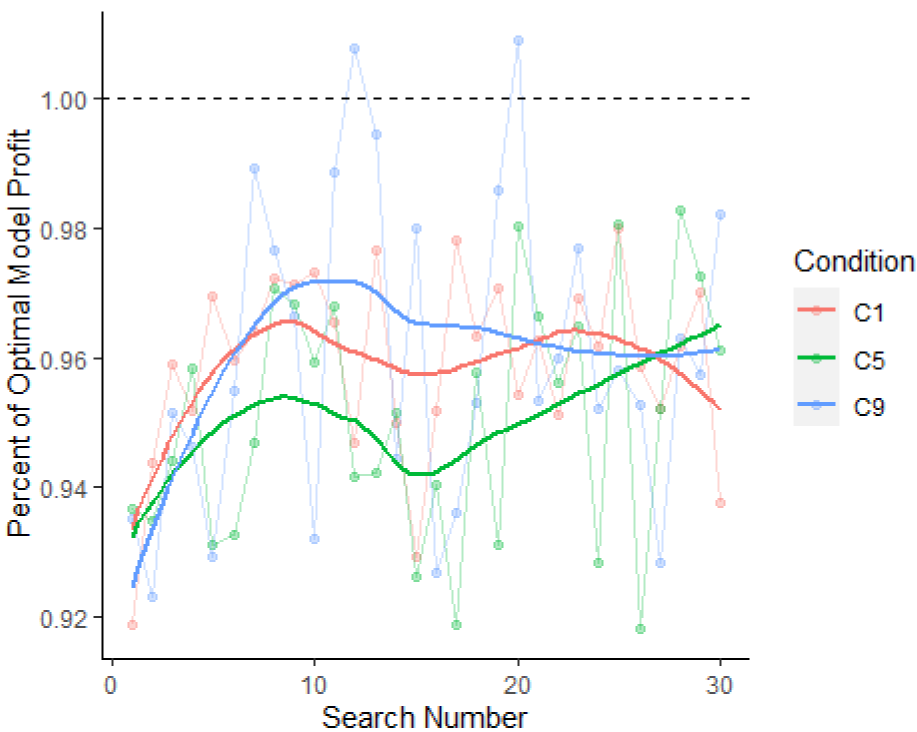
Note. Trend lines are LOESS lines.

Both the linear and the log model suggest that profits increase as participants repeat the search (linear: b=1.609, t(4496) = 3.271, p=0.001; log: b=1.569, t(4496)=4.764, p<0.001). The BIC suggests that log model fits the data better, which comports with the trends in Figure 4.1.1. We present full regression in Appendix Table 6.5.1. These results suggest that participants learn early in the task and at a declining rate. As a final test, we run our log model on the first 10 searches, the middle 10 searches, and the last 10 searches. This analysis allows us to see whether learning is isolated to certain phases of the task. Overall, increasing search number is associated with increasing profits in the first 10 searches, b = 3.34, t(1496) = 4.914, p<0.001. This coefficient suggests that in the first 3 searches participants gain an average of 3.3 points of profit, and then gain another 3.3 points of profit in (roughly) the next 5 searches. There was no relationship between block number and profit in the middle or late searches (middle: b=1.609, $t(1496) = 0.632$, p=0.527; late: b=-6.374, $t(1496)$=1.664, p=0.096). In late periods, the average may even decline, although increased variation in these final rounds makes estimates less reliable.

Since we recorded the draws for options that participants did not inspect, we can also compare participants' performance to the performance of the optimal model. Here, we divide participant profits by optimal model performance to create an index, where 1 is equivalent to optimal model performance. We present these trends in Figure 4.2. We regress these indices onto search numbers controlling for condition and with cluster robust standard errors to account for the within-subjects design. We run one model that uses the raw search number and another that uses the log of that number. We find that this index increases as participants gain experience (linear: b=0.004, $t(4496) = 1.846$, $p = 0.065$; log: b=0.006, $t(4496) = 2.511$, p= 0.012). The log model has a slightly lower BIC than the linear model (BIC_{log}=-4718.7; BIC_{linear}=4714.89), suggesting that learning slows as participants search more. Full regression tables are available in Appendix Table 6.5.2. Essentially, we replicate the early learning to a flat asymptote pattern in the first analysis of raw profits.

Figure 4.1.2.Percent of Optimal Model Profits by Search Number and Condition





Note. Trend lines are LOESS lines. Dashed line is parity between the optimal model profits and participant profits.

Finally, we look at individual changes in profit by running the log model on raw profits for each participant. Overall, 66% of participants had positive coefficients on search number, which suggests an increase in profits as they gain experience. When we look only at participants who had coefficients significant at the $p<0.1$ level, 10.7% had positive coefficients and only 1.3% had negative coefficients.

Since profits are calculated by subtracting search costs from the maximum value, we can think of two extremes in learning. One is that participants learn to find higher maximum values without increasing their search costs. The other is that participants inspect fewer options as they gain more experience but still manage to find the same maximum value. In the next two sections, we look at trends in maximum value and search durations (which for each stimulus set is equivalent to cost) to better understand the sources of learning.

Maximum Values. Unsurprisingly, the trends for the maximum value found looked very similar to trends for profits (see. Appendix Figure 6.5.1). A visual inspection showed that the mean maximum found increased in the early searches before leveling off, matching the pattern for profits. To examine this trend, we regressed the maximum value found onto (i) the search number and (ii) the log of the search number. We again controlled for condition with fixed effects and used cluster robust standard errors (see these models in Appendix Table 6.5.3).

Having completed a greater number of searches is associated with finding greater maximum values in both the linear and log models (linear: $b=0.125$, $t(4496) = 3.811$, $p<0.001$; log: $b=1.569$, $t(4496)=4.764$, $p<0.001$). As with profits, the log model had a superior fit, measured by BIC ($BIC_{log}=38514.64$; $BIC_{linear}=38495.65$). When we apply the log model to the early, middle, and late phases separately, we find the same pattern that we find in profits. Early on, participants increase the maximum values that they find, $b=3.921$, $t(1496) = 5.687$, $p<0.001$.

There was no relationship between block number and profit in the middle or late searches (middle: b=1.881, $t(1496) = 0.816$, p=0.414; late: b=-6.721, $t(1496)$=1.898, p=0.058). Again, in the late phase, maximum values appear to decrease.

Finally, looking at individual regressions, 58% have positive coefficients on search number suggesting increasing maximum values as they search more. When we look only at participants with coefficients that are significant at the $p<0.10$ level, 18.6% have positive coefficients while only 5.3% have negative coefficients. These results suggest that participants are finding higher maxima as they gain experience, and that the gains trail off. Given that these trends look similar to those for profits, it is likely that learning to find higher maxima is one source of increased profits.

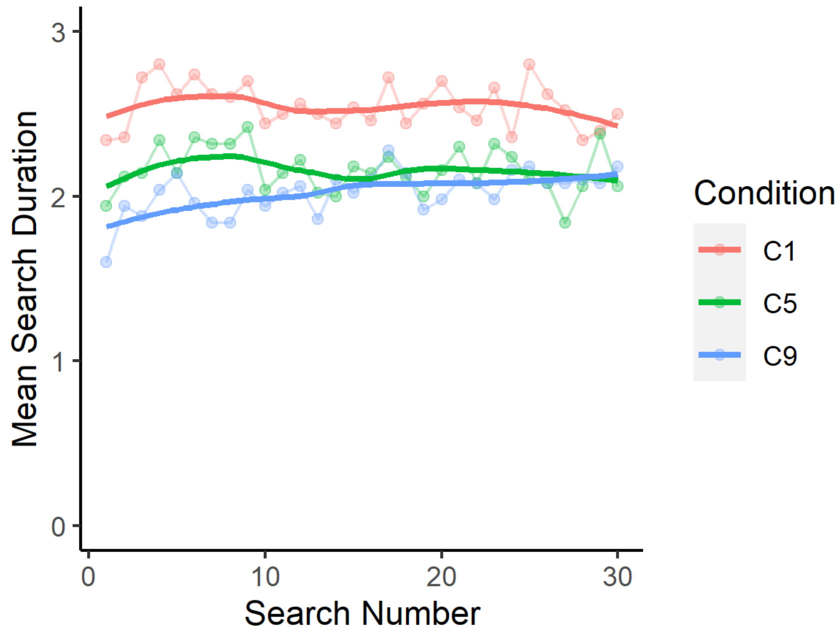
Search Durations (Costs). When we plotted search number against search durations, we did not see the trend that we observed for maximum values or profits (Figure 4.1.3). As for profits and maximum values, we regressed search duration onto (i) search number and (ii) the log of search number. We again controlled for condition with fixed effects and used cluster robust standard errors. There was no significant relationship between search number and search duration in either the linear or log models (see full regression models in Appendix Table 6.5.4).

When we run regressions of duration on block for each participant, 54.0% had negative coefficients suggesting that they reduced their search durations[1]. When we looked only at participants who had significant coefficients at the $p<0.1$ level, we found that 15.9% decreased search durations and 18.6% increased them. However, when we look at the conditions separately, we see a similar pattern as in our overall analysis. Participants in the Cost=1 and

[1] We removed 5 participants from this analysis because they had the same duration for all searches.

Cost=5 were slightly more likely to decrease the lengths of their searches, while participants in the Cost=9 condition were more likely to increase their search durations.

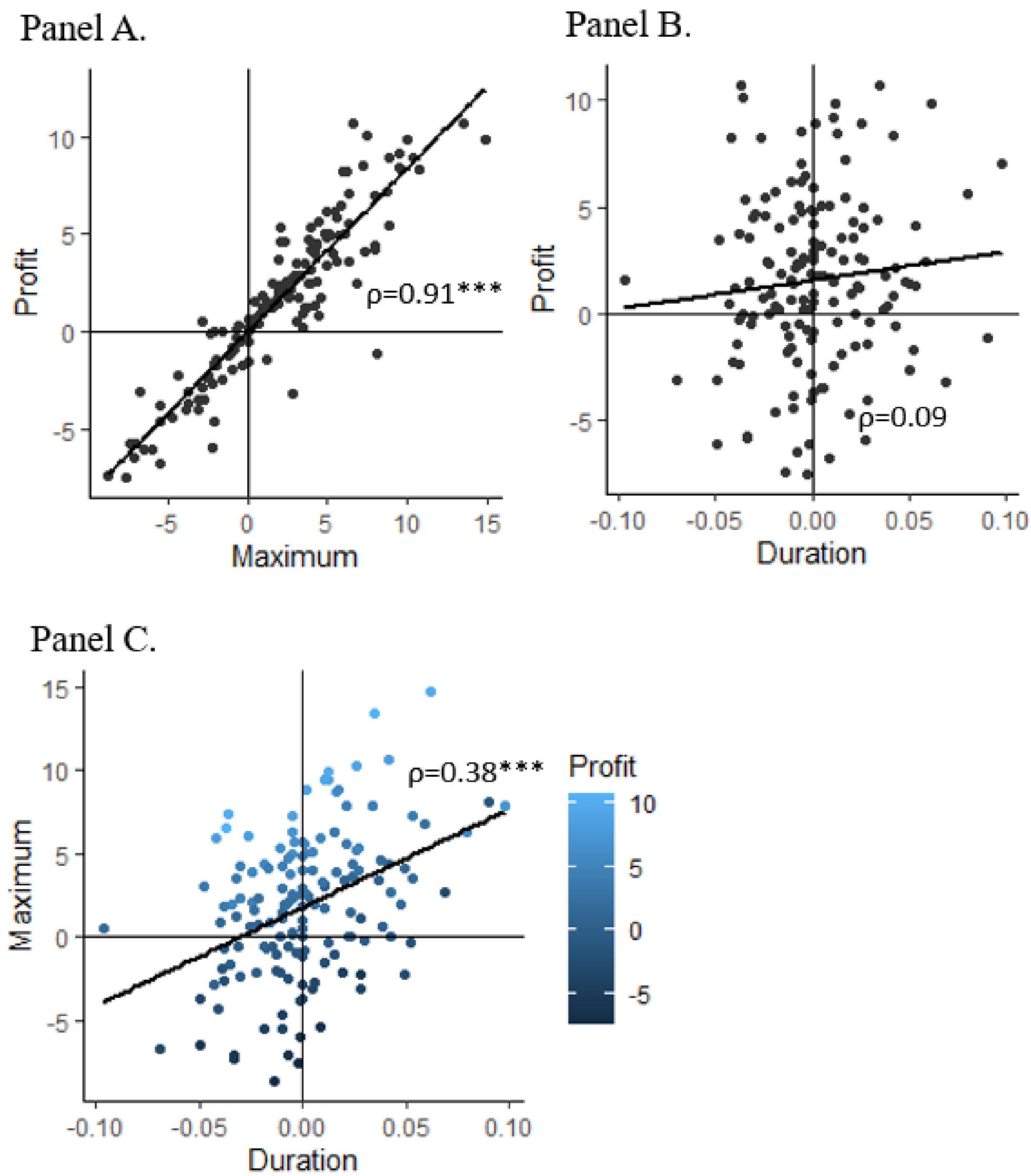
Figure 4.1.3. Mean Search Duration by Search Number



Note. Trend lines are LOESS lines.

Relationship Between Learning Based on the Maximum and Cost. As a final step, we look at the relationship between increasing profits and systematic changes in maximum found and search duration for individuals. We take the individual-level coefficients from our analyses of profits, maximum found, and duration and look at their relationship. We plot these relationships in Figure 4.1.4. Overall, learning (in the form of increased profits) is positively correlated with learning to find higher maximum values (spearman $\rho=0.91$, $p<0.001$). On the other hand, there is not a significant correlation between changes in search duration and increasing profits ($\rho=0.09$, $p<0.285$). Finally, there is a positive correlation between duration and finding the highest maximum ($\rho=0.38$, $p<0.001$). These relationships suggest that most of the increased profits come from participants learning to find higher maxima. Part of this ability may stem from finding better items by searching longer (as suggested by the positive relationship between coefficients on maximum and duration). But the weak relationship between increased duration and profit suggests that search length is not the primary driver of increased profits.

Figure 4.1.4. Relationships Between Individual Level Coefficients for (A) Maximum and Profit, (B) Duration and Profit, and (C) Duration and Maximum



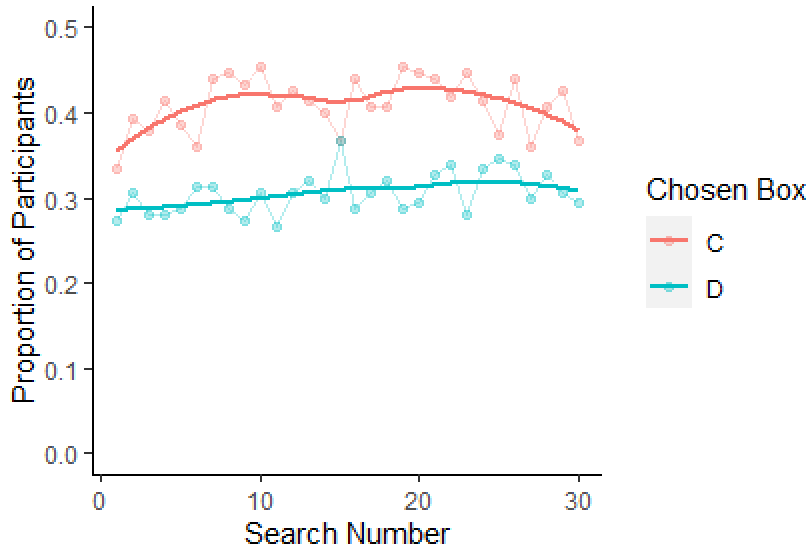
Note. Each point is a set of regression coefficients an individual.

Search Strategies. If participants find higher maxima without lengthening their searches by too much, they must be improving their strategies. In this section, we examine the evolution of participants' strategies.

As a first step, we plot the percentage of participants selecting each option for each search number (Figure 4.1.5). In all conditions, D is the choice with the choice with the highest reservation value and C is the choice with the highest expected value. In our analysis, we focus on these two options. We find no clear trend in our data. To confirm this lack of trend, we next run linear probability models where we regress the choice of box onto search number controlling for condition. We use cluster robust standard errors to account for the within-subjects design. We run separate models for the choice of box C and the choice of box D. Since we found non-linear trends in our earlier analyses, we also ran models with the log of search number. And we also run logistic regressions with the same variables. For the sake of simplicity, since all models were similar, we present the linear probability model where we do not take the natural log of search number. None of the models showed significant relationships (Box C: b=0.001, $t(4496)$=0.510, p=0.610; Box D: b=0.001, $t(4496)$=1.240, p=0.215). We present models in Appendix Table 6.5.5.

Next, when we do a similar visual inspection separated by condition, we see some trends, although they differ by condition. In the Cost=5 condition, where we saw only a modest increase in profits, we do not see much change in strategy. On the other hand, in the Cost=1 condition we see increased visiting of Box C (High EV) early and in the Cost=9 condition, we see increased visiting of Box D (High RV) early. To examine these trends, we repeat our analysis separated by condition. For ease of interpretability, we discuss linear probability models which have qualitatively similar results to logistic regressions (see both sets of models in Appendix Table 6.5.6).

Figure 4.1.5. Choice of First Box by Search Number



Note. Trend lines are LOESS lines. Box C is the maximum expected value choice while Box D is the maximum reservation value choice.

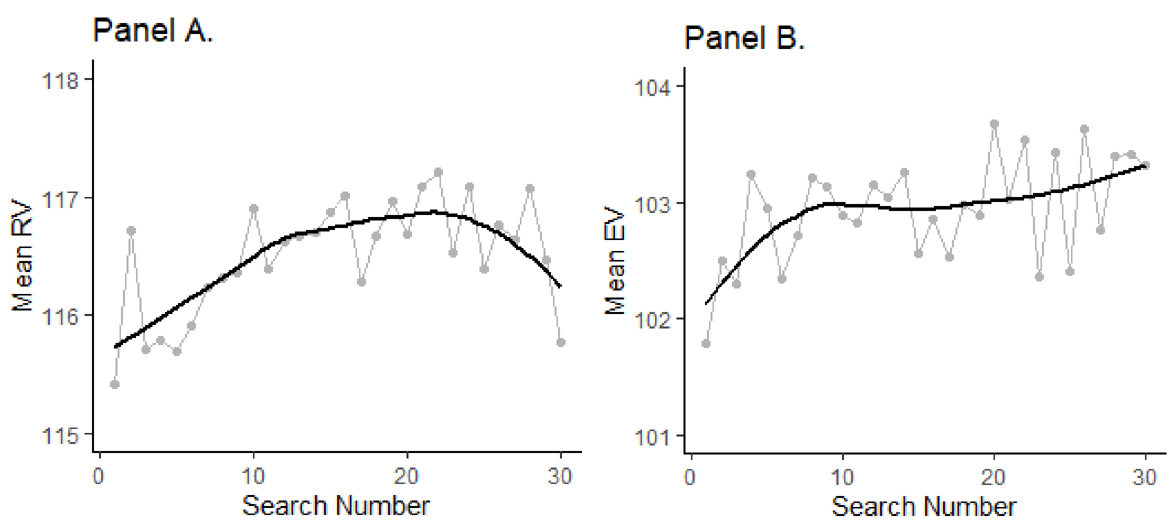
In the Cost=1 condition, we find weak evidence that the participants are more likely to select Box C first as they gain more experience. Models that regress the choice of Box C onto the natural logarithm search number find a marginally significant increase, $b = 0.038$, $t(1498) = 1.834$, $p=0.067$. We find no relationship between search number and opening Box D. In the Cost 5 condition, we find no relationship between search number and opening Box C or Box D first. In the Cost=9 condition, we find that the probability of opening Box D (the high RV box) increases as participants gain experience according to both the linear search number model ($b = 0.003$, $t(1498) = 3.191$, $p = 0.001$) and the log search number model ($b = 0.040$, $t(1498) = 3.470$, $p < 0.001$).

These results provide mixed evidence about strategies. We find no evidence of strategy change in the Cost=5 condition. There is some evidence of strategy learning In the Cost=1 and

Cost=9 conditions. Participants in the Cost=1 condition tend to increasingly visit high expected value options first while those in the Cost=9 condition tend to increasingly visit the high reservation value option. It is worth noting that even though the high EV option isn't the optimal choice, particularly in the Cost=1 and Cost=5 conditions, it has the second highest reservation value.

Even if participants are not visiting options in the order predicted by the optimal model, they may still be visiting those options in some order. To see whether this is the case, we look at the average RV and EV of the items that participants visit during a search. First, we plot the average RV and EV against the search number in Figure 4.1.6. There are modest increases in both Mean RV and Mean EV as participants gain more experience.

Figure 4.1.6. Mean RV (Panel A) and Mean EV (Panel B) by Search Number



Note. Trend lines are LOESS lines. The y-axes on the two panels differ but have the same scale.

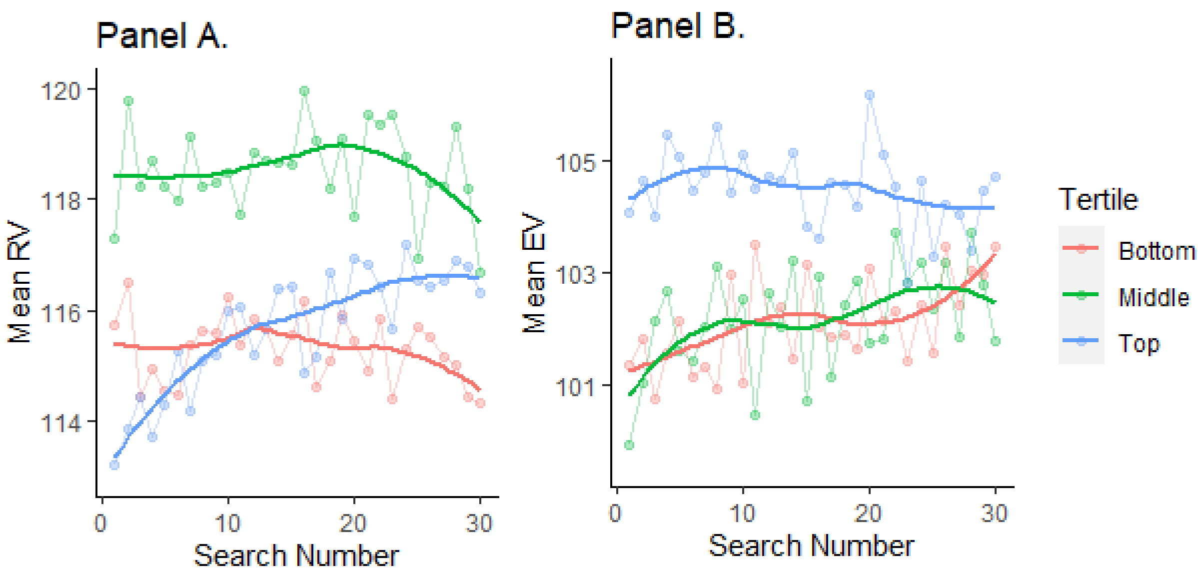
We regress the mean RV onto search number, controlling for condition and the length of search and with clustered standard errors for participants. Controlling for the length of search is important in this case because we would expect mean RV to converge towards the mean RV of the whole set as search lengths increase. There is a weak positive trend, *b*=0.028, *t*(4495) =

1.912, p =0.056. A similar model using the natural logarithm of search number finds a similar result, b=0.377, $t(4495)$ = 2.251, p =0.024. BICs for both models are similar. These results show a modest increase in the average RV of visited options as participants gain more experience. Next, we repeat these regressions with the average EV. In the model that uses the search number directly, there is an increasing trend, b=0.027, $t(4495)$ = 2.062, p =0.039. Using the log of search number also shows a significant increase, b=0.367, $t(4495)$ = 2.592, p =0.010. The log model has a slightly lower BIC. These results show that average EV also increases modestly with experience (see models in Appendix Table 6.5.7).

To explore EV versus RV strategy differences more sensitively, we divide our participants into three groups based on how much their performance (as measured in profits) increased over the 30 searches. We take the coefficients from our individual regressions of profit on search number and separate them into learning tertiles. More positive coefficients, in this case, indicate greater increases in profits with experience. We plot mean RV and mean EV by block separated by learning tertile in Figure 4.1.7. Overall, participants who improved their profits the most increased the mean RV of the items they inspected, while other participants did not. In contrast, participants who did not increase their profits as much increase the average EV of the items they inspect. Participants who increased profits did not increase average EVs. In Appendix Figure 6.5.2 we also separate these trends by condition. A visual inspection shows the same overall pattern for RVs in the Cost=1 and Cost=5 conditions, but not the Cost=9 condition, where average RVs increase rapidly in the first few searches and then stabilize.

We run regressions of Mean RV and Mean EV onto the natural logarithm of search number for each tertile separately. We use the model with the natural log because it had a lower BIC in our previous analysis. Participants who improved their profits the most increased the average RV of items they visited as they gained experience, *b*=1.160, *t*(1495)=4.008, *p*<0.001, but they did not increase the average EV of items, *b*=-0.058, *t*(1495)=0.282, *p*=0.777. In contrast, participants in both the middle and bottom tertiles (who did not increase profits as much), did not increase the average RV of the items they visited as they gained experience (middle: *b*=0.111, *t*(1495)=0.494, *p*=0.621; bottom: *b*=-0.119, *t*(1495)=0.383, *p*=0.702). In contrast, both groups moderately increased the average EV of the items they visited as they gained experience (middle: *b*=0.677, *t*(1495)=2.373, *p*=0.018; bottom: *b*=0.486, *t*(1495)=2.131, *p*=0.033). We present these models in Appendix Table 6.6.6.Table 2.1.1

Figure 4.1.7. Mean RV (Panel A) and Mean EV (Panel B) by Search Number and Learning Tertile



Note. Trend lines are LOESS lines. The y-axes on the two panels differ but have the same scale. Colors represent tertiles of profit increase such that participants in the top tertile (Blue) increased their profits the most over the 30 searches.

4.1.3. Discussion

Participants in this study increased their performance (measured in profit) early in the task. This increase in performance comes mostly from being able to find higher maximum values rather than from reducing the lengths of their searches. When we examined whether specific strategies could explain these gains, we find some suggestive evidence that participants whose profits increase more are shifting to the RV strategy. Participants in all three conditions were more likely to look at certain options later in the task: in the Cost = 1 and Cost = 5 conditions, participants increasingly visited the high expected value options as they gained experience. When we compare participants who increase their profits the most we find that they increase the average reservation value of the items they visit over time, but not the average expected value. Other participants, who do not increase their profits as much, do not increase the average RV of the items they visit, but do increase the average EV of those items. Overall, participants who learn the most from this task increase the quality of the options that they visit while others do not.

4.2. Study 2. Learning with 12 Locations

In Study 2, we extend our investigation into learning by having participants repeatedly search when there are 12 locations rather than 5. Having 12 locations raises the complexity of search because participants are offered more choices. A participant searching according to the Weitzman model, for example, must now calculate and track the reservation values of 12 locations instead of 5. In Chapter 3, we saw that increasing the number of boxes can change the patterns that people use while searching. As in Study 1 of this chapter, we have participants conduct 30 searches with the same set of 12 locations. Unlike in Study 1, we varied search costs while holding the distributions from which participants drew values constant.

4.2.1. Method

Participants. We recruited 131 adults with the goal of sampling 100 participants. We discarded data for 5 participants who did not sample an item from one or more locations in all 30 searches. We removed another 23 participants who indicated that the 12 boxes did not fit on their display screens. Our final sample was 103 participants (64% male, 35% female) on Prolific Academic.

The experimental sessions averaged approximately 26 minutes. Participants were paid a participation fee of $3.50 and an average bonus of $1.45.

Design. The design for Study 2 was nearly identical to that of Study 1. Participants completed 30 searches using the same set of 12-boxes. Participants were randomly assigned to one of 2 possible cost conditions (Cost =1 or 4). We present our stimulus set in Table 4.2.1. For the remainder of this study, we refer to these 2 conditions by their costs (C1, C4). Draws from the location, the vertical order of the boxes, and the draws used to graphically depict the possible outcomes for each box were randomly generated for each search.

Table 4.2.1. Stimulus Sets for Study 2

	Box Range		RV		
Box Name	Lower Bound	Upper Bound	Cost=1	Cost=4	EV
A	42	191	173.8	156.5	116.5
B	62	174	159.0	144.1	118
C	119	161	151.9	142.7	140
D	56	142	128.9	115.7	99
E	101	124	117.2	110.4	112.5
F	33	129	115.2	101.3	81
G	111	121	116.5	112.0	116
H	83	125	115.8	106.7	104
I	61	126	114.6	103.2	93.5
J	21	133	118.0	103.1	77
K	55	126	114.1	102.2	90.5
L	97	121	114.9	107.1	109

Procedure. We told participants that they were picking a movie. In each search, they were considering one of twelve movies. As in Study 1, in each search, participants saw a graphic depicting the ratings of 30 people that had seen the movie. They were told that to learn how much they would like the movie, they could spend some time inspecting the movie, for example by watching a trailer. They had to inspect at least one movie but could inspect about all of them if they wanted to. When they inspected a movie, the graphic depicting others' ratings for the brand would be replaced by their personal numerical rating (on a scale from 0 to 200). As in previous studies this numerical rating was randomly drawn each time a participant chose to learn more about a movie from the same uniform distribution represented in the graphic. Participants were charged a search cost that summarized the effort, time, and money required to inspect a movie.

Participants had 25 seconds to complete each of the 30 searches. In a question asked at the end of the study, 96% of participants stated that they had enough time or too much time to complete each search. For each search, a participant's point total was the maximum rating that they found in their inspections minus their total search costs. We told participants that we would randomly select one of the full set of 30 searches (as opposed to one of the last 5 searches, as in Study 1). We paid participants 1% of their profits in USD (Maximum Value – Search Costs) from the randomly selected search.

4.2.2. Results and Analysis

Global Descriptive Statistics. We present global statistics overall and by condition in Table 4.2.2. Overall, participants inspected 2.33 options on average ($SD_{participant}$=1.89). When we separate our 2 conditions, we find that higher search costs lead to shorter search durations, $t(69.8)=3.23$, $p<0.002$. The mean per-search value obtained (without accounting for cost) was 146.52 ($SD_{participant}$=7.2), and there were no differences by condition, $t(101.0)=1.39$, $p=0.167$. The average profit (Maximum Found – Total Search Cost) was 141.3 ($SD_{participant}$=7.84) was unsurprisingly higher when search costs were lower, $t(100.97) = 5.17$, $p<0.001$. Finally, since we recorded the draws for movies that participants did not inspect, we can also compare participants' performance to the performance of the optimal model. Here, we divide participant profits by optimal model performance to create an index, where 1 is equivalent to optimal model performance. When we compare the participant profits to the profits made by the optimal model, participants earn 96.2% of the optimal model profits. There was no difference between cost conditions, $t(98.9)=0.814$, $p=0.418$.

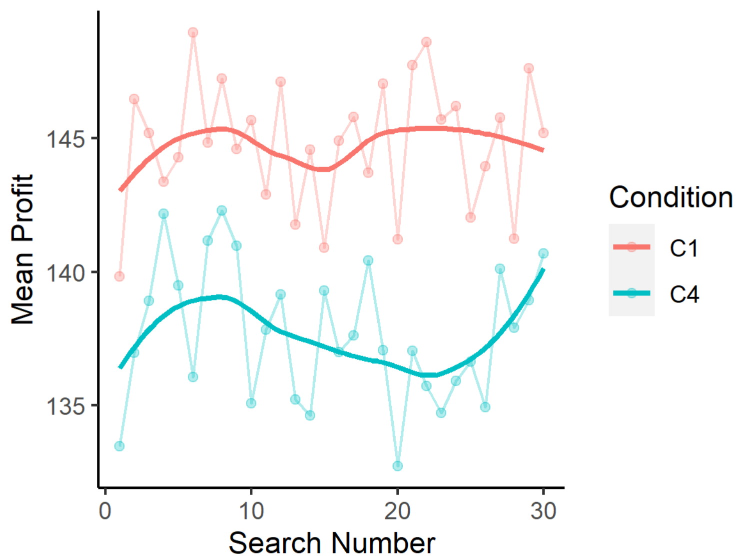
Table 4.2.2. Global Descriptive Statistics for Searches in Study 2

	Duration		Max Found		Profits		% Optimal Profit	
	M	SD	M	SD	M	SD	M	SD
C1	2.68	1.47	147.5	7.31	144.8	7.02	96.5%	3.8%
C4	1.96	0.64	145.5	7.07	137.7	6.99	95.9%	4.4%
Total	2.32	1.19	146.52	7.23	141.28	7.84	96.2%	4.1%

Note. Means and standard deviations are of mean individual values

Profits. As a first step, we plotted the profit at each search number for each of our stimulus sets (see Figure 4.2.1). A visual inspection showed no clear signs of learning. As in Study 1, we fit two models to this data. The first was a linear model regressing profit onto search number (for example, for the 15^{th} search, we would regress the profit of the 15^{th} search onto 15), controlling for stimulus set and with cluster robust standard errors. Given that learning usually slows in later searches, the second model used the natural logarithm of search number. Both the linear and the log model suggest that profits do not increase as participants gain experience with the task (linear: b=-0.001, $t(3087) = -0.239$, $p = 0.811$; log: b=0.124, $t(3087) = 0.282$, $p = 0.778$, see models in Appendix Table 6.6.1).

Figure 4.2.1. Mean Profits by Search Number in Study 2

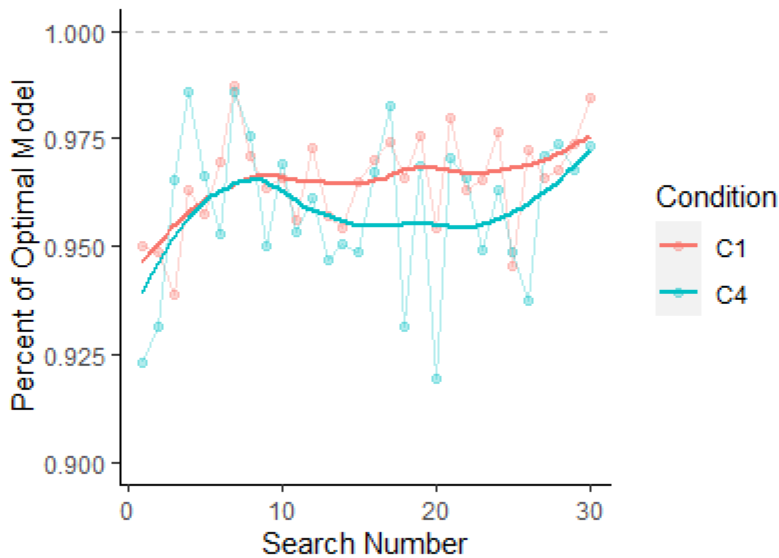


Note. Trend lines are LOESS lines.

Finally, we look at individual changes in profit by running the log model for each participant. We choose the log model because it fit best in the previous study. Overall, 46.6% of participants had positive coefficients on search number, which suggests an increase in profits as they gain experience. When we look only at participants who had coefficients significant at the p<0.1 level, 11.6% had positive coefficients and 6.8% had negative coefficients. In contrast to Study 1, participants do not increase raw profits as they gain experience with the task.

Comparison to Optimal Profits. We present participant performance compared to the optimal model for each search number in Figure 4.2.2. We regress these indices onto search numbers controlling for condition and with cluster robust standard errors to account for our within-subjects design. We run one model that uses the raw search number and another that uses the log of that number. This index increases as participants gain experience (linear: b=0.000, $t(3087) = 1.793$, $p = 0.127$; log: b=0.005, $t(3087) = 2.048$, p= 0.040). Next, we add an interaction between the condition and the search number to look for differences in learning between conditions. We do not find any signs of an interaction in either the linear or log models (linear: *b*=-0.000, *t*(3086) = 0.813, *p* = 0.416; log: *b*=0.005, *t*(3086) = 0.453, *p*= 0.650). So, we find small learning effects on one measure of profit improvement (see Appendix Table 6.6.2).

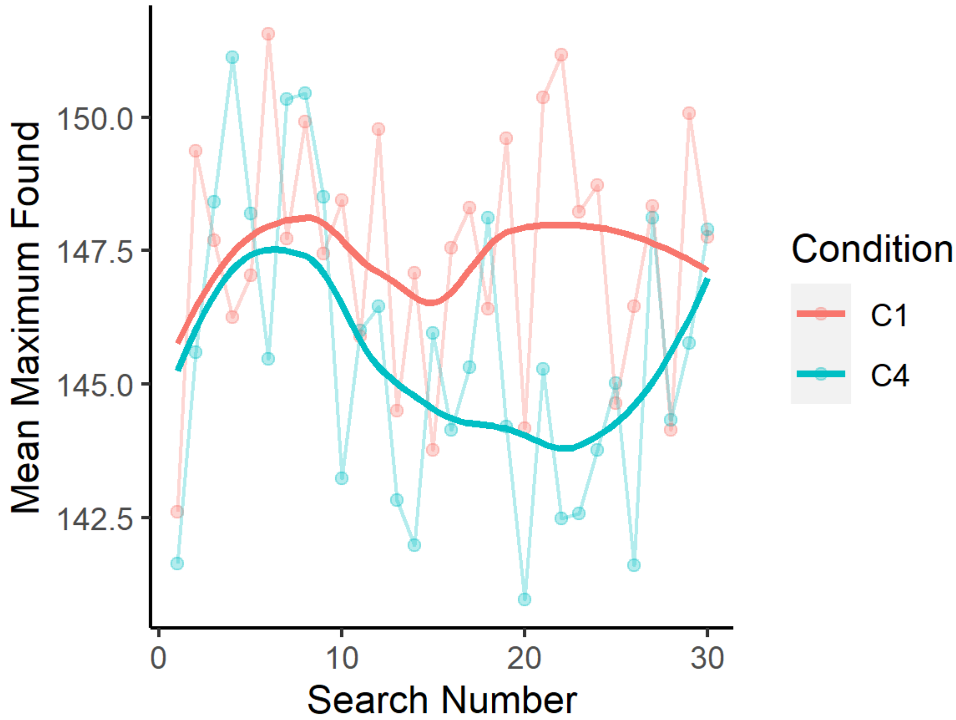
Figure 4.2.2. Percent of Optimal Model Profits by Search Number and Condition in Study 2



Note. Trend lines are LOESS lines. Dashed line is parity between the optimal model profits and participant profits.

Maximum Values. As with profits, we plotted the maximum value found at each search number for each of our stimulus sets (Figure 4.2.3). A visual inspection shows that the mean maximum found increased in the early searches in both conditions. In the Cost=4 condition, there is a substantial dip in the maximum value found in the middle of the task followed by an increase in maximum value towards the end. In the Cost=1 condition, we do not see a similar dip. before leveling off in a similar pattern to profits. When we run a regression of maximum value found on search number (and log search number), we find no linear relationships, (linear: *b*=-0.044, *t*(3086) = -0.449, *p* = 0.654; log: *b*=-0.004, *t*(3086) = -0.992, *p*= 0.321, see models in Appendix Table 6.6.3).

Figure 4.2.3.Mean Maximum Values Found by Search Number in Study 2



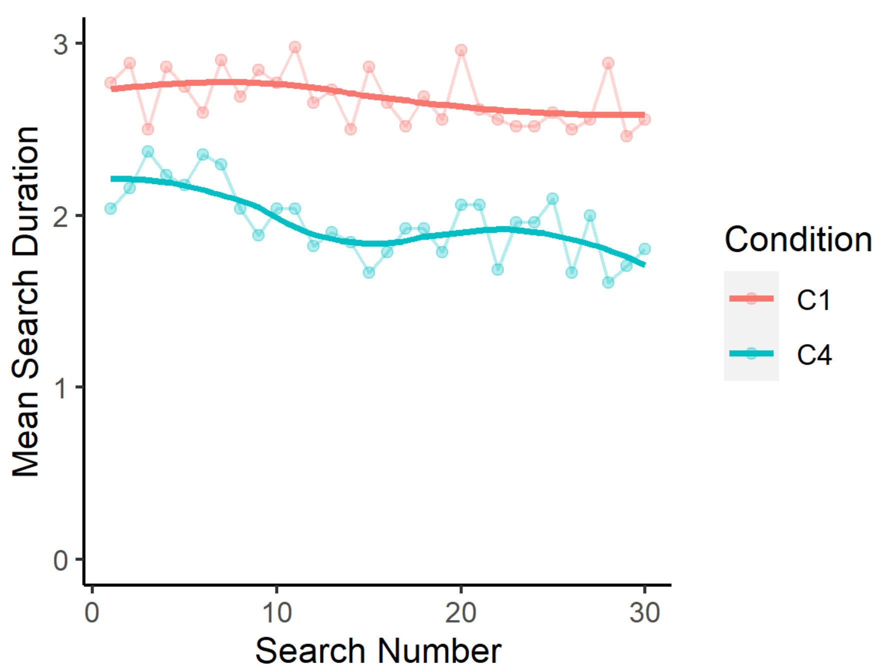
Note. Trend lines are LOESS lines.

Looking at individual regressions, 47.6% have positive coefficients on search number suggesting increasing maximum values as they search more. When we look only at participants with coefficients that are significant at the *p*<0.10 level, 8.7% have positive coefficients while 12.6% have negative coefficients. Unlike Study1, these results suggest that participants are not finding higher maxima as they gain experience. Overall, improvements in profits and maxima from learning are isolated in a small "getting acquainted with the task" effect in the first 5 searches. There are no systematic increases in these measures in the last 25 searches.

Search Durations. When we plotted search number against search durations, we did not see the trend that we saw for maximum values or profits (Figure 4.2.4). We ran two models. In one, we regressed search duration onto search number and condition. In the second model, we included the interaction between search number and condition because a visual inspection of the data suggested different trends by condition. We did not run models with the log of search number because a visual inspection of the data suggested a linear trend. We again used cluster robust standard errors to account for our within-subjects design. Overall, there was a decrease in search duration as participants gained experience, $b = -0.012$, $t(3087) = 3.197$, $p = 0.001$. When we include the interaction of search number and condition, there is no significant interaction, $b = -0.007$, $t(3086) = -1.097$, $p = 0.273$, see the models in Appendix Table 6.6.4).

When we run regressions of duration on search number for each participant, 68.0% had negative coefficients suggesting that they reduced their search durations. When we looked only at participants who had significant coefficients at the p<0.1 level, we found that 26.2% decreased search durations and 7.8% increased them. Overall, we find decreased search durations across both of our conditions, unlike the pattern in Study 1.

Figure 4.2.4. Mean Search Duration by Search Number in Study 2

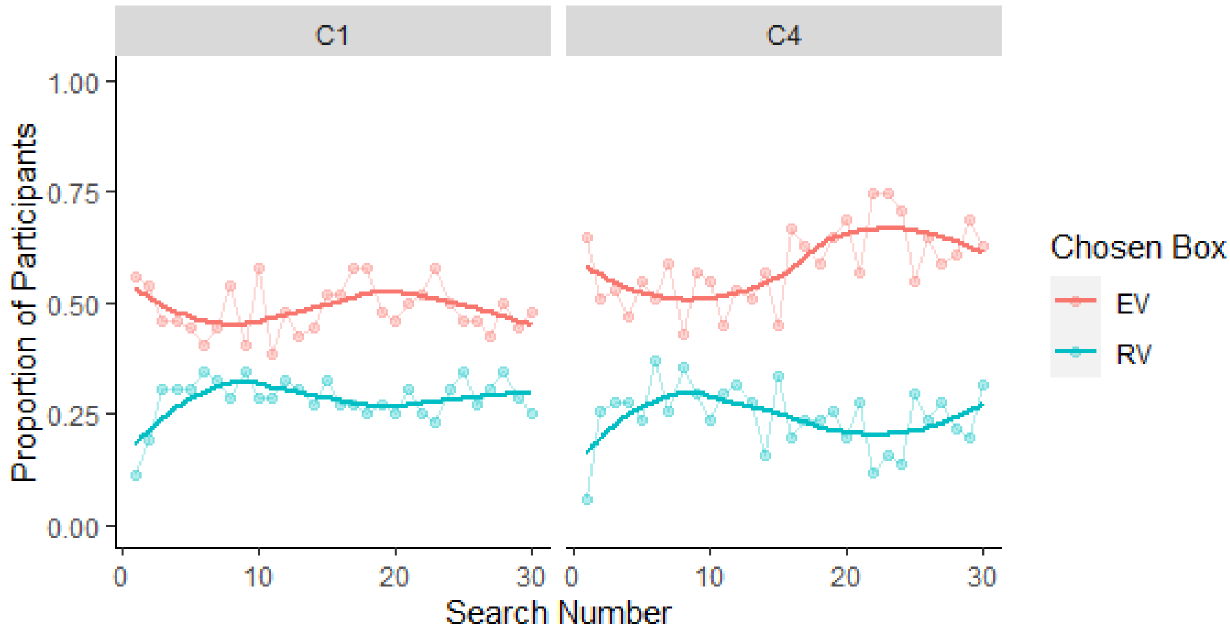


Note. Trend lines are LOESS lines.

Relationship Between Learning Based on the Maximum and Cost. As with Study 1, we again look at the relationship between increasing profits and systematic changes in maximum found and search duration for individuals. We take the individual-level coefficients from our analyses of profits, maximum found, and duration and look at their relationship. Overall, learning (in the form of increased profits) is positively correlated with learning to find higher maximum values (spearman $\rho=0.98$, $p<0.001$). On the other hand, there is not a significant correlation between changes in search duration and increasing profits ($\rho=0.13$, $p<0.199$). Finally, there is a positive correlation between duration and finding the highest maximum ($\rho=0.27$, $p=0.006$). These relationships suggest that, as in Study 1, most of the increased profits come from participants learning to find higher maxima. Part of this ability may stem from searching for longer (as suggested by the positive relationship between coefficients on maximum and duration). But the weak relationship between increased duration and profit suggests that participants are learning to find higher maxima without substantially increasing search lengths.

Search Strategies. In this section, we examine the evolution of participants' strategies as they gain experience. As a first step, we plot the percentage of participants selecting the maximum RV and EV locations for each search number (Figure 4.2.5). In our analysis, we focus on these two boxes. Overall, a visual inspection does not suggest a strong trend. It is worth pointing out, however, that participants in both conditions select the high EV box first often.

Figure 4.2.5. Choice of First Location by Search Number and Condition in Study 2



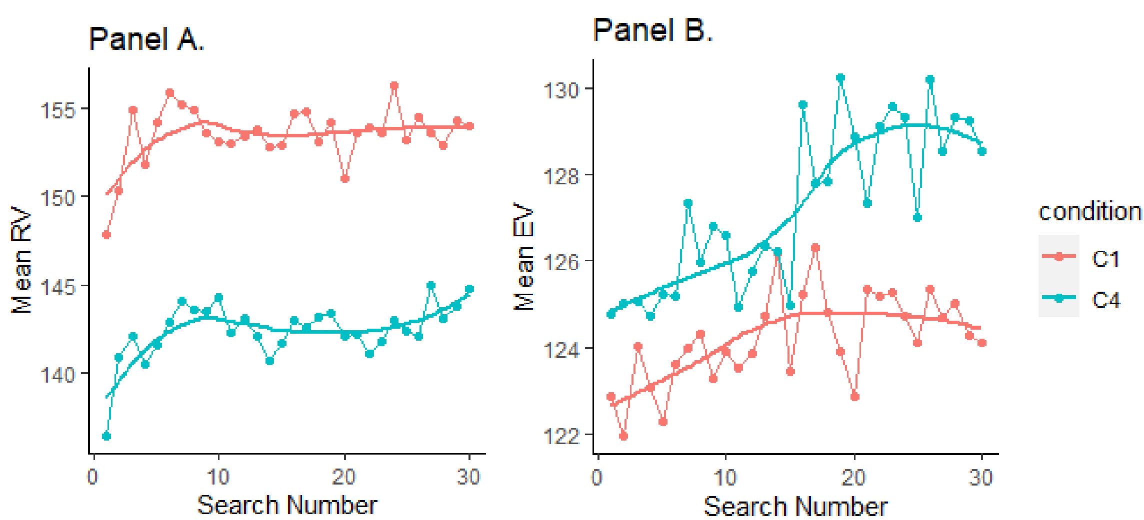
Note. Trend lines are LOESS lines. The percentage of participants choosing the high EV box are in red while those choosing the high RV box are in blue.

We run linear probability models and logistic regressions where we regress the choice of movie onto search number controlling for condition. We use cluster robust standard errors to account for the within-subjects design. We run separate models for the choice of the high EV movie and the choice of the high RV movie. We report the linear probability model here, although we also include the logistic regression in Appendix Table 6.6.5. There is no change in how often participants select the high RV movie first, b=-0.000, $t(3087)$=0.171, p=0.865. On the other hand, there is a marginally significant increase in participants selecting the high EV movie first as they gain experience, b=-0.003, $t(3087)$=1.883, p=0.060.

A visual inspection of our trends suggests that participants in the Cost=4 (vs Cost=1) condition may be more likely visit the high EV movie early. To examine this relationship, we repeat our linear probability model but add an interaction between the condition and the search number. There is a marginally significant interaction, $t(3086) = 1.801$, $p = 0.072$. When we examine the simple effects of search number on first choice of the high EV movie, there is a significant increase with experience in the high cost (Cost=4) condition, b = 0.005, t(3086) = 2.359, p =0.018, but not in the low cost (Cost=1) condition, b = 0.000, t(3086) = 0.102, p =0.919.

Even if participants are not inspecting options in the order predicted by the optimal model, they may still be visiting those options in some order. To see whether this is the case, we look at the average RV and EV of the options that participants visit during a search. First, we plot the average RV and EV against the search number in Figure 4.2.6. We then regress the average RV of each search onto the search number with a fixed effect for condition and search length with clustered standard errors. We repeat this regression with average EV.

Figure 4.2.6. Mean RV (Panel A) and Mean EV (Panel B) by Search Number in Study 2



Note. Trend lines are LOESS lines.

The trend for average RV is best captured by a model that regresses average RV onto the log of search number (BIC_{linear}=23700.1; BIC_{log}=23692.9) whereas linear and log search number models are equivalent for the average EV (BIC_{linear}=22096.4; BIC_{log}=22095.6). For consistency, we present log models for both RV and EV below. Overall, both average RV and average EV show signs of learning (Average RV: b=0.753, t(3086)=2.242, p=0.025; Average EV: b= 0.817,

$t(3086)=3.104$, $p=0.002$). A visual inspection suggests sharp increases in mean RV in the first 5 searches and consistent increases in mean EV over the first 20 searches. Unsurprisingly we see a strong fixed effect of condition on the average RV. This is expected because the RVs of all options are lower when search costs are higher. Based on a visual inspection of the plots, we also run models with interactions between log of search number and condition. When we regress average RV onto log search number with an interaction controlling for condition and search length, there is no significant interaction, $t(3085)=0.014$, $p = 0.989$. When we add an interaction to the regression of average EV onto search number controlling for condition and search length, there is also no interaction, $t(3085)=1.106$, $p = 0.269$. We present these regressions in Appendix Table 6.6.6.

Spatial Strategies. In Study 1 in Chapter 3, we saw that when there are many locations, participants are more likely to engage in spatial strategies. Participants were more likely to scan from top to bottom. In this section, we look at whether this type of behavior changes as participants gain experience and whether search costs impact the prevalence of spatial strategies.

Overall, when looking for evidence of spatial strategies, we can divide the evidence into positional evidence and transitional evidence. Positional evidence involves individual choices that participants make as they search. For example, if participants search from top-to-bottom, we expect that the position of early inspections would be closer to the top of an array of choices than for later inspections. Transitional evidence involves looking at how participants move from one inspection to the next. For example, if participants search from top to bottom, we expect that a participant would inspect items below the previously inspected item. If a participant inspects the 5^{th} item from the top first, we expect them to subsequently inspect an option below the 5^{th} item. Transitional evidence was weak an ambiguous, so we do not report these analyses.

First, to understand how common spatial strategies are, we took the Spearman correlation between the vertical position of choices on the screen and the order of the item that a participant clicked on for each search number. For example, if a participant clicked on the 5th option from the top followed by the 9th box from the top, we would correlate [5, 9] with [1, 2]. A positive correlation is an indication that participants are moving from top to bottom as they search. When we look at the average of these correlations for each search number, we see that in the Cost=1 condition, there was an average correlation of 0.108 which is significantly greater than 0, $t(29)=5.71$, $p<0.001$. These correlations are similar to those we saw in Study 1 in Chapter 3, when there was a cost of 1 and 12 boxes. In the Cost=4 condition, there was a correlation of 0.025, which is not significantly different from 0, $t(29)=1.189$, $p=0.244$.

If participants are searching from top to bottom, we would also expect that participant's first inspections would be more likely to be towards the top of the screen. For example, a participant's first inspection is more likely to be the 1st option than the 12th. To see whether this pattern exists, we look at the percentage of times that each participant's first inspection was among the top half (i.e., top 6) of the options. Overall, participants inspected one of the top 6 boxes first 54.2% of the time, significantly more than 50%, $t(102)=3.255$, $p=0.002$. When we compare conditions, we find that participants in the low cost condition were more likely to select among the top half of boxes first than those in the high cost condition, $M_{c1}=0.580$, $M_{c4}=0.502$, $t(96.65)=3.158$, $p=0.002$.

An additional piece of positional evidence we can examine is the vertical position of choices. In general, if participants are searching from top to bottom, we should expect that the average vertical position of their first choices should be higher than that of a participant choosing to inspect items based purely on potential value. Overall, participants' first choices had an average vertical position of 6.08, which was significantly above the midpoint, 6.5, $t(102)=4.602$, $p<0.001$. When we separate vertical position of the first item by cost condition, participants in the low cost condition have significantly lower (i.e., higher up the screen) average first inspections than those in the high cost condition, $t(98.62)=3.140$, $p=0.002$. When we compare the average vertical position of first choices for each condition against the midpoint of the screen, participants in the low cost condition had mean vertical positions significantly higher up the screen than the center, $M=5.82$, $t(51)=5.184$, $p<0.001$. On the other hand, participants in the high cost condition did not have mean vertical positions different from the midpoint, $M=6.36$, $t(50)=1.265$, $p=0.212$. The positional measures suggest there is a slight preference for options in the top half of the vertical display only for low cost (Cost = 1) searches. Across our measures, we did not find any evidence that participants changed their use of spatial strategies as they gained experience.

4.2.3. Discussion

In Study 2, we increased the number of movies that participants could inspect in each search. Overall, participants learned only modestly as they gained experience over 30 searches. Most of this learning happens in early searches, where participants are getting acquainted with the task. In both high and low search cost conditions, we also did not find substantial changes in strategy use, with participants generally preferring to look at high expected value options rather than high reservation value options. One notable trend was that participants in low cost conditions were more likely to rely on a “top-to-bottom” spatial strategy than participants in high cost conditions. When search costs were low (vs. high), participants were more likely to inspect items in the upper part of the screen early in their search and to inspect items below previously inspected items. And, this tendency increased as participants gained experience.

4.3. General Discussion

The two studies reported in Chapter 4 are an initial exploration of learning to search. Participants repeated similar 5- and 12-option searches over 30 trials. They received simple diagnostic feedback by observing their overall outcomes at the end of each search. These conditions did not produce dramatic learning effects. In most cases, there appeared to be an initial "getting to know the task" improvement in performance over the first five to ten searches, but little apparent learning beyond the early familiarization phase (in both studies participants averaged 96% of the optimal model's profits).

To the degree that there was learning, participants increased the maximum values that they found without substantially increasing search durations. We manipulated number of options in the search sets, intending to increase task difficulty. We thought increased difficulty might extend the learning process across more trials and tell us more about changes in strategies that could be attributed to learning. Ultimately, we were disappointed in the informativeness of both studies on the topic of learning to search.

Detailed descriptive statistics reinforced our earlier conclusion that reservation value and expected value-based navigation strategies predominated. But, the modest amount of improvement in performance observed in these studies did not suggest that the average participants shifted toward the optimal RV strategy over the heuristic EV strategy. We did, however, find that the participants who improved their performance the most increasingly visited high RV locations. Finally, we found suggestive evidence that participants rely more on spatial strategies when search costs are low (vs. high), although we did not find any significant changes with experience.

In short, the studies reported in Chapter 4 were a disappointment to us. But these studies can help us generate ideas for more discerning experiments. For example, in Study 2, there is some suggestion that participants in the high-cost condition continued to learn throughout the task. Thus, it is possible that we did not provide participants with the kind of learning experience that could best improve performance. Increasing the number of learning trials may reveal more learning. And it is possible that the type of feedback might matter. For example, providing participants with a maximum performance benchmark (like the performance of the optimal model) might encourage them to experiment more and facilitate learning.

5. Chapter 5: General Discussion

Across 7 studies, we have examined how consumers search. In Chapter 2, we report three studies focused on value-based search. In Study 1, the average participants' stated reservation values roughly conformed to predictions of the optimal model. However, there was also substantial heterogeneity, with many participants using a strategy best described by heuristics based on expected value rather than reservation value. In Study 2, participants engaged in a simple search task that required navigation. We again observed heterogeneity, with some participants best described as navigating based on the reservation value and others based on the expected value. Participants using both navigation strategies achieved profits only slightly lower than the maximum achieved by the optimal model. However, participants using different search strategies often ended up with different products. In Study 3, we replicated the findings of Study 2 using graphical rather than numerical representations of possible outcomes for each location.

In Chapter 3, we focused on how complexity impacts the use of strategies. In Study 1, we examined the impact of increasing the number of products participants could inspect. When more products were available, participants were more likely to search from top to bottom. Interestingly, the percentage of searches best described by the optimal strategy, while modest, was not impacted by the number of available options. On the other hand, the percentage of searches best described by the expected value heuristic declined as the number of products increased. In Study 2 we manipulated both the spatial layout of locations (i.e., linear or circular) and whether cues to value were presented graphically or numerically. Participants' search strategies depended on the difficulty of accessing cues. When cues to value were difficult to access due to spatial layout (circular) and the data were presented (graphically), the use of the RV strategy declined while the reliance on spatially-based strategies increased. When data were presented numerically, spatial layout had little impact because numerical information comparisons are similar in difficulty for both spatial layouts.

Finally, in Chapter 4, we conducted an initial exploration of learning to search with simple diagnostic feedback. In Study 1, participants repeated similar 5-box searches for 30 trials. We did not find dramatic learning effects. In general, participants' profits increased in the first 5 to 10 trials as they grew acquainted with the task, before leveling off over the remaining 20 trials. This increase in profits was the result of finding higher maxima without increasing the number of searches. Participants who improved the most tended to inspect locations with higher reservation values. In Study 2, we increased the complexity of the task by having participants repeat similar 12-box searches for 30 trials. As in Study 1, participants learned in early trials, but there were no systematic changes in performance afterward. In Study 2, we also manipulated search costs and found that participants were more likely to rely on a "top-to-bottom" spatial strategy when search costs were low (versus high).

5.1.1. How do consumers' search behaviors compare to the predictions of optimal models?

Across all 7 of our studies, we compared participants' behaviors to predictions made by the risk-neutral optimal model. While perfect adherence to the optimal model was rare, many participants did navigate or stop according to its principles, particularly when search environments were simple and the cues to value were easily accessible. In Chapter 2, most participants reduced stated reservation values in response to rising search costs and roughly a third of participants' stopping rules were best described by the optimal model. In the simple navigation situations in Studies 2 and 3 of Chapter 2, between one-half (with numerical stimuli) and one-third (with graphical stimuli) of searches had navigation paths consistent with the optimal model. In Study 1 of Chapter 3, where we manipulated the number of available options with graphical stimuli, we found similar results. Interestingly, the percentage of searches consistent with the optimal model was not sensitive to the number of options. Our explanation for this invariance, despite seemingly higher complexity, is that the graphical conditions make cues to value highly accessible since participants can take advantage of rapid, parallel visual comparisons to detect the high reservation value option. In Study 2 of Chapter 3, when we make these visual comparisons difficult, we find substantially less adherence to the optimal model. Finally, in Chapter 4, in which we focused on learning, about a third of searches fit the optimal model. Crucially, though, we find little sign of learning beyond the first few searches.

5.1.2. When consumers are not searching optimally, what are they doing instead?

When consumers were not searching optimally, they relied both on alternative value-based heuristics and spatial heuristics. In Chapter 2, which focused on value-based strategies, non-optimal participants were best described by strategies based on the expected value. In contrast to the reservation value strategy, which (with sufficiently low costs) places more weight on the top end of a distribution of possible outcomes, an expected value strategy places equal weight on all parts of a distribution of possible outcomes. In Study 1 of Chapter 2 roughly half of participants' stated reservation values were best described by a heuristic based on the expected value. In Studies 2 and 3 of Chapter 2, between 20% and 40% of navigation paths were consistent with an expected value-based heuristic. Searches best described by an expected value model perform only slightly worse than those navigating (but not necessarily stopping) according to the optimal model. Crucially, however, participants navigating according to the expected value often end up "purchasing" products from different locations than those using the optimal model. The results of Chapters 3 and 4 largely replicate findings around the use of expected value-based heuristics. Importantly, in Chapter 4, participants who perform relatively poorly tend to increasingly visit higher expected value items as they gain experience while those who perform best do not.

In Chapter 3, we find that some participants use spatial heuristics. For example, as we increase the number of locations in Study 1, participants are more likely to search from top-to-bottom rather than using a value-based criterion, either the reservation value or expected value. Similarly in Study 2, participants tend to inspect options clockwise when cues to value are not easily accessible.

5.1.3. When are consumers likely to use different strategies?

Our overall conclusion is that as the cognitive effort required to execute a strategy increases, participants will shift to cognitively simpler strategies. However, the cognitive effort demanded by a strategy can be non-obvious. In Study 1 of Chapter 3, we increase complexity by adding locations to our graphically formatted choice set. Roughly a quarter of participants used navigated consistently with the optimal strategy in each of our conditions. We did not find participants shifting away from the optimal strategy. Instead, the use of the expected value heuristic declined. This is because participants must estimate the expected value of each location and hold those estimates in memory, if they attempt to rely on an expected value heuristc. Participants following a reservation value strategy can take advantage of the fast, precise, and parallel visual system to pick out locations *with the highest maximum* (which mimics the reservation value strategy in our stimuli). Stated another way, our graphical stimulus displays made it easy to pick the high reservation value item regardless of the number of options. This result suggests that if we make the information needed to infer reservation values less accessible, participants would shift to other cues and strategies.

In Study 2 of Chapter 3, we presented information about locations both graphically and numerically while varying their spatial layout. When stimuli were presented graphically, more searches were described by expected value-based heuristics (and fewer by the optimal reservation value strategy) than when they were presented numerically, especially when locations were displayed in a linear layout. Our interpretation of these results is that it was easier for participants to use expected values to navigate when they could do so visually, particularly when a linear layout facilitated comparisons between locations. In the numerical condition, instead participants had to calculate an average (expected value) based on two numbers. Importantly, in the same study, fewer searches used the reservation value strategy when we made the reservation values difficult to compare by changing the visual layout of our stimuli. When distributions of outcomes were presented graphically, searches were less likely to be described by the reservation value strategy when the layout was circular rather than linear. In numerical conditions, the layout did not impact the use of the optimal strategy. Our interpretation of these results is that the circular condition made comparisons more difficult in the graphical conditions, where participants had to rely more heavily on visual comparisons but did not change the difficulty of comparisons in the numerical conditions, where comparisons always took place by comparing numbers in memory.

Finally, consumers switched to spatial strategies when it was too difficult to use value-based strategies. For example, in Study 1 of Chapter 3, the use of the expected value heuristic declines as the number of locations increases and the use of a "top-to-bottom" spatial strategy increases. And in Study 2 of Chapter 3, when we make comparisons between locations difficult (by presenting distributions graphically and displaying locations as a circle), we see fewer searches relying on the optimal strategy and more searches relying on a "clockwise" spatial strategy. Finally, in Study 2 of Chapter 4, we find suggestive evidence that participants are more likely to use spatial strategies when search costs are low (vs. high). Overall, we conclude that consumers switch to spatial strategies when value-based strategies are not deemed to be worth the additional effort. In Chapter 3, participants switched to spatial strategies as we made determining the value of a location or comparing locations more difficult. We interpret the results in Chapter 4 as participants deciding that using the inefficient, but low effort spatial strategy is too "expensive" in the high-cost condition, but not in the low-cost condition.

5.2. Limitations and Scope of Research

Our research has several important limitations which make us qualify our advice about specific marketing applications. Most importantly, our research uses stimuli that minimize the complexity of a typical consumer experience to isolate the essential factors driving search behaviors. For example, we present participants with well-defined distributions of possible outcomes where participants in most consumer situations would learn them from unsystematic samples of their shopping experiences. We also collapse the value of a product into a single number even though in most shopping situations consumers integrate multiple cues to infer value (e.g., price and quality-relevant attributes). Given these differences between our controlled consumer search tasks and realistically complex search tasks, we can make only suggestive claims about how our results will apply to non-laboratory consumer markets.

Another key limitation is that we do not have fine-grained process data on how our participants are searching. We observed participants' navigation and stopping decisions but not what locations participants were thinking about visiting. In Chapter 2, when we categorize people as using optimal or expected value-based strategies, we infer their plausible strategies from their behaviors. But it is always possible, that participants are navigating based on different cognitive operations. Fine-grained process data would also be useful for identifying spatial strategies. In the experiments reported in Chapter 3, for example, eye-tracking measures could provide more evidence about whether participants are scanning locations from top to bottom (or clockwise) and making navigation decisions without considering some of the locations at all, violating any concept of optimal search. And in Chapter 4, process data might reveal subtle changes in attention over time that our current measures miss. More generally, process data would show us whether participants are making a limited number of pairwise comparisons rather than holding the value of visited locations in memory and making multiple comparisons for each alternative.

To shed some light on process, we had a small sample (n = 22) of participants "think-aloud" as they searched using our five-location, graphical task (as in Chapter 2, Study 3). We observed that participants typically rely on a series of pair-wise comparisons. Furthermore, when deciding whether to visit locations, participants often either focus on the expected value and shade their valuation of the location upward or downward based on the range of possible outcomes or focus on the maximum possible value and shade down based on the range. Future studies could build upon these think-alouds by collecting process data using a more advanced interface (like those provided in Mouselab; Johnson et al., 1989) or eye-tracking. Adding these types of data could provide more evidence with which to evaluate our claims about participants' strategies.

Here are a few suggestive implications of our findings for marketing practice. A key finding from Chapter 1 is that participants employ different search strategies, and when they do so, end up purchasing products at different locations depending on the strategy they are using. More precise identification of consumers' search styles will allow marketers to predict where consumers will purchase their products. For example, consumers who navigate using expected value are likely to be attracted to everyday low price (EDLP) retailers while those using the optimal strategy are likely to visit Hi-Lo retailers early in their searches.

Modern store shelves are often packed with products. For example, consumers have dozens of choices of toothpaste or pasta sauce. These results suggest that people will search through these alternatives in different ways depending on how many alternatives are presented and on how easy it is to access information about those alternatives. When the number of alternatives is small, participants are likely to navigate based on value rather than spatial features, but when the number of options increases, navigation based on spatial layout increases. Furthermore, navigation based on spatial locations increases when it is difficult to make comparisons between alternatives. Visually salient attributes (for example in a table display of shirts) or easily comparable attributes that can be used to predict how much a product will be liked (e.g., televisions have a few well-defined features), are likely to promote value-based navigation. On the other hand, difficult to access information (for example in a rack of shirts, where individual product attributes are hidden by the display format) or for products without easily comparable benefits or salient cues to value (e.g., different brands of vitamin E) may result in increased reliance on spatial strategies.

5.3. Conclusion

Our results extend the existing empirical literature on consumer search by exploring optimal and heuristic navigation strategies more thoroughly than any prior research program. An empirical literature has tested some general implications of these models using consumer surveys (e.g., Moorthy et al., 1997) and market data (e.g., Honka & Chintagunta, 2017; Ursu, 2018). Highly controlled laboratory studies, like the ones described in the present dissertation, complement this literature by supporting stronger causal conclusions and providing more detail on the cognitive processes that underlie observable behaviors.

A key implication of the present research is that consumers use reasonable strategies while searching, particularly when cognitive costs are accounted for. Given a complex search task, consumers will use adaptive heuristics to simplify the task. Often these heuristics involve integrating the available information using simpler calculations than those required by the optimal model. Depending on the complexity of the search, consumers may also reduce cognitive costs associated with holding information in memory by taking advantage of the spatial structure of the environment.

More broadly, we suggest that consumers trade-off between different types of search strategies depending on the costs and benefits of each, including cognitive information-processing costs. When it is relatively easy or beneficial to do so, consumers will rely on value-based strategies that guide them to high-value locations. The optimal strategy (based on the reservation value) is the most effective of these strategies but is also cognitively effortful. Therefore, consumers employ simpler value-based strategies adapted to available cues.

When value-based strategies are difficult to employ, either due to lack of information or to the complexity of the task, consumers will shift to alternative strategies to navigate. One family of these strategies uses proximity to guide navigation. Certain spatial layouts (i.e., linear or clustered) allow consumers to reduce cognitive costs associated with holding information in working memory. Consumers use strategies that take advantage of these spatial layouts when they deem the savings in cognitive costs outweigh the loss of effectiveness. Proximity in memory may also form the basis for navigation when cues in the environment are missing. So, for example, consumers may evaluate options based on similarity, a non-spatial form of proximity.

Another family of alternative search strategies relies on social information. Rather than expending effort in forming their own evaluations, consumers will often rely on others' evaluations. In some online settings, consumers may ignore other cues-to-value and rely completely on individual reviews or summaries of other consumers' evaluations. Similarly, they may base navigation decisions on simple popularity, believing that popularity is correlated with value. And sometimes consumers rely on role models' tastes to guide them (e.g., celebrities, influencers, friends), especially for products that carry significant personal identity or status signals.

Search is a central part of the modern consumer experience. We conclude that consumers are adaptive and guided by rational principles; they balance the costs of expending effort, both physical and cognitive, against the benefits of outcomes achieved from various search strategies. They skillfully use the structure their environments and their cognitive resources to find good products. The ultimate purpose of searching for products is to use those products. By searching effectively consumers can find good products and services without searching forever.

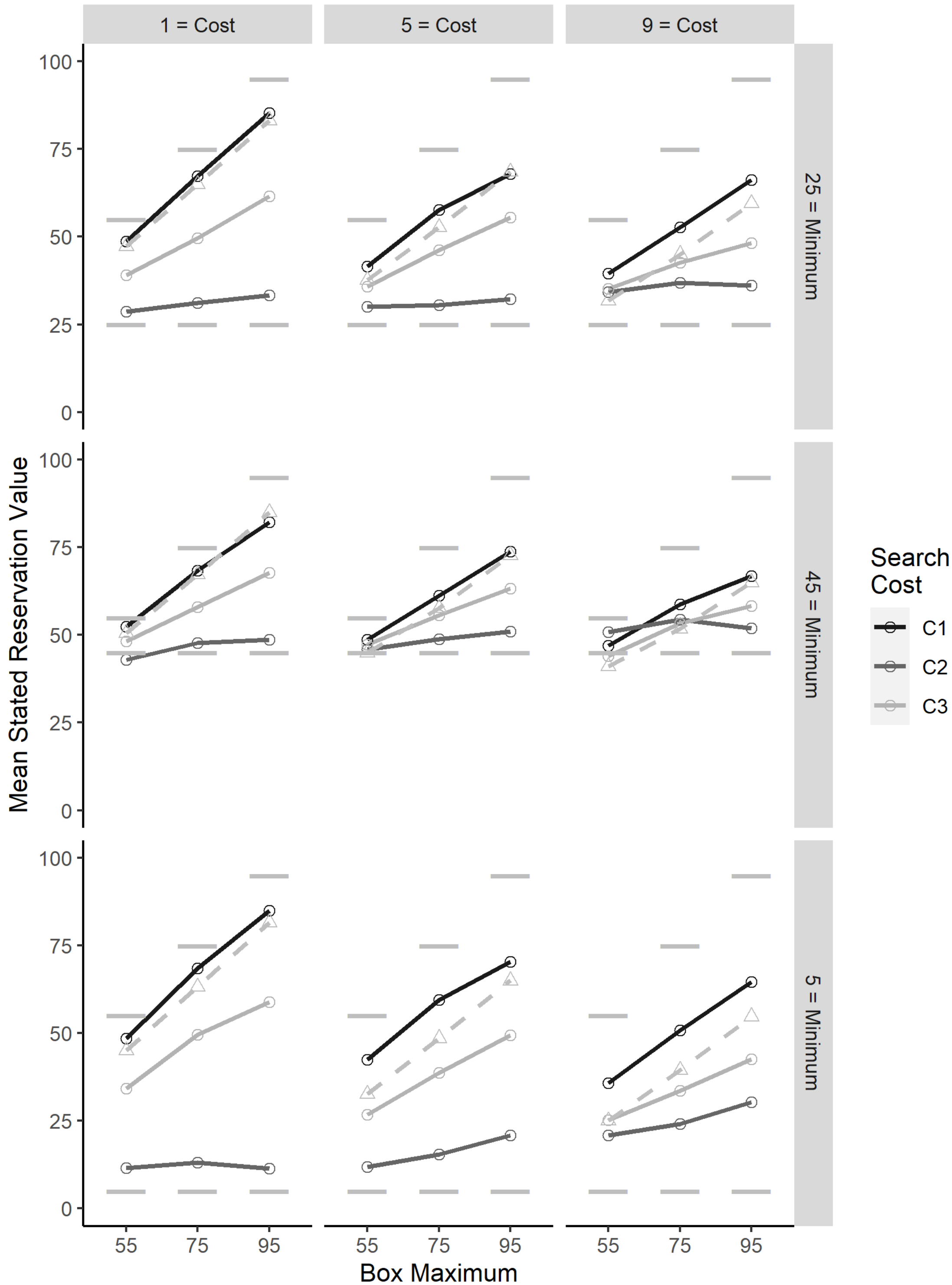
6. Appendix

6.1. Supplementary Materials for Chapter 2, Study 1

Table 6.1.1 Study 1 stimuli and average stated RVs (reservation values)

Cost	Min	Max	RV	N	Mean	Median	SD
1	5	55	45.00	100	36.07	38.50	14.45
1	5	75	63.15	100	51.23	55.00	19.47
1	5	95	81.58	100	61.50	60.00	28.13
1	25	55	47.27	100	40.92	40.00	9.31
1	25	75	65.00	100	53.07	51.00	15.02
1	25	95	83.18	100	65.75	65.50	19.88
1	45	55	50.53	100	48.89	50.00	5.12
1	45	75	67.25	100	60.03	60.00	10.03
1	45	95	85.00	100	70.10	70.00	16.28
5	5	55	32.64	100	30.00	30.00	13.19
5	5	75	48.54	100	42.61	44.50	19.04
5	5	95	65.00	100	52.70	52.50	23.25
5	25	55	37.68	100	36.99	35.00	8.10
5	25	75	52.64	100	47.96	49.50	12.87
5	25	95	68.54	100	56.69	55.50	17.47
5	45	55	45.00	100	47.48	47.00	3.14
5	45	75	57.68	100	56.63	55.00	7.58
5	45	95	72.65	100	65.14	65.00	12.74
9	5	55	25.00	100	28.02	28.00	14.02
9	5	75	39.51	100	38.01	40.00	17.79
9	5	95	54.75	100	48.15	50.00	23.22
9	25	55	31.76	100	36.41	35.00	9.03
9	25	75	45.00	100	45.12	41.00	12.84
9	25	95	59.50	100	52.51	50.00	18.33
9	45	55	41.00	100	45.77	45.00	8.99
9	45	75	51.76	100	55.16	55.00	8.55
9	45	95	65.00	100	60.23	60.00	13.96

Figure 6.1.1. Mean Stated RVs By Cluster in Study 1



Note. Cluster 1 contained 32% of participants, Cluster 2 contained 12%, and Cluster 3 contained 56%. Dashed gray lines with hollow triangles are risk-neutral optimal responses.

6.2. Supplementary Materials for Chapter 2, Study 2 and Study 3

Table 6.2.1. Study 2 and 3 Stimuli

		Bounds of Uniform Distribution				Search Orders	
Stimulus Set	Box Name	Lower	Upper	RV	EV	RV	EV
1	A	55	135	106.8	95.0	2.5	2.5
	B	79	111	93.1	95.0	4	2.5
	C	76	130	106.8	103.0	2.5	1
	D	40	146	113.4	93.0	1	4
	E	44	118	90.8	81.0	5	5
2	A	45	146	114.2	95.5	2.5	2.5
	B	81	110	93.0	95.5	4	2.5
	C	98	133	114.2	115.5	2.5	1
	D	28	157	121.0	92.5	1	4
	E	40	115	87.6	77.5	5	5
3	A	36	121	91.8	78.5	2.5	2.5
	B	44	113	86.7	78.5	4	2.5
	C	54	117	91.8	85.5	2.5	1
	D	26	126	94.4	76.0	1	4
	E	30	105	77.6	67.5	5	5
4	A	52	164	130.5	108.0	2.5	2.5
	B	91	125	106.6	108.0	4	2.5
	C	109	151	130.5	130.0	2.5	1
	D	32	181	142.4	106.5	1	4
	E	46	132	102.7	89.0	5	5
5	A	41	117	89.4	79.0	2.5	2.5
	B	49	109	84.5	79.0	4	2.5
	C	57	113	89.4	85.0	2.5	1
	D	32	124	93.7	78.0	1	4
	E	34	111	83.3	72.5	5	5
6	A	42	122	93.7	82.0	2.5	2.5
	B	66	98	80.1	82.0	4	2.5
	C	63	117	93.7	90.0	2.5	1
	D	27	133	100.4	80.0	1	4
	E	31	105	77.8	68.0	5	5

Note. In search orders, ties are denoted by halves. For example, 2.5 means that this box is tied for second

6.3. Supplementary Materials for Chapter 3, Study 1

Table 6.3.1. Stimuli in Studies 1 and 2

	1				2			
	MIN	MAX	RV	EV	MIN	MAX	RV	EV
Base Items	80	186	**171.4**	133.0	68	197	**180.9**	132.5
	95	178	*165.1*	*136.5*	85	186	*171.7*	*135.5*
	135	170	161.6	**152.5**	141	173	165.0	**157.0**
	84	158	145.8	121.0	80	155	142.8	117.5
	117	141	134.1	129.0	121	147	139.8	134.0
	85	141	130.4	113.0	95	150	139.5	122.5
9 Item	120	140	133.7	130.0	114	144	136.3	129.0
	90	138	128.2	114.0	129	141	136.1	135.0
	110	142	134.0	126.0	55	153	139.0	104.0
12 Item	115	141	133.9	128.0	126	145	138.8	*135.5*
	95	143	133.3	119.0	97	146	136.1	121.5
	65	139	126.8	102.0	68	153	140.0	110.5

	3				4			
	MIN	MAX	RV	EV	MIN	MAX	RV	EV
Base Items	66	166	**151.8**	116.0	42	191	**173.8**	116.5
	76	161	*147.6*	118.5	62	174	*159.0*	*118.0*
	130	152	145.4	**141.0**	119	161	151.9	**140.0**
	84	155	143.1	119.5	56	142	128.9	99.0
	78	140	128.9	109.0	101	124	117.2	112.5
	122	130	126.0	126.0	33	129	115.2	81.0
9 Item	97	134	125.9	115.5	111	121	116.5	116.0
	51	142	128.5	96.5	83	125	115.8	104.0
	113	133	126.7	123.0	61	126	114.6	93.5
12 Item	67	140	127.9	103.5	21	133	118.0	77.0
	82	139	128.3	110.5	55	126	114.1	90.5
	106	135	127.4	120.5	97	121	114.9	109.0

Table 6.3.1. Stimuli in Studies 1 and 2 continued

	5				6			
	MIN	MAX	RV	EV	MIN	MAX	RV	EV
Base Items	73	169	**155.1**	121.0	67	173	**158.4**	120.0
	132	159	*151.6*	**145.5**	82	165	*152.1*	123.5
	97	155	144.2	126.0	103	157	146.6	*130.0*
	120	149	141.4	*134.5*	130	145	139.5	**137.5**
	90	141	130.9	115.5	108	137	129.4	122.5
	117	136	129.8	126.5	85	139	128.6	112.0
9 Item	55	144	130.7	99.5	118	132	126.7	125.0
	87	140	129.7	113.5	55	142	128.8	98.5
	104	138	129.8	121.0	89	137	127.2	113.0
12 Item	120	134	128.7	127.0	63	139	126.7	101.0
	97	140	130.7	118.5	94	137	127.7	115.5
	59	144	131.0	101.5	115	135	128.7	125.0

Note. Bolded numbers indicate the highest value based on the RV or EV strategy while italicized items indicate the second highest value.

Table 6.3.2. Individual Level Comparison of First Inspection by Position to Random Choices

Number of Locations	Vertical Position	Mean	*t(49)*	*p*
6	1	15.0%	-0.982	0.331
	2	14.3%	-1.658	0.104
	3	19.5%	1.456	0.152
	4	**13.5%**	**-2.435**	**0.019**
	5	17.5%	0.579	0.565
	6	*20.2%*	*1.979*	*0.053*
9	*1*	*14.2%*	*1.902*	*0.063*
	2	12.2%	0.859	0.395
	3	11.5%	0.341	0.734
	4	11.5%	0.334	0.740
	5	10.0%	-0.807	0.424
	6	10.3%	-0.598	0.553
	7	9.0%	-1.950	0.057
	8	10.5%	-0.469	0.641
	9	10.8%	-0.215	0.830
12	**1**	**15.5%**	**2.617**	**0.012**
	2	10.2%	1.224	0.227
	3	8.7%	0.220	0.827
	4	8.0%	-0.320	0.751
	5	8.0%	-0.270	0.788
	6	7.0%	-1.391	0.170
	7	**6.3%**	**-2.209**	**0.032**
	8	7.7%	-0.546	0.588
	9	7.3%	-1.145	0.258
	10	7.7%	-0.501	0.618
	11	6.8%	-1.328	0.190
	12	6.8%	-1.549	0.128

Note. Positions significantly greater than chance at the $p<0.05$ level are bolded while those at the $p<0.10$ level are italicized.

Figure 6.3.1. Log of Ratio of Up vs Down Choices by Condition and Vertical Position after the First Inspection in Study 1

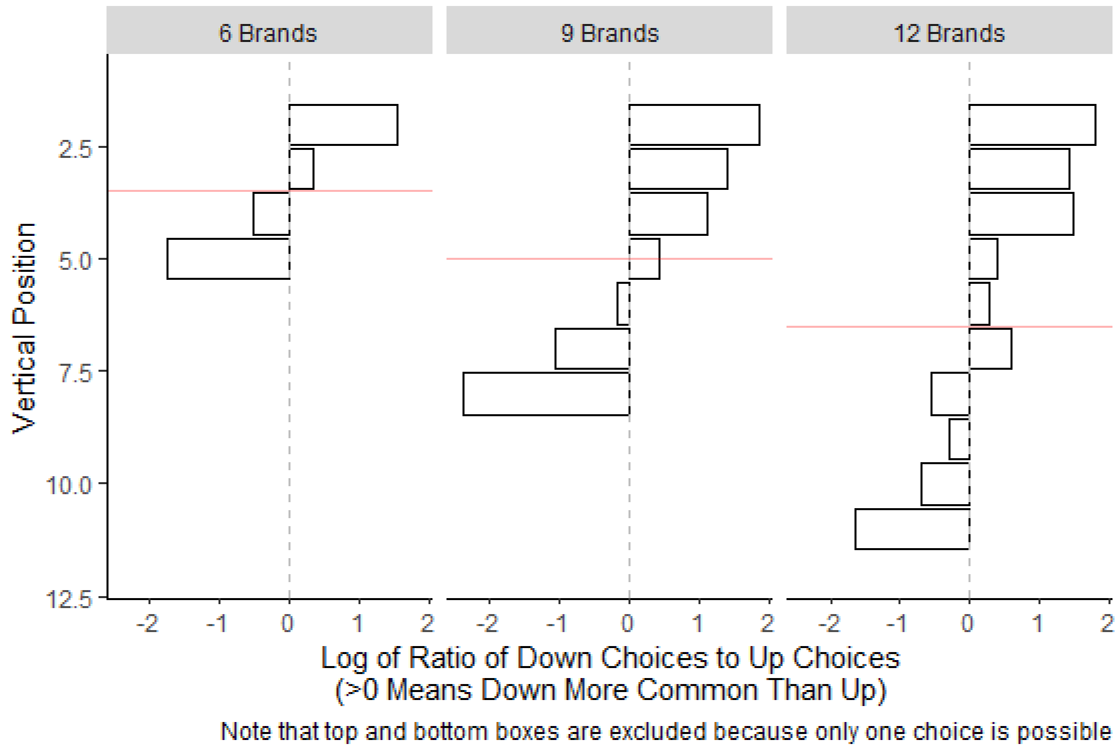


Table 6.3.3. Models of Up and Down Moves on Number of Boxes

	Linear Probability Model		Logistic Model	
	All Data	Truncated Data	All Data	Truncated Data
Number of Boxes	0.063***	0.057***	0.383***	0.289***
	(0.005)	(0.006)	(0.040)	(0.038)
Vertical Position	-0.100***	-0.080***	-0.576***	-0.394***
	(0.005)	(0.006)	(0.039)	(0.037)
Second Half	-0.011	-0.017	-0.048	-0.069
	(0.019)	(0.025)	(0.115)	(0.122)
Choice Number	0.052***	0.050***	0.278***	0.250***
	(0.013)	(0.036)	(0.068)	(0.067)
Intercept	0.405***	0.382***	-0.635	-0.632
	(0.055)	(0.066)	(0.324)	(0.330)
Clustered Ses	X	X	X	X
Search FE	X	X	X	X
Obs	3051	2315	3051	2315
R^2	0.34	0.18		

Table 6.3.4.Models with Number Remaining Above and Below the Most Recently Inspected Brand

	Linear Probability Model		Logistic Model	
	All Data	Truncated Data	All Data	Truncated Data
Number of Boxes	-0.001	-0.020	-0.147*	-0.101
	(0.004)	(0.017)	(0.072)	(0.094)
Vertical Position	0.020**	0.065*	0.421**	0.326⁺
	(0.007)	(0.032)	(0.136)	(0.184)
Second Half	-0.005	-0.005	-0.035	-0.027
	(0.013)	(0.021)	(0.108)	(0.105)
Choice Number	-0.013***	-0.016	-0.071	-0.078
	(0.004)	(0.013)	(0.062)	(0.073)
Percent Below	1.084***	1.393***	8.977***	6.647***
	(0.053)	(0.294)	(1.207)	(1.175)
Intercept	-0.062	-0.247	-4.801***	-3.569***
	(0.046)	(0.191)	(0.820)	(0.785)
Clustered Ses	X	X	X	X
Search FE	X	X	X	X
Obs	3051	1915	3051	1915
R^2	0.477	0.192		

Table 6.3.5. Coefficient on Binary Variable of Double Downward Inspections (Present = 1) for Inspection Order 1 through 4

Inspection Order	Coefficient	*df*	*t*	*p*
1	0.022	487	2.135	0.033
2	0.049	206	2.618	0.010
3	0.067	114	3.005	0.003
4	0.057	60	2.517	0.015

6.4. Supplementary Materials for Chapter 3, Study 2

Table 6.4.1. Regression of (1) Duration, (2) Maximum Value, and (3) Profit on Conditions

	(1) Duration	(2) Maximum Value	(3) Profit
Graphical/Numerical	0.113	-0.851	-0.964^{+}
	(0.159)	(0.581)	(0.574)
Circular/Linear	-0.220	0.763	0.983^{+}
	(0.159)	(0.581)	(0.574)
Set Shown: S1	0.041	5.463^{***}	5.422^{***}
	(0.055)	(0.629)	(0.645)
Set Shown: S2	0.082	-14.203^{***}	-14.285^{***}
	(0.062)	(0.571)	(0.585)
Set Shown: S3	-0.003	-8.326^{***}	-8.323^{***}
	(0.059)	(0.739)	(0.758)
Set Shown: S4	-0.146^{*}	-9.463^{***}	-9.317^{***}
	(0.058)	(0.521)	(0.532)
Set Shown: S5	0.147^{*}	-13.119^{***}	-13.266^{***}
	(0.060)	(0.565)	(0.578)
Block: Late = 1	-0.125^{*}	0.271	0.395
	(0.051)	(0.397)	(0.403)
Set Presentation Order	-0.021^{+}	0.188^{+}	0.209^{+}
	(0.012)	(0.104)	(0.106)
Graphical/Numerical:Circular/Linear	-0.226	1.955^{+}	2.181^{+}
	(0.318)	(1.162)	(1.148)
Constant	3.064^{***}	155.360^{***}	152.297^{***}
	(0.097)	(0.653)	(0.657)

$^{+}p<0.1|^{*}p<0.05|^{**}p<0.01|^{***}p<0.001$

Note. The Variables for condition were coded -0.5 and 0.5 which allows the coefficients on the conditions to be interpreted roughly as average effects.

Table 6.4.2. Regression of Maximum RV first choices (Models 1 and 3) and EV first choices (Models 2 and 4) on Conditions

	(1) RV lpm	(2) EV lpm	(3) RV logit	(4) EV logit
Graphical/Numerical	-0.027	0.053^{*}	-0.129	0.236^{+}
	(0.026)	(0.027)	(0.121)	(0.122)
Circular/Linear	0.005	0.040	0.025	0.177
	(0.026)	(0.027)	(0.121)	(0.122)
Set Shown S1	-0.002	-0.030	-0.012	-0.132
	(0.020)	(0.020)	(0.097)	(0.090)
Set Shown S2	0.004	0.004	0.018	0.016
	(0.019)	(0.020)	(0.095)	(0.086)
Set Shown S3	0.097^{***}	-0.015	0.443^{***}	-0.068
	(0.019)	(0.019)	(0.085)	(0.085)
Set Shown S4	0.044^{*}	0.034^{+}	0.209^{*}	0.147^{+}
	(0.020)	(0.020)	(0.094)	(0.086)
Set Shown S5	0.034^{+}	-0.121^{***}	0.164^{+}	-0.586^{***}
	(0.019)	(0.020)	(0.092)	(0.098)
Block: Late = 1	0.011	0.009	0.054	0.043
	(0.012)	(0.013)	(0.058)	(0.058)
Set Presentation Order	-0.002	0.0004	-0.009	0.002
	(0.003)	(0.003)	(0.016)	(0.016)
Graphical/Numerical: Circular/Linear	0.041	0.053	0.193	0.229
	(0.052)	(0.053)	(0.242)	(0.243)
Constant	0.280^{***}	0.353^{***}	-0.945^{***}	-0.610^{***}
	(0.022)	(0.023)	(0.105)	(0.103)

$^{+}p<0.1 | ^{*}p<0.05 | ^{**}p<0.01 | ^{***}p<0.001$

Note. The Variables for condition were coded -0.5 and 0.5 which allows the coefficients on the conditions to be interpreted roughly as average effects.

Table 6.4.3. Comparison of First Inspection by Spatial Position to Random Choices

Graph/Num	Circ/Lin	Total	Position	N	Percent	$X^2(1)$	p-value
Graphical	Circular	1284	***1***	***147***	***11.4%***	***15.83***	***<0.001***
			2	***137***	***10.7%***	***8.81***	***0.003***
			3	***146***	***11.4%***	***15.03***	***<0.001***
			4	***168***	***13.1%***	***37.19***	***<0.001***
			5	*124*	*9.7%*	*2.74*	*0.098*
			6	92	7.2%	2.17	0.141
			7	***84***	***6.5%***	***5.20***	***0.023***
			8	***84***	***6.5%***	***5.20***	***0.023***
			9	***62***	***4.8%***	***20.25***	***<0.001***
			10	***81***	***6.3%***	***6.67***	***0.010***
			11	***71***	***5.5%***	***12.90***	***<0.001***
			12	*88*	*6.9%*	*3.52*	*0.061*
	Linear	1260	***1***	***146***	***11.6%***	***16.96***	***<0.001***
			2	***130***	***10.3%***	***6.19***	***0.013***
			3	90	7.1%	2.21	0.137
			4	108	8.6%	0.06	0.806
			5	95	7.5%	0.95	0.329
			6	107	8.5%	0.02	0.885
			7	117	9.3%	1.35	0.245
			8	118	9.4%	1.60	0.206
			9	***83***	***6.6%***	***4.84***	***0.028***
			10	100	7.9%	0.22	0.640
			11	95	7.5%	0.95	0.329
			12	***71***	***5.6%***	***11.71***	***<0.001***
Numerical	Circular	1332	1	114	8.6%	0.06	0.811
			2	97	7.3%	1.81	0.178
			3	*91*	*6.8%*	*3.77*	*0.052*
			4	***136***	***10.2%***	***5.85***	***0.016***
			5	101	7.6%	0.90	0.342
			6	111	8.3%	0.00	1.000
			7	96	7.2%	2.09	0.148
			8	100	7.5%	1.10	0.294
			9	101	7.6%	0.90	0.342
			10	116	8.7%	0.19	0.662
			11	***148***	***11.1%***	***13.02***	***<0.001***
			12	121	9.1%	0.87	0.351
	Linear	1188	***1***	***124***	***10.4%***	***6.57***	***0.010***
			2	***130***	***10.9%***	***10.19***	***0.001***
			3	104	8.8%	0.22	0.643
			4	103	8.7%	0.13	0.720
			5	97	8.2%	0.03	0.868
			6	89	7.5%	1.01	0.315
			7	***74***	***6.2%***	***6.65***	***0.010***
			8	*80*	*6.7%*	*3.80*	*0.051*
			9	93	7.8%	0.34	0.558
			10	95	8.0%	0.14	0.707
			11	101	8.5%	0.02	0.881
			12	98	8.2%	0.00	0.952

Note. Bolded values are significant at p<0.05. Italicized values significant at the p<0.1.

Table 6.4.4.Individual Level Comparison of First Inspection by Position to Random Choices

Graph/Num	Circ/Lin	df	Position	Percent	*t*	*p-value*
Graphical	Circular	106	**1**	**11.4%**	**2.30**	**0.024**
			2	**10.7%**	**2.36**	**0.020**
			3	**11.4%**	**3.31**	**0.001**
			4	**13.1%**	**4.47**	**<0.001**
			5	9.7%	1.41	0.161
			6	7.2%	-1.48	0.142
			7	**6.5%**	**-2.51**	**0.014**
			8	**6.5%**	**-2.54**	**0.013**
			9	**4.8%**	**-5.70**	**<0.001**
			10	**6.3%**	**-2.80**	**0.006**
			11	**5.5%**	**-4.30**	**0.000**
			12	**6.9%**	**-2.10**	**0.038**
	Linear	104	**1**	**11.6%**	**2.93**	**0.004**
			2	**10.3%**	**2.22**	**0.029**
			3	7.1%	-1.69	*0.094*
			4	8.6%	0.29	0.769
			5	7.5%	-1.05	0.295
			6	8.5%	0.19	0.849
			7	9.3%	1.18	0.240
			8	9.4%	1.23	0.222
			9	**6.6%**	**-2.38**	**0.019**
			10	7.9%	-0.54	0.592
			11	7.5%	-1.08	0.285
			12	**5.6%**	**-3.87**	**<0.001**
Numerical	Circular	110	1	8.6%	0.25	0.801
			2	7.3%	-1.46	0.147
			3	**6.8%**	**-2.26**	**0.026**
			4	**10.2%**	**1.99**	**0.049**
			5	7.6%	-1.01	0.315
			6	8.3%	-0.01	0.993
			7	7.2%	-1.59	0.114
			8	7.5%	-1.38	0.170
			9	7.6%	-0.94	0.351
			10	8.7%	0.46	0.646
			11	**11.1%**	**2.73**	**0.007**
			12	9.1%	0.90	0.368
	Linear	98	**1**	**10.4%**	**2.26**	**0.026**
			2	**10.9%**	**2.39**	**0.019**
			3	8.8%	0.54	0.587
			4	8.7%	0.36	0.720
			5	8.2%	-0.24	0.810
			6	7.5%	-1.23	0.223
			7	**6.2%**	**-3.20**	**0.002**
			8	6.7%	-2.05	0.043
			9	7.8%	-0.61	0.542
			10	8.0%	-0.40	0.688
			11	8.5%	0.18	0.857
			12	8.2%	-0.10	0.918

Table 6.4.5. Regression of First Choices in Top 3 Positions on Conditions

	(1) lpm	(2) logit
Graphical/Numerical	0.049^{**}	0.245^{**}
	(0.016)	(0.077)
Circular/Linear	0.015	0.090
	(0.016)	(0.077)
Set Shown S1	0.032	0.146
	(0.022)	(0.102)
Set Shown S2	-0.050^{*}	-0.245^{*}
	(0.022)	(0.107)
Set Shown S3	-0.058^{**}	-0.288^{*}
	(0.022)	(0.112)
Set Shown S4	-0.028	-0.137
	(0.022)	(0.105)
Set Shown S5	-0.047^{*}	-0.232^{*}
	(0.021)	(0.102)
Block: Late = 1	0.005	0.023
	(0.013)	(0.064)
Set Presentation Order	0.001	0.004
	(0.004)	(0.018)
Graphical/Numerical: Circular/Linear	-0.119^{***}	-0.596^{***}
	(0.031)	(0.153)
Constant	0.308^{***}	-0.818^{***}
	(0.022)	(0.105)

$^{+}p<0.1|^{*}p<0.05|^{**}p<0.01|^{***}p<0.001$

Note. The Variables for condition were coded -0.5 and 0.5 which allows the coefficients on the conditions to be interpreted roughly as average effects.

Table 6.4.6. Regression of Average RV of Searches (Model 1) and Average EV of Searches (Model 2) on Conditions

	(1) Average RV	(2) Average EV
Graphical/Numerical	0.118	0.283
	(0.587)	(0.436)
Circular/Linear	0.804	0.283
	(0.589)	(0.438)
Number of Boxes Opened	-2.094***	-1.906***
	(0.089)	(0.106)
Set Shown S1	5.262***	1.713***
	(0.329)	(0.286)
Set Shown S2	-15.652***	-12.921***
	(0.286)	(0.319)
Set Shown S3	-5.341***	-16.165***
	(0.361)	(0.303)
Set Shown S4	-12.022***	-6.584***
	(0.292)	(0.294)
Set Shown S5	-13.197***	-12.456***
	(0.293)	(0.271)
Block: Late = 1	0.667**	0.095
	(0.207)	(0.195)
Set Presentation Order	0.155**	0.048
	(0.055)	(0.053)
Graphical/Numerical: Circular/Linear	2.078^{+}	1.059
	(1.174)	(0.871)
Constant	164.177***	143.569***
	(0.561)	(0.530)

$^{+}p<0.1$ | $^{*}p<0.05$ | $^{**}p<0.01$ | $^{***}p<0.001$

Note. The Variables for condition were coded -0.5 and 0.5 which allows the coefficients on the conditions to be interpreted roughly as average effects.

Table 6.4.7. Regression of Up vs Down Moves onto Graphical/Numerical Conditions in Linear Conditions

	lpm	logit
Condition: Numerical	0.009	0.042
	(0.019)	(0.098)
Order	-0.002	-0.004
	(0.007)	(0.043)
Percent Locations Below	2.296^{***}	12.590^{***}
	(0.273)	(2.076)
Vertical Position	0.134^{***}	0.788^{***}
	(0.024)	(0.181)
Set.ShownS1	0.019	0.119
	(0.027)	(0.147)
Set.ShownS2	-0.057^{*}	-0.286^{*}
	(0.027)	(0.141)
Set.ShownS3	-0.028	-0.137
	(0.027)	(0.141)
Set.ShownS4	-0.036	-0.176
	(0.028)	(0.142)
Set.ShownS5	-0.028	-0.134
	(0.025)	(0.131)
Last Half	-0.016	-0.090
	(0.016)	(0.082)
Order of Presentation	0.002	0.010
	(0.005)	(0.024)
Constant	-1.463^{***}	-11.148^{***}
	(0.297)	(2.226)
Note:	$^{+}p<0.1$ | $^{*}p<0.05$ | $^{**}p<0.01$ | $^{***}p<0.001$	

Table 6.4.8. Regression of Up vs Down Moves onto Graphical/Numerical Conditions in Linear Conditions by Screen Position

			Graphical			Numerical			Difference		
Inspections Completed	Remaining Below	df	*b*	*t*	*p*	*b*	*t*	*p*	*b*	*t*	*p*
1	1	160	0.01	0.39	0.70	0.04	1.07	0.29	0.03	0.52	0.60
	2	148	0.06	1.34	0.18	0.00	-0.07	0.94	-0.07	-0.95	0.34
	3	134	0.08	1.22	0.23	0.04	0.71	0.48	-0.03	-0.40	0.69
	4	161	0.05	1.06	0.29	0.08	1.13	0.26	0.03	0.33	0.74
	5	154	0.03	0.57	0.57	**0.17**	**2.57**	**0.01**	*0.14*	*1.67*	*0.10*
	6	152	*-0.10*	*-1.83*	*0.07*	0.11	1.81	0.07	**0.21**	**2.56**	**0.01**
	7	144	0.09	1.53	0.13	0.06	1.06	0.29	-0.03	-0.38	0.71
	8	171	-0.01	-0.28	0.78	-0.05	-0.90	0.37	-0.04	-0.48	0.63
	9	153	0.02	0.62	0.54	0.03	0.69	0.49	0.00	0.07	0.94
	10	220	-0.03	-0.68	0.49	**0.05**	**3.10**	**0.00**	*0.08*	*1.96*	*0.05*
2	1	123	0.05	1.29	0.20	0.00	0.00	1.00	-0.05	-0.94	0.35
	2	109	0.01	0.24	0.81	-0.01	-0.28	0.78	-0.03	-0.37	0.72
	3	96	**0.13**	**2.15**	**0.03**	0.11	1.41	0.16	-0.02	-0.21	0.83
	4	109	0.12	1.50	0.14	0.06	0.88	0.38	-0.06	-0.56	0.58
	5	118	-0.02	-0.37	0.72	0.06	0.91	0.36	0.08	0.90	0.37
	6	113	0.02	0.39	0.70	0.09	1.49	0.14	0.07	0.82	0.41
	7	119	0.00	-0.07	0.94	-0.01	-0.10	0.92	0.00	-0.03	0.97
	8	122	0.03	0.48	0.63	-0.05	-0.82	0.41	-0.07	-0.93	0.36
	9	89	-0.07	-1.14	0.26	-0.04	-0.76	0.45	0.02	0.29	0.77
3	1	48	0.15	1.44	0.16	0.15	1.97	0.05	-0.01	-0.04	0.97
	2	46	*0.22*	*1.81*	*0.08*	0.03	0.32	0.75	-0.19	-1.28	0.21
	3	56	*0.21*	*1.99*	*0.05*	0.10	1.18	0.24	-0.10	-0.74	0.46
	4	52	0.13	1.15	0.26	*0.13*	*1.78*	*0.08*	0.00	0.03	0.97
	5	62	0.08	0.64	0.53	-0.08	-0.92	0.36	-0.16	-1.06	0.29
	6	56	0.09	1.11	0.27	-0.05	-0.47	0.64	-0.14	-1.08	0.28
	7	40	-0.03	-0.26	0.80	0.03	0.41	0.69	0.06	0.45	0.66
	8	45	-0.01	-0.21	0.84	-0.11	-1.14	0.26	-0.09	-0.80	0.43

Note. Statistics under Graphical and Numerical describe the intercept in the regressions for each position. A positive coefficient indicates that participant down moves are more common than simulated down moves for that position. The difference column shows the coefficient on the condition variable. Coefficients significant at the $p \leq 0.05$ level are in bold while those significant at the $p \leq 0.10$ level are italicized.

Table 6.4.9. Regression of Clockwise vs Counterclockwise Moves onto Graphical/Numerical Conditions in Circular Conditions

	(1) lpm	(2) logit
Graphical/Numerical	-0.051^{+}	-0.206^{+}
	(0.026)	(0.107)
Set Shown S1	0.024	0.097
	(0.040)	(0.165)
Set Shown S2	-0.056	-0.228
	(0.039)	(0.159)
Set Shown S3	-0.036	-0.145
	(0.033)	(0.134)
Set Shown S4	-0.049	-0.198
	(0.038)	(0.154)
Set Shown S5	-0.029	-0.117
	(0.037)	(0.150)
Block: Late = 1	0.018	0.071
	(0.024)	(0.096)
Set Presentation Order	0.006	0.024
	(0.006)	(0.023)
Search Number	0.006	0.024
	(0.006)	(0.024)
Vertical Position: 2	0.004	0.016
	(0.033)	(0.135)
Vertical Position: 3	0.022	0.087
	(0.036)	(0.146)
Vertical Position: 4	0.032	0.130
	(0.034)	(0.137)
Vertical Position: 5	0.025	0.102
	(0.034)	(0.136)
Vertical Position: 6	0.031	0.127
	(0.037)	(0.149)
Vertical Position: 7	-0.012	-0.048
	(0.037)	(0.148)
Vertical Position: 8	-0.006	-0.023
	(0.037)	(0.150)
Vertical Position: 9	0.005	0.022
	(0.039)	(0.156)
Vertical Position: 10	-0.031	-0.125
	(0.036)	(0.144)
Vertical Position: 11	-0.022	-0.090
	(0.038)	(0.154)
Vertical Position: 12	0.051	0.210
	(0.040)	(0.165)
Constant	0.528^{***}	0.114
	(0.053)	(0.216)
Note:	$^{+}p<0.1$ | $^{*}p<0.05$ | $^{**}p<0.01$ | $^{***}p<0.001$	

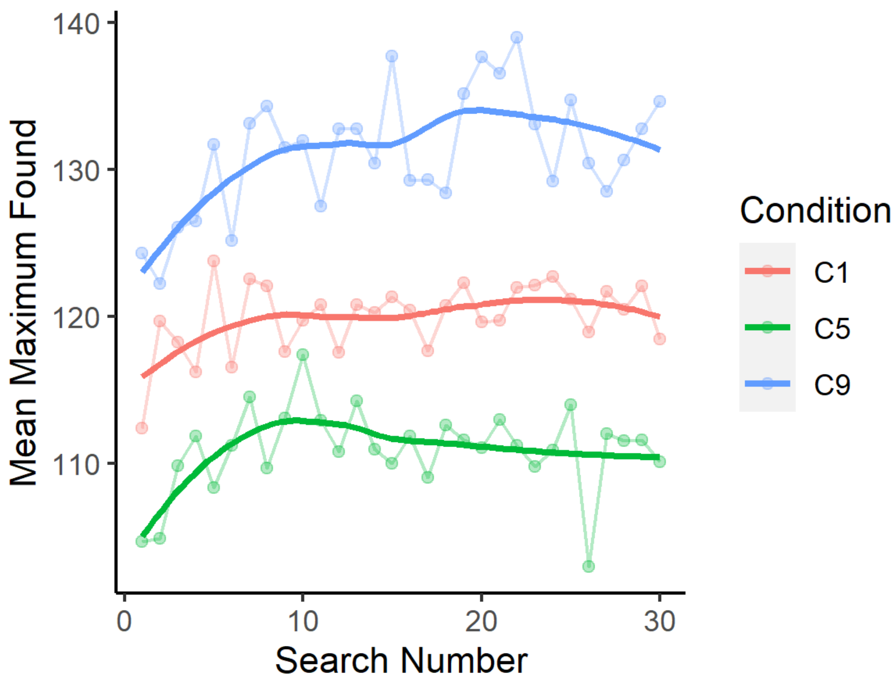
Table 6.4.10. Regression of Clockwise vs Counterclockwise Moves onto Graphical/Numerical Conditions in Circular Conditions

		Graphical Intercept			Numerical Intercept			Difference		
Position	df	*b*	*t*	*p*	*b*	*t*	*p*	*b*	*t(4983)*	*p*
1	380	**0.098**	**2.321**	**0.021**	-0.065	-1.534	0.126	**-0.164**	**-2.724**	**0.007**
2	423	**0.092**	**2.533**	**0.012**	-0.054	-1.296	0.196	**-0.146**	**-2.638**	**0.009**
3	428	*0.073*	*1.955*	*0.051*	0.014	0.343	0.732	-0.059	-1.060	0.290
4	491	**0.074**	**2.145**	**0.032**	0.034	1.088	0.277	-0.040	-0.873	0.383
5	422	**0.087**	**2.206**	**0.028**	0.013	0.321	0.748	-0.074	-1.336	0.182
6	403	**0.078**	**1.966**	**0.050**	0.041	1.139	0.255	-0.037	-0.687	0.492
7	398	0.029	0.685	0.494	0.000	0.000	1.000	-0.029	-0.490	0.625
8	415	**0.075**	**2.137**	**0.033**	-0.037	-1.044	0.297	**-0.112**	**-2.252**	**0.025**
9	400	0.019	0.442	0.659	0.039	1.108	0.268	0.020	0.354	0.724
10	396	0.025	0.781	0.436	-0.040	-1.318	0.188	-0.065	-1.471	0.142
11	418	0.008	0.193	0.847	-0.007	-0.186	0.852	-0.014	-0.268	0.789
12	389	0.041	1.037	0.301	**0.109**	**2.779**	**0.006**	0.068	1.215	0.225

Note. Statistics under Graphical and Numerical describe the intercept in the regressions for each position. A positive coefficient indicates that clockwise moves are more common than counterclockwise moves. The difference column shows the coefficient on the condition variable. Coefficients significant at the $p \leq 0.05$ level are in bold while those significant at the $p \leq 0.10$ level are italicized.

6.5. Supplementary Materials for Chapter 4, Study 1

Figure 6.5.1. Mean Maximum Values Found by Search Number in Study 1



Note. Trend lines are LOESS lines.

Table 6.5.1. Regression of Profit onto Search Number in Study 1

	Linear Model	Log Model
Search Number	0.104***	
	(0.031)	
ln(Search Number)		1.569***
		(0.329)
Cost = 5	-17.305***	-17.305***
	(1.027)	(1.027)
Cost = 9	-4.426***	-4.426***
	(1.180)	(1.180)
Intercept	115.842***	113.545***
	(0.980)	(1.189)
BIC	39013.45	39001.24
Nobs	4500	
Participants	1500	

Note. $^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001

Table 6.5.2. Regression of Ratio of Optimal to Participant Profit onto Search Number in Study 1

	Linear (1)	Log (2)
Search Number	0.0004^{+}	
	(0.0002)	
I(log(Search Number))		0.007*
		(0.003)
Cost = 5	-0.008	-0.008
	(0.007)	(0.007)
Cost = 9	0.001	0.001
	(0.010)	(0.010)
Constant	0.952***	0.942***
	(0.007)	(0.009)
BIC	-4714.89	-4718.74
Nobs	4500	
Participants	150	
Note:	$^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001	

Table 6.5.3. Regression of Maximum Value Found onto Search Number in Study 1

	Linear Model	Log Model
Search Number	0.125***	
	(0.033)	
ln(Search Number)		1.870***
		(0.356)
Cost = 5	-9.054***	-9.054
	(1.177)	(1.177)
Cost = 9	11.238***	11.238***
	(1.375)	(1.375)
Intercept	118.060***	11.238***
	(1.050)	(1.374)
BIC	38514.64	38495.65
Nobs	4500	
Participants	150	

Note. $^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001

Table 6.5.4. Regression of Search Duration onto Search Number in Study 1

	Linear Model	Log Model
Search Number	0.001	
	(0.002)	
ln(Search Number)		0.031
		(0.025)
Cost = 5	-0.391**	-0.391**
	(0.151)	(0.151)
Cost = 9	-0.528**	-0.528**
	(0.171)	(0.171)
Intercept	2.534	2.475***
	(0.128)	(0.135)
BIC	14367.8	14366.02
Nobs	4500	
Participants	1500	

Note. $^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001

Table 6.5.5. Regression of RV and EV First Choices onto Search Number in Study 1

	lpm		logistic	
	RV	EV	RV	EV
Block	0.001	0.001	0.006	0.002
	(0.001)	(0.001)	(0.005)	(0.005)
Display Order: 2	-0.004	-0.012	-0.018	-0.051
	(0.023)	(0.025)	(0.114)	(0.107)
Display Order: 3	0.013	0.009	0.064	0.038
	(0.021)	(0.023)	(0.100)	(0.098)
Display Order: 4	-0.016	0.020	-0.082	0.083
	(0.023)	(0.026)	(0.112)	(0.109)
Display Order: 5	0.008	-0.001	0.039	-0.002
	(0.022)	(0.026)	(0.104)	(0.108)
Cost = 5	0.077	-0.160^{***}	0.336	-0.674^{***}
	(0.049)	(0.047)	(0.217)	(0.196)
Cost = 9	-0.112^{**}	-0.056	-0.591^{**}	-0.225
	(0.041)	(0.057)	(0.206)	(0.232)
Constant	0.299^{***}	0.471^{***}	-0.856^{***}	-0.120
	(0.039)	(0.042)	(0.181)	(0.172)
Nobs	4500			
Participants	150			
Note.	$^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001			

Table 6.5.6. Linear Probability Model (Table A) and Logistic Regression Models (Table B) of RV and EV First Choices Onto Block By Condition

Table A

	Linear Probability Model					
	Cost = 1		Cost = 5		Cost = 9	
	RV	EV	RV	EV	RV	EV
Block	0.0002	0.002	-0.0005	0.001	0.004**	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Display Order: 2	0.067+	-0.065	-0.087*	0.047	0.001	-0.012
	(0.039)	(0.046)	(0.042)	(0.041)	(0.036)	(0.043)
Display Order: 3	0.040	0.025	0.005	-0.001	-0.005	0.004
	(0.039)	(0.041)	(0.036)	(0.034)	(0.031)	(0.045)
Display Order: 4	-0.026	0.034	-0.026	0.017	-0.008	0.016
	(0.045)	(0.056)	(0.039)	(0.040)	(0.034)	(0.041)
Display Order: 5	0.058	-0.038	0.008	-0.009	-0.045	0.048
	(0.039)	(0.052)	(0.041)	(0.038)	(0.030)	(0.039)
Constant	0.286***	0.458***	0.421***	0.289***	0.159***	0.445***
	(0.043)	(0.051)	(0.044)	(0.043)	(0.031)	(0.055)
Nobs	1500		1500		1500	
Participants	50		50		50	
Note:		+=<0.1, *=<0.05, **=<0.01, ***=<0.001				

Table B

	Logistic Regression					
	Cost = 1		Cost = 5		Cost = 9	
	RV	EV	RV	EV	RV	EV
Block	0.001	0.009	-0.002	0.007	0.023^{**}	-0.007
	(0.009)	(0.008)	(0.007)	(0.009)	(0.007)	(0.007)
Display Order: 2	0.309^{+}	-0.264	-0.375^{*}	0.212	0.006	-0.049
	(0.180)	(0.187)	(0.185)	(0.184)	(0.216)	(0.178)
Display Order: 3	0.190	0.099	0.021	-0.005	-0.031	0.017
	(0.183)	(0.164)	(0.146)	(0.157)	(0.185)	(0.187)
Display Order: 4	-0.129	0.138	-0.107	0.079	-0.051	0.065
	(0.228)	(0.225)	(0.162)	(0.186)	(0.205)	(0.168)
Display Order: 5	0.270	-0.154	0.034	-0.040	-0.295	0.195
	(0.179)	(0.210)	(0.167)	(0.177)	(0.201)	(0.159)
Constant	-0.914^{***}	-0.169	-0.319^{+}	-0.899^{***}	-1.655^{***}	-0.221
	(0.207)	(0.203)	(0.181)	(0.202)	(0.199)	(0.225)
Nobs	1500		1500		1500	
Participants	50		50		50	

Note: $^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001

Table 6.5.7. Regression of Mean RV and EV onto Search Number in Study 1

	RV		EV	
	Linear	Log	Linear	Log
Search Number	0.028^{+}		0.027^{*}	
	(0.015)		(0.013)	
Ln(Search Number)		0.377^{*}		0.367^{**}
		(0.167)		(0.142)
Cost = 5	-21.842^{***}	-21.844^{***}	-6.637^{***}	-6.639^{***}
	(0.766)	(0.766)	(0.695)	(0.695)
Cost = 9	-8.809^{***}	-8.811^{***}	12.499^{***}	12.497^{***}
	(0.749)	(0.750)	(0.813)	(0.813)
Search Length	-1.033^{***}	-1.037^{***}	-2.173^{***}	-2.177^{***}
	(0.175)	(0.176)	(0.204)	(0.204)
Constant	128.597^{***}	128.111^{***}	105.441^{***}	104.958^{***}
	(0.919)	(0.924)	(1.026)	(1.022)
Nobs	4500			
Participants	150			

Note: $^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001

Figure 6.5.2. Mean RV (Panel A) and Mean EV (Panel B) by Search Number and Learning Tertile by Condition in Study 1

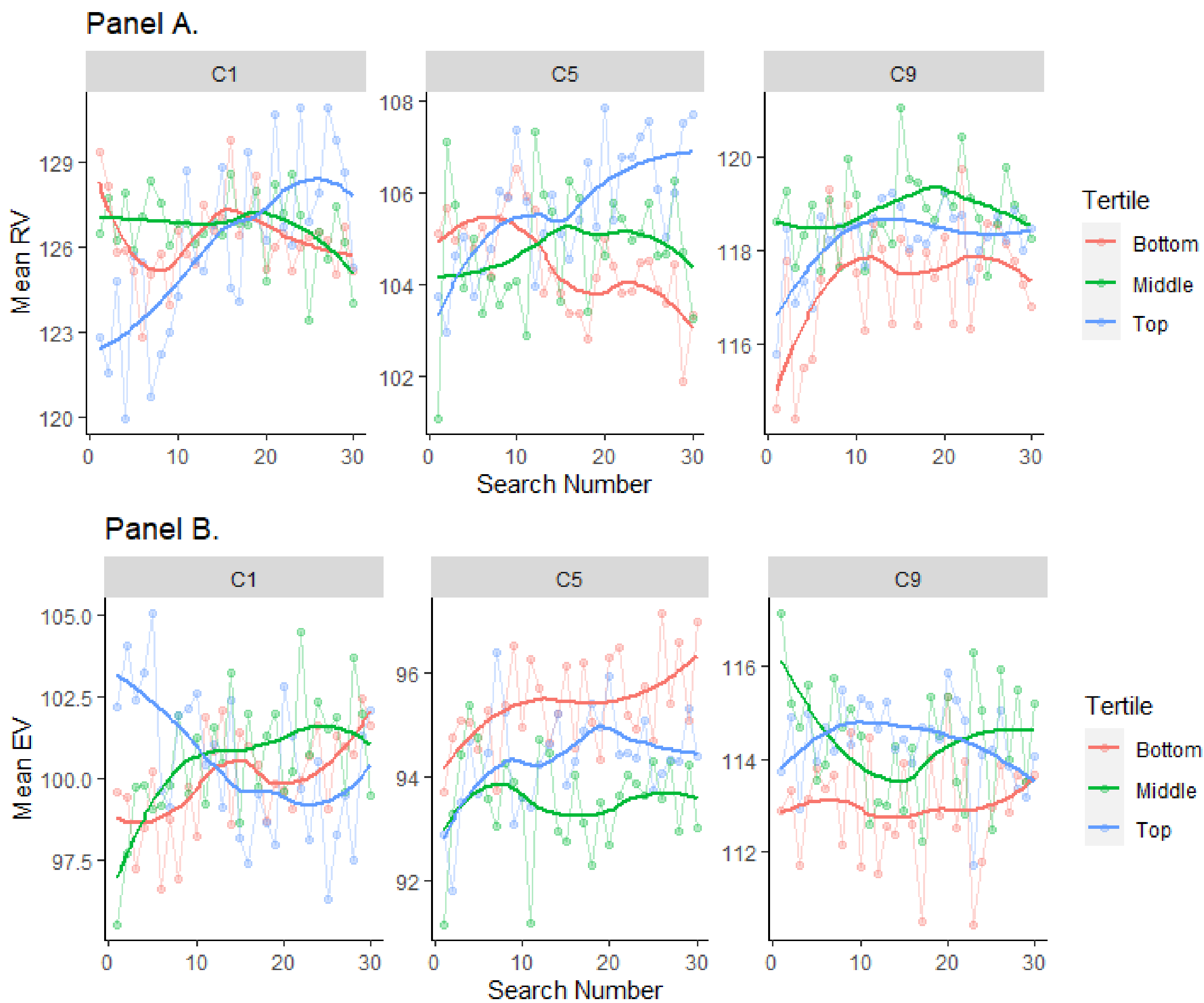


Table 6.5.8. Regression of Mean RV and EV onto Search Number by Tertile in Study 1

	Average RV			*Average EV*		
	Top	Middle	Bottom	Top	Middle	Bottom
ln(Search Number)	1.160***	0.111	-0.119	-0.058	0.677*	0.486*
	(0.289)	(0.224)	(0.310)	(0.206)	(0.285)	(0.228)
Cost = 5	-20.587***	-22.322***	-22.117***	-6.650***	-7.594***	-5.775***
	(1.258)	(1.232)	(1.637)	(1.041)	(1.198)	(1.268)
Cost = 9	-8.219***	-9.015***	-9.088***	12.996***	12.063***	12.170***
	(1.257)	(1.054)	(1.626)	(1.184)	(1.447)	(1.558)
Search Length	-0.987***	-1.566***	-0.541	-1.962***	-2.462***	-2.025***
	(0.217)	(0.276)	(0.377)	(0.264)	(0.367)	(0.41)
Constant	125.602***	130.340***	127.961***	105.561***	104.997***	104.056***
	(1.598)	(1.256)	(1.954)	(1.514)	(1.805)	(1.852)
Nobs	1500	1500	1500	1500	1500	1500
Participants	50	50	50	50	50	50
Note:	$^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001					

6.6. Supplementary Materials for Chapter 4, Study 2

Table 6.6.1. Regression of Profit onto Search Number in Study 2

	Linear Model	Log Model
Search Number	-0.010	
	(0.041)	
ln(Search Number)		0.124
		(0.439)
Cost = 4	-7.144	-7.144
	(1.374)	(1.374)
Intercept	144.968***	144.509***
	(1.374)	(1.489)
BIC	26789.84	26789.80
Nobs	3090	
Participants	1500	

Note. $^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001

Table 6.6.2. Regression of Ratio of Optimal to Participant Profit onto Search Number

	Linear Model		Log Model	
	No Interaction	Interaction	No Interaction	Interaction
Search Number	0.0004	0.001^{+}		
	(0.0002)	(0.0003)		
ln(Search Number)			0.005^{*}	0.006^{+}
			(0.003)	(0.003)
Cost = 4 Condition	-0.007	-0.002	-0.007	-0.001
	(0.008)	(0.012)	(0.008)	(0.016)
Interaction		-0.0003		-0.002
		(0.0005)		(0.005)
Constant	0.960^{***}	0.957^{***}	0.952^{***}	0.950^{***}
	(0.007)	(0.008)	(0.009)	(0.011)
BIC	-5810.27	-5802.89	-5813.5	-5805.78
Nobs	3090			
Participants	103			

Note: $^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001

Table 6.6.3. Regression of Maximum Found onto Search Number in Study 1

	Linear	Log
Search Number	-0.044	
	(0.044)	
ln(Search Number)		-0.211
		(0.469)
Cost = 4 Condition	-1.975	-1.975
	(1.411)	(1.411)
Constant	148.177***	148.023***
	(1.239)	(1.576)
BIC	26623.53	26624.62
Nobs	3090	
Participants	103	
Note:	$^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001	

Table 6.6.4. Regression of Search Duration onto Search Number in Study 1

	No Interaction	Interaction
Search Number	-0.012**	-0.008*
	(0.004)	(0.004)
Cost = 4 Condition	-0.719**	-0.609*
	(0.222)	(0.257)
Interaction		-0.007
		(0.007)
Constant	2.861***	2.806***
	(0.218)	(0.225)
BIC	11524.83	11531.66
Nobs	3090	
Participants	103	
Note:	$^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001	

Table 6.6.5.Regression of (A) RV and (B) EV First Box Choice and onto Search Number in Study 2

Table A (RV)	Linear Probability Model		Logistic Regression	
	No Interaction	Interaction	No Interaction	Interaction
Block	-0.000	0.001	-0.001	0.004
	(0.001)	(0.002)	(0.007)	(0.008)
Cost = 4 Condition	-0.041	-0.012	-0.213	-0.058
	(0.046)	(0.052)	(0.240)	(0.267)
Interaction		-0.002		-0.010
		(0.003)		(0.013)
Constant	0.289***	0.274***	-0.901***	-0.975***
	(0.034)	(0.035)	(0.168)	(0.174)
BIC	3737.28	3744.21	3588.9	3595.79
Nobs	3090			
Participants	103			
Note:		+=<0.1, *=<0.05, **=<0.01, ***=<0.001		

Table B (EV)	Linear Probability Model		Logistic Regression	
	No Interaction	Interaction	No Interaction	Interaction
Block	0.003+	0.000	0.011+	0.001
	(0.001)	(0.002)	(0.006)	(0.007)
Cost = 4 Condition	0.100+	0.018	0.403+	0.065
	(0.056)	(0.064)	(0.229)	(0.258)
Interaction		0.005+		0.022+
		(0.003)		(0.012)
Constant	0.442***	0.482***	-0.236	-0.070
	(0.044)	(0.045)	(0.177)	(0.178)
BIC	4464.26	4465.7	4254.61	4255.83
Nobs	3090			
Participants	103			
Note:		+=<0.1, *=<0.05, **=<0.01, ***=<0.001		

Table 6.6.6. Regression of Average RV and EV onto Search Number in Study 2

	RV			*EV*		
	linear	log	log w/ interaction	linear	log	log w/ interaction
Search No.	0.038			0.077^{**}		
	(0.031)			(0.025)		
Ln(Search No.)		0.753^{*}	0.748		0.817^{**}	0.521^{+}
		(0.336)	(0.526)		(0.263)	(0.286)
Cost=4	-12.808^{***}	-12.800^{***}	-12.823^{***}	0.5	0.499	-0.984
	(1.73)	(1.73)	(2.213)	(1.083)	(1.083)	(1.511)
Search No. X			0.009			0.597
Condition			(0.661)			(0.54)
Search Length	-2.442^{***}	-2.431^{***}	-2.431^{***}	-3.531^{***}	-3.532^{***}	-3.529^{***}
	(0.264)	(0.266)	(0.266)	(0.496)	(0.496)	(0.495)
Constant	159.479^{***}	158.162^{***}	158.174^{***}	132.483^{***}	131.653^{***}	132.378^{***}
	(1.63)	(1.759)	(1.924)	(1.558)	(1.639)	(1.534)
Nobs	6180			6180		
Participants	103			103		
BIC	23700.11	23692.90	23700.94	22096.42	22095.60	22101.03

Note. $^{+}$=<0.1, *=<0.05, **=<0.01, ***=<0.001

7. References

Anderson, N. H. (1981). *Foundations of Information Integration Theory*. Academic Press.

Anderson, N. H. (2014). *A Functional Theory of Cognition*. Psychology Press.

Baddeley, A. (2001). Is Working Memory Still Working? *American Psychologist*, *56*(11), 849–851.

Baddeley, A., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation - Advances in Research and Theory*, *8*(C), 47–89.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*(4), 723–767.

Benhabib, J., & Bull, C. (1983). Job Search: The Choice of Intensity. *Journal of Political Economy*, *91*(5), 747–764.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive Consumer Choice Processes. *Journal of Consumer Research*, *25*(3), 187–217.

Brannon, J. I., & Gorman, M. F. (2002). The effects of information costs on search and convergence in experimental markets. *Journal of Economic Behavior and Organization*, *47*(4), 375–390.

Braunstein, Y. M., & Schotter, A. (1982). Labor Market Search: An Experimental Study. *Economic Inquiry*, *20*(1), 133–144.

Brown, M., Flinn, C. J., & Schotter, A. (2011). Real-time search in the laboratory and the market. *The American Economic Review*, *101*(2), 948–974.

Butler, D., & Loomes, G. (1997). Quasi-rational search under incomplete information: Some evidence from experiments. *Manchester School*, *65*(2), 127–144.

Caplin, A., Dean, M., & Martin, D. (2011). Search and satisficing. *American Economic Review*, *101*(7), 2899–2922.

Clark, A., & Chalmers, D. (1998). The Extended Mind. *Analysis*, *58*(1), 7–19.

Cox, J. C., & Oaxaca, R. L. (1989). Laboratory experiments with a finite-horizon job-search model. *Journal of Risk and Uncertainty*, *2*(3), 301–329.

Cox, J. C., & Oaxaca, R. L. (1992). Direct Tests of the Reservation Wage Property. *The Economic Journal*, *102*(415), 1423–1432.

Danchin, E.; Giraldeau, L. & Cezilly, F. (2008). Behavioural Ecology. New York: Oxford University Press.

De Lillo, C., Aversano, M., Tuci, E., & Visalberghi, E. (1998). Spatial constraints and regulatory functions in monkeys' (Cebus apella) search. *Journal of Comparative Psychology*, *112*(4), 353–362.

De Lillo, C., Kirby, M., & James, F. (2014). Spatial working memory in immersive virtual reality foraging: path organization, traveling distance and search efficiency in humans (Homo sapiens). *American Journal of Primatology*, *76*(5), 436–446.

De los Santos, B., Hortaçsu, A., & Wildenbeest, M. R. (2012). Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior. *The American Econonomic Review*, *102*(6), 2955–2980.

Dellaert, B. G. C., & Häubl, G. (2012). Searching in choice mode: Consumer decision processes in product search with recommendations. *Journal of Marketing Research*, *49*(2), 277–288.

Droll, J. A., & Hayhoe, M. M. (2007). Trade-offs Between Gaze and Working Memory Use. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(6), 1352–1365.

Dunn, T. L., & Risko, E. F. (2016). Toward a Metacognitive Account of Cognitive Offloading. *Cognitive Science*, *40*(5), 1080–1127.

Einav, L. (2005). Informational asymmetries and observational learning in search. *Journal of Risk and Uncertainty*, *30*(3), 241–259.

Friedman, D., Pommerenke, K., Lukose, R., Milam, G., & Huberman, B. A. (2007). Searching for the sunk cost fallacy. *Experimental Economics*, *10*(1), 79–104.

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, *40*(1), 35–42.

Gabaix, B. X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly Information Acquisition : Experimental Analysis of a Boundedly Rational Model. *The American Economic Review*, *96*(4), 1043–1068.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the Fast and Frugal Way : Models of Bounded Rationality. *Psychological Review*, *103*(4), 650–669.

Gray, W. D., Sims, C. R., & Schoelles, M. J. (2006). The Soft Constraints Hypothesis : A Rational Analysis Approach to Resource Allocation for Interactive Behavior. *Psychological Review*, *113*(3), 461–482.

Häubl, G., Dellaert, B. G. C., & Donkers, B. (2010). Tunnel vision: Local behavioral influences on consumer decisions in product search. *Marketing Science*, *29*(3), 438–455.

Hey, J. D. (1982). Search for rules for search. *Journal of Economic Behavior and Organization*, *3*, 65–81.

Hey, J. D. (1987). Still searching. *Journal of Economic Behavior and Organization*, *8*, 137–144.

Hogarth, R. M., Lejarraga, T., & Soyer, E. (2015). The Two Settings of Kind and Wicked Learning Environments. *Current Directions in Psychological Science*, *24*(5), 379–385.

Honka, E., & Chintagunta, P. (2017). Simultaneous or sequential? Search strategies in the U.S. auto insurance industry. *Marketing Science*, *36*(1), 21–42.

Houser, D., & Winter, J. (2004). How Do Behavioral Assumptions Affect Structural Inference? Evidence from a Laboratory Experiment. *Journal of Business and Economic Statistics*, *22*(1), 64–79.

Hutchins, E. (1995). *Cognition in the wild*. MIT Press.

Hutchins, E. (2010). Cognitive Ecology. *Topics in Cognitive Science*, *2*, 705–715.

Jansma, J. M., Ramsey, N. F., De Zwart, J. A., Van Gelderen, P., & Duyn, J. H. (2007). fMRI study of effort and information processing in a working memory task. *Human Brain Mapping*, *28*(5), 431–440.

Johnson, E. J., & Payne, J. W. (1985). Effort and Accuracy in Choice. *Management Science*, *31*(4), 395–414.

Johnson, E. J., Payne, J. W., Bettman, J. R., & Schkade, D. A. (1989). Monitoring information processing and decisions: The mouselab system. *Duke University Center for Decision Studies*.

Kirsh, D. (1995). Artificial Intelligence The intelligent use of space. *Artificial Intelligence*, *73*(94), 31–68.

Kirsh, D. (2010). Thinking with external representations. *AI and Society*, *25*(4), 441–454.

Kirsh, D., & Maglio, P. (1994). On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science*, *18*(4), 513–549.

Kogut, C. A. (1990). Consumer search behavior and sunk costs. *Journal of Economic Behavior and Organization*, *14*(3), 381–392.

Land, M. F. (2014). Do we have an internal model of the outside world? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1636).

Lee, M. D. (2006). A hierarchical Bayesian model of human decision-making on an optimal stopping problem. *Cognitive Science*, *30*(3), 1–26.

Logie, R. H. (1995). *Visuo-Spatial Working Memory*. Psychology Press.

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science*, *20*(4), 240–245.

Martin, A., & Moon, P. (1992). Purchasing decisions, partial knowledge, and economic search experimental and simulation evidence. *Journal of Behavioral Decision Making*, *5*(4), 253–266.

McCall, B. P., & McCall, J. J. (2008). *The Economics of Search* (First). Routledge.

McCall, J. J. (1970). Economics of Information and Job Search. *The Quarterly Journal of Economics1*, *84*(1), 113–126.

Moon, P., & Martin, A. (1990). Better Heuristics for Economic Search — Experimental and Simulation Evidence. *Journal of Behavioral Decision Making*, *3*(3), 175–193.

Moorthy, S., Ratchford, B. T., & Talukdar, D. (1997). Consumer Information Search Revisited: Theory and Empirical Analysis. *Journal of Consumer Research*, *23*(4), 263.

Mortensen, D. (1970). Job Search, the Duration of Unemployment, and the Phillips Curve. *American Economic Review*, *60*(5), 847–862.

Mortensen, D. (1987). Job search and labor market analysis. In O. Ashenfelter & R. Layard (Eds.), *Handbook of Labor Economics vol 2* (pp. 849–919). Elsevier.

O'Regan, J. K. (1992). Solving the "Real" Mysteries of Visual Perception: The World as an Outside Memory. *Canadian Psychology*, *46*(3), 461–488.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The Adaptive Decision Maker*. Cambridge University Press.

Pirolli, P. (2010). Rational analysis of information foraging on the web. Cognitive Science, 29(3), 343-373.

Reinholtz, N. (2015). Persistence in Consumer Search. *Unpublished Manuscript*.

Risko, E. F., & Dunn, T. L. (2015). Storing information in-the-world: Metacognition and cognitive offloading in a short-term memory task. *Consciousness and Cognition*, *36*, 61–74.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive Offloading. *Trends in Cognitive Sciences*, *20*(9), 676–688.

Risko, E. F., Medimorec, S., Chisholm, J., & Kingstone, A. (2014). Rotating With Rotated Text: A Natural Behavior Approach to Investigating Cognitive Offloading. *Cognitive Science*, *38*(3), 537–564.

Rothschild, M. (1974). Searching for the Lowest Price When the Distribution of Prices Is Unknown. *Journal of Political Economy*, *82*(4), 689–711.

Russo, J. E., & Dosher, B. (1983). Strategies for Multiattribute Binary Choice. *Journal of Experimental Psychology: Learning Memory and Cognition*, *9*(4), 676–696.

Schonpflug, W. (1988). Retrieving texts from an external store: the effects of an explanatory context and of semantic fit between text and address. *Psychological Research*, *50*, 19–27.

Schotter, A., & Braunstein, Y. M. (1981). Economic Search: An Experimental Study. *Economic Inquiry*, *19*(1), 1–25.

Schunk, D. (2009). Behavioral heterogeneity in dynamic search situations: Theory and experimental evidence. *Journal of Economic Dynamics and Control*, *33*(9), 1719–1738.

Schunk, D., & Winter, J. (2009). The relationship between risk attitudes and heuristics in search tasks: A laboratory experiment. *Journal of Economic Behavior and Organization*, *71*(2), 347–360.

Shah, A., & Oppenheimer, D. (2008). Heuristics Made Easy: An Effort Reduction Framework. *Psychological Bulletin*, *134*(2), 207–222.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *Quarterly Journal of Economics*, *69*(1), 99–118.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138.

Smith, A. D., & De Lillo, C. (2022). Sources of variation in search and foraging: A theoretical perspective. *Quarterly Journal of Experimental Psychology*, *75*(2), 197–231.

Sonnemans, J. (1998). Strategies of search. *Journal of Economic Behavior & Organization*, *35*, 309–332.

Stigler, G. J. (1961). The Economics of Information. *Journal of Political Economy*, *69*(3), 213–225.

Stigler, G. J. (1962). Information in the Labor Market. *Journal of Political Economy*, *70*(5), 94–105.

Tversky, A. (1972). Elimination by Aspects: A Theory of Choice. *Psychological Review*, *79*(4), 281–299.

Urbany, J. E. (1986). An Experimental Examination of the Economics of Information. *Journal of Consumer Research*, *13*(2), 257.

Ursu, R. M. (2018). The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science*, *37*(4), 530–552.

Valenzuela, A., Dhar, R., & Zettelmeyer, F. (2009). Contingent Response to Self-Customization Procedures: Implications for Decision Satisfaction and Choice. *Journal of Marketing Research*, *46*(12), 754–763.

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica*, *47*(3), 641–654.

Zwick, R., Rapoport, A., Chung Lo, A. K., & Muthukrishnan, A. V. (2003). Consumer Sequential Search: Not Enough or Too Much? *Marketing Science*, *22*(4), 503–519.