THE USE OF CUSTOMIZATION TOOLS TO FACILITATE FINDING AND RE-FINDING OF INFORMATION IN A COMPUTER DESKTOP ENVIRONMENT

BY

YI WANG

THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Human Factors in the Graduate College of the University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Adviser:

Assistant Professor Wai-Tat Fu
ABSTRACT

We present results from three experiments that studied how people would use customization tools to help them offload information indexing to the external environment to augment finding and re-finding of information in a computer desktop environment. The results of the first experiment showed that participants were sensitive to the cost and benefit of customization. In general, participants performed more customization when the cost was low and when the benefit was high. Customization was also found to influence their information indexing strategies. The results of the second experiment showed that the use of customization tools helped participants develop a mix of internal (memory) and external (customization cues) indexing of information to adapt to dominance and dispersion structures of information needs and performance was significantly impacted by these information indexing strategies. The results of the third experiment showed that less convenient and efficient communication method and more help-needed questions respectively made participants ask their partners for less help and it thereby led to a lower proportion of optimal answers and a higher proportion of sub-optimal answers. As a result, it finally led to a worse performance. We concluded that there was much benefit for customization tools to be designed to adapt to different information environment in order to facilitate finding and re-finding information.
To Father and Mother
ACKNOWLEDGEMENTS

This thesis would not have been possible without the support of many people. Many thanks to my adviser, Assistant Professor Wai-Tat Fu, who inspired my research interests and cleared my confusion. Also thanks to Professor Alex Kirlik and Associate Professor Jason S. McCarley who taught me a lot of knowledge regarding Human Factors. Thanks to Peter Vlach, Jane Lawhead, and Karen Ayers of the Human Factors Division for their great help. Thanks to Philip Casella and Cassandra Medina who helped me run the third experiment and transcribed audio to text for data analysis. Most importantly, thanks to my family, Wenzhi Wang and Shunying Cao, for always being there with support and love.
# TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION ................................................................................................... 1  
1.1 Three Motivating Examples.......................................................................................... 1  
1.2 Research Background................................................................................................. 2  
1.3 Contributions............................................................................................................. 3  
1.4 Thesis Outline.......................................................................................................... 3

CHAPTER 2: LITERATURE REVIEW ....................................................................................... 4  
2.1 Information Finding, Re-finding and Indexing.......................................................... 4  
2.2 Customization............................................................................................................ 5  
2.3 Multidimensional Information, Dominant Dimensions and Dispersion of Information Needs.................................................................................................................... 6  
2.4 Collaborative Search................................................................................................. 7

CHAPTER 3: RESEARCH QUESTIONS AND HYPOTHESES ................................................... 9  
3.1 The First Set of Research Questions and Hypotheses.............................................. 9  
3.2 The Second Set of Research Questions and Hypotheses...................................... 11  
3.3 The Third Set of Research Questions and Hypotheses........................................... 17

CHAPTER 4: THE FIRST EXPERIMENT ............................................................................... 20  
4.1 Method...................................................................................................................... 20  
4.2 Results...................................................................................................................... 22  
4.3 Conclusion and Discussion....................................................................................... 26

CHAPTER 5: THE SECOND EXPERIMENT ....................................................................... 28  
5.1 Method...................................................................................................................... 28  
5.2 Results...................................................................................................................... 30  
5.3 Conclusion and Discussion....................................................................................... 42
CHAPTER 1
INTRODUCTION

Although a number of studies (Alvarado et al, 2003; Barreau & Nardi, 1995; Capra & Pérez-Quíñones, 2003; Dumais et al, 2003) investigated how people find and re-find information in different contexts, there have been relatively few studies on how people will make use of customization tools to facilitate finding and re-finding information in a computer desktop environment. However, systematic understanding of how people utilize the tools to index information in the external environment deserves more consideration as our personal document collections grow constantly and we become more and more reliant on repeatedly assessing electronic items in our personal archives, practically using them as our external memory stores (Chen et al, 2010). Studies on how people utilize the tools to index information in the external environment for later access will therefore shed light on how we can design better tools that help us more effectively offload information to the external world and that enable us quickly re-find the information when it is needed again.

1.1 THREE MOTIVATING EXAMPLES

Example 1. Imagine you are a psychologist who gives diagnoses to patients with a potential psychological disorder. You have a wide range of information concerning these disorders (e.g., dissociative disorders, mood episodes, and schizophrenia) filed on your computer desktop. How would you use your distributed information resources (e.g., electronic documents) to make diagnosis for different patients? It is likely that you may have to access a subset of information repeatedly to inform decisions, or to find and re-find relevant information when patients with similar symptoms appear at different intervals. Having a tool that helps you to index useful information so that you can easily re-find them when they are needed will greatly enhance your task efficiency.
Example 2. Imagine you are a travel agent and your job is to recommend or provide details of travel itineraries based on your customers’ requests. Depending on the season and other factors, there may be more people asking about travel plans to certain popular cities or looking for plans that fit particular time or monetary budgets, which may make certain itineraries more popular than others. Imagine that details of these travel plans are stored in multiple documents that you have access on your desktop, how would you customize accesses to these documents to better serve your customers? You may have to find and re-find information about travel plans that vary in their levels of demand. Having a set of customization tools that help you adapt to the structures of information demands can significantly improve the effectiveness of information indexing - i.e., utilizing cues that help you re-find information in a desktop environment.

Example 3. Imagine you are such a travel agent described in Example 2, but instead of working individually, you will work with one of your colleagues. Each of you will own one computer and similar but different documents regarding travel plans are filed on your computer desktops. How would you and your colleague exchange information to better serve your customers? You two may use the customization tools to index information in order to facilitate finding and re-finding of useful information.

1.2 RESEARCH BACKGROUND

Indeed, this kind of information finding and re-finding described in the above examples is common. We have done extensive field interviews to know about real-life situations. First, we investigated healthcare centers and found that they were using exactly this kind of interface to have multiple patient records presented as icons on a computer desktop, from which nurses needed to find, organize and re-find patients’ records to retrieve useful information. In addition, we also interviewed local travel agencies to understand how they were performing routine information organization and sharing. There were about ten representatives; each one was in charge of several travel itineraries. We observed that they tended to directly put the travel folder icons on the desktop and each representative tended to customize their icons differently depending on the types of customers and tours they handled. This motivated us to study how
people would customize (if customization options were available) their own desktop environment to index information for later use.

1.3 CONTRIBUTIONS

This thesis reports three experiments. The first experiment was about an individual desktop search task with an icon size editing tool. We found that people’s customization behavior was sensitive to the cost and benefit of customization. Specifically, lower cost and higher benefit of customization would lead to more customization behavior. We therefore performed the second experiment – an individual desktop search task with a set of icon editing tools – in which more task questions and customization options were designed to encourage more customization behavior. We did find participants fully utilized the tools to index information. Additionally, we also found people used the tools to adapt to the structures of information needs and developed a mix of internal and external information indexing. To understand how people made use of customization tools to index information and worked together, we conducted the third experiment, a two-participant collaborative desktop search task with a set of icon editing tools. We found different communication methods led to different search performance and the participants who needed less help asked comparatively more questions than their partners. These studies made unique contributions to the research field of desktop search since they made us have a better understanding of how people would take advantage of customization tools to find re-find and exchange information in individual or collaborative desktop search tasks.

1.4 THESIS OUTLINE

The thesis is organized as follows. Chapter 2 reviews some past and ongoing research in desktop search with customization tools. Chapter 3 proposes research questions and hypotheses. Chapter 4, Chapter 5 and Chapter 6 report the first, second and third experiments in detail, respectively. Chapter 7 draws conclusions from the three experiments. Chapter 8 made general discussion and Chapter 9 talks about limitations and suggests future research direction.
CHAPTER 2
LITERATURE REVIEW

In this section, I introduce previous research regarding (1) information finding, re-finding and indexing, (2) customization, (3) multidimensional information, dominant dimensions, and dispersion of information needs, and (4) collaborative search to provide the main theoretical frameworks behind our current studies.

2.1 INFORMATION FINDING, RE-FINDING AND INDEXING

Each day we access a significant number of new electronic items such as web pages, emails, documents, spreadsheets, digital photos and mp3 files (Chen et al, 2010) and file them in our local archives for accessing them again in the future (Dumais et al, 2003). Thus it is important to understand how people find and re-find electronic information in their local archives. Finding electronic information is well understood to be a complex, multi-stage process (Bates, 1989; Belkin, 1993; Belkin et al, 1993; Marchionini, 1995; O’Day & Jeffries, 1993; Teevan, 2007). Although finding typically involves simple searches for information that is known in advance, the search behavior follows the broader information-seeking pattern. Several studies of finding behaviors (Nardi & Barreau, 1995; Ravasio et al, 2004; Teevan et al, 2004) suggest that people prefer to perform even the most directed searches by orienteering via small, local steps using their contextual knowledge as a guide, rather than by teleporting, or jumping directly to it using a keyword-search utility.

Similar to finding new information, re-finding electronic information is also common (Byrne, 1999; Cockburn et al, 2003; Tauscher & Greenberg, 1997). Compared to finding, one distinguishing feature of re-finding is that the searcher often knows a lot of meta-information (or contextual information) about the target, such as its author, title, date created, color, or style of text (Capra & Pérez-Quiñones, 2005; Dumais et al, 2003; Lansdale & Edmonds, 1992; Ringel et al, 2003). Contextual knowledge actually plays a more important role in re-finding than finding
(Teevan, 2007). It is important that search results appear where expected during re-finding (Teevan, 2007).

Specifically, in a computer desktop environment, characteristics of interface icons can serve as contextual cues and users can easily employ these cues to internally (cognitively) or externally index information to help them find and re-find relevant information. The icons can actually convey indicative meaning with slight alteration of their visual features (Ehret, 2002; Houde & Salomon, 1993). For example, icon luminance was designed to visually represent cues such as frequency and recency of accesses to these icons (Moon & Fu, 2009), and icon size was used to represent cues such as the importance of the icons and the categories the icons belong to.

2.2 CUSTOMIZATION

Nowadays, more and more customizable features in application software are designed for the end users, providing specific mechanisms for people to specify individual preferences about the software and how they will interact with it. These customizations, along with choices about which applications to use, constitute the unique “customizable application environment” for each individual.

Mackay (1991) argued that usually, people’s customizations need one or more reasons. She found that people were most likely to customize when they discovered that they were doing something repeatedly and chose to automate the process. Also very common was customization for the purpose of stopping something that was annoying or slow. In addition, she reported that unless the user is bored or just learning a new system, customizations that make the software environment aesthetically pleasing or more interesting are generally avoided.

It is theoretically possible for people to freely utilize any customization options, but the general observation is that many people choose not to customize. They spent most of their time simply “using” the software (Mackay, 1991). Mackay (1988) showed that users often resist using software features, and simply providing a set of customization features does not ensure that users will take advantage of them to achieve a performance improvement.
People are busy and customizing takes time, so they only customize when they deem it worth the trouble (Findlater & McGrenere, 2004; Mackay, 1991; McGrenere et al, 2002). One theoretical framework to systematically study how and why people customize is to cast the decision as a trade-off between a short-term investment and a long-term potential benefit (Fu & Gray, 2000; Gray & Fu, 2001; Gray et al, 2006). In general, one can assume that it is desirable if the long-term benefits of customizing justify the short-term cost of doing so (Fu & Gray, 2006; Fu & Pirolli, 2007; Gray et al, 2006). Therefore, one would expect that the more demanding the task is, the more likely people would like to use customization tools to offload indexing to the external environment.

2.3 MULTIDIMENSIONAL INFORMATION, DOMINANT DIMENSIONS, AND DISPERSION OF INFORMATION NEEDS

Information is often organized as some clusters of “records” or “objects”, each of them containing multiple fields or dimensions that contain information related to each other (Morrison et al, 2001; Xie, 2009). For example, a patient’s record can include her/his name, gender, age, SSN (social security number), address and phone number. Among them, name, gender, and address are non-quantitative information dimensions, while age, SSN and phone numbers are quantitative information dimensions. These dimensions are probably followed by a detailed personal history of diseases. Other examples include an introduction to a university (non-quantitative dimensions such as name and location as well as quantitative dimensions such as year of establishment and number of students are probably followed by details regarding faculty and academics), a travel plan (non-quantitative dimensions such as travel destination and travel agency as well as quantitative dimensions such as travel cost and travel duration are probably followed by details concerning scenic spots and historical sites at the destination), and so on.

In the above examples, one may find that some dimensions (dominant dimensions) may be used or requested more often than others (Lansdale & Edmonds, 1992; Teevan, 2007). For example, a university’s name (or a travel destination) is probably used more often than its location because the name (or destination) is usually used as an index of information record to retrieve other dimensions of the information related to that record. In addition, one may also find that some
information records are probably requested more often than others (dispersion of information needs) (Teevan, 2007). For example, there are probably more queries about the trip to city A than the trip to city B due to the fact that more places of interest are located in city A. Indeed, multidimensional information, dominant dimensions, and dispersion of information needs are very common in many naturally occurring information ecologies (Dourish et al, 2000; Teevan, 2007). It is therefore valuable to understand how customization tools could be designed with respect to these statistical structures of information needs to facilitate finding and re-finding of information.

2.4 COLLABORATIVE SEARCH

Computer based search is generally considered a solitary task. However, there are many situations in which people desire to collaborate on search tasks. Students often want to collaboratively search information to complete homework assignments or group projects (Amershi & Morris, 2009; Amershi & Morris, 2008; Large et al, 2002; Twidale et al, 1997). Friends or families planning a vacation to Hawaii might want to search information together to identify cheap airfares, appropriate hotels and interesting tourist activities. Colleagues in a research lab who are working together to write a scholarly article probably want to perform a joint literature search (Morris, 2008). Actually, just like the evidence provided by Morris’s investigations (Morris, 2008), collaborative searching is surprisingly commonplace.

There are some previous studies regarding collaborative search. However, they are mainly about working together to search the web, and very few studies are related to collaboratively searching the computer desktop. For example, SearchTogether (Morris & Horvitz, 2007) and Cerchiamo (Pickens et al, 2008) support real-time collaboration on web search tasks amongst a group of users who each have their own computer. The Sociable Web (Donath and Robertson, 1994) allows a user to know that others were currently viewing the same webpage, and to communicate with those people. WebTagger (Keller et al, 1997) allows users to share bookmarks or favorites lists. GroupWeb (Greenberg and Roseman, 1996) provides yoked browsing and telepointers, as well as allowing users to associate comments with jointly-viewed Web pages. Gatherer
(Schraefel, 2002) and Google’s Notebook application allow users to collect snippets of content from several webpages and combine them in a single document.

To share information and facilitate collaborative search, HCI (human-computer interaction) and IR (information retrieval) researchers have begun to design systems with diverse communication methods (Amershi & Morris, 2008; Diamadis & Polyzos, 2004; Freyne & Smyth, 2006; Morris & Horvitz, 2007; Pickens et al, 2008). Specifically, some systems were designed for very specialized domains and/or devices, and the search tasks are performed among groups of users who know each other and are working together toward a shared goal. For example, TeamSearch (Morris et al, 2006) is a system that enables co-located groups of up to four people to simultaneously search collections of digital photographs, using a visual query language designed for a multi-user interactive tabletop. MUSE (Krishnappa, 2005) is a system that supports synchronous, remote collaboration between two people searching a medical database. Users of this system perform standard, single-user searches, but have a built-in textual chat facility as well as the ability to press a “share” button that sends some metadata about what they have found to the other user. However, the most common methods reported for collaborating on search were still traditionally emailing links back and forth, using instant messaging software to exchange links and query terms, and speaking with a collaborator on the phone while viewing a Web browser (Morris & Horvitz, 2007).
CHAPTER 3
RESEARCH QUESTIONS AND HYPOTHESES

In this section, three sets of research questions and hypotheses are proposed to explore desktop search with customization tools. Based on previous studies, we believe that the cost and benefit of customization might greatly influence people’s search performance, willingness to customize and information indexing strategies. We thereby propose the first set of research questions and hypotheses to test them. Different structures of information needs are pretty common in our daily life. We believe that these structures of information needs might interact with customization tools to influence people’s search performance, customization behavior, search behavior and information indexing strategies. Therefore, we propose the second set of research questions and hypotheses to test them. Desktop search is generally considered to be a solitary activity. All major applications in desktop environment are designed for solo use. However, we found many tasks can benefit from the ability to collaboratively search the desktops with others. Our intuition was that such situations might be commonplace. Therefore, we were very interested in collaborative desktop search. To better study it, we included the third set of research questions and hypotheses in our studies. Specifically, we would investigate how some particular factors in a collaborative search task would influence search performance and communication behavior.

3.1 THE FIRST SET OF RESEARCH QUESTIONS AND HYPOTHESES

One theoretical framework to systematically study how and why people customize is to cast the decision as a trade-off between a short-term investment and a longer-term potential benefit. In general, one can assume that it is desirable if the long-term benefits of customizing justify the short-term cost of doing so. Indeed, this kind of tradeoff has been studied in previous human factors research. For example, Gray et al (2006) proposed the theory of soft constraints to characterize this tradeoff, which states that the decision on when to act is sensitive to the time costs of the alternative actions. They found that people might even adopt suboptimal actions, when they somehow perceive that the short-term costs are not justified by the long-term benefit;
however, with experience, the perception of benefit may change and their decision on actions may tend to approach optimality across time (Fu & Gray, 2006; Gray et al, 2006). Therefore, although anecdotal evidence seems to suggest that people do not customize their desktop for re-finding of information, there seems to be other factors that may influence their willingness to do so, and one such important factor is the cost-benefit tradeoff involved in customization (Fu & Gray, 2006).

3.1.1 Research Questions

We therefore conducted an experiment (To be consistent with the studies described later in this thesis, the current one is called the first experiment) and directly manipulated the cost and benefit of customization, and tested the following research questions (To be consistent with the set of research questions proposed later in the thesis, the current one are called the first set of research questions):

(1) Do the cost and benefit of customization influence search performance?
(2) How do the cost and benefit of customization influence the willingness to customize?
(3) How do the cost and benefit of customization influence information indexing strategies (i.e., how to re-access found information)?

The first research question tested whether the variation of cost and benefit of customization would influence people’s search performance. It is worthwhile because if the answer to this question is positive, then it implies that we can design customization tools with appropriate cost and benefit to enhance people’s desktop search performance. The second research question tested how the variation of cost and benefit of customization would influence people’s willingness to customize. It is important because we need to know whether, to what extent, and how people are sensitive to the cost and benefit of customization, and based on this finding, we can design the subsequent studies to encourage more customization behavior and further study it. The third research question tested how the variation of cost and benefit of customization would lead to different information indexing strategies. It is also important because better understanding of how people develop their own information indexing strategies to adapt to the
cost and benefit of customization would greatly help the design of novel and helpful customization tools and finally enhance people’s search performance.

3.1.2 Hypotheses

To test the three research questions, we propose the following hypotheses:
(1) Lower cost and higher benefit of customization lead to finding and re-finding more accurate information in shorter time.
(2) Lower cost and higher benefit of customization lead to stronger willingness to customize.
(3) Lower cost and higher benefit of customization lead to more diverse information indexing strategies.

Since the amount of accurate information retrieval and the corresponding time could be used to measure the search performance mentioned in the first research question, the first hypothesis allowed us to test the first research question. We wanted to see whether lower cost and higher benefit of customization would result in more accurate information retrieval in shorter time. Related to the second research question, the second hypothesis tested whether decreasing cost and increasing benefit would increasing people’s willingness to customize. The third hypothesis corresponded to the third research question. We expected that decreasing cost and increasing benefit of customization would make people try more diverse information indexing strategies.

I will present the method and results regarding to the first set of research questions and hypotheses in Chapter 4 (the first experiment).

3.2 THE SECOND SET OF RESEARCH QUESTIONS AND HYPOTHESES

Finding and repeatedly re-finding information in a computer desktop environment is very commonplace. For example, healthcare center often has multiple patient records presented as icons on computer desktops, from which nurses often need to find and re-find a particular patient’s record to retrieve information. Similarly, real estate agents often need to find and re-find property records to retrieve necessary information that fits the buyers/sellers’ requests
(Moon & Fu, 2009). In other words, it exists in domains where the operators have to manage a large number of similar records that contain multi-dimensional information (e.g., travel destination, travel cost, departure date, and so on) that are frequently requested and need to be found repeatedly by the operators.

Given the limited capacity of human working memory, it will quickly become hard to mentally index all information as the complexity and size of desktop icons keeps increasing (Chen et al, 2010). One common solution to the problem could be externally indexing information. To enable people to externally index information, some customization tools will be designed such that people can make use of the tools to adjust the visual features of desktop icons. By making this adjustment, desktop icons will become organized based on their categories, sizes, or other attributes, and thereby contextual cues will be generated and people’s information finding and re-finding will be facilitated. However, when the customization tools are not available, people will have to rely purely on their internal information indexing. Comparing performance between the two groups will provide baseline information about the difference in external and internal information indexing. Indeed, even though the customization tools are available, one often cannot completely offload information indexing to the environment. Instead, the user has to rely on effectively developing a mix of internal and external indexing of information.

It is therefore possible that other factors than customization options exist in influencing information indexing strategies. We believed one such important factor is the statistical structure of information needs and hence we were interested in how information indexing strategies would be influenced by different structures of information needs. Specifically speaking, we were interested in knowing how information indexing strategies would change when there was a subset of information dimensions (e.g., travel destination, costs, etc) that were used as indexing cues more often than others. For example, it was possible that most people would like to know more about plans to a particular travel destination, and the travel agents could more easily find relevant information for customers if their records were indexed based on destination, such that all information related to any destination could be retrieved. It was thereby reasonable that when some dominant dimensions exist in the distribution of information needs, both internal and external indexing strategies could change. In addition to dominant dimensions, when some travel
plans were requested more often than others, it was possible that people would highlight or encode those plans so that they could be retrieved more easily in the future. Therefore, we were also interested in the influence of the dispersion of information needs – i.e., whether there was a skewed or uniform distribution of information needs among a set of documents – on information indexing strategies.

3.2.1 Research Questions

For the above concerns, we conducted an experiment (It is called the second experiment for the consistent reason) and tested the research questions (They are called the second set of research questions for the consistent reason) as follows.

(1) How do customization tools, dominance and dispersion structures of information needs influence the performance of finding and re-finding information?
(2) How do dominance and dispersion structures of information needs influence customization behavior?
(3) How do customization tools, dominance and dispersion structures of information needs influence search behavior?
(4) How do customization tools, dominance and dispersion structures of information needs influence developing a mix of internal and external information indexing strategies?

The first research question tested how the use of customization tools and different structures of information needs would influence people’s search performance. It is important because the answer to this question may enlighten us to better design tools for different structures of information needs to enhance people’s search performance. The second research question tested how different structures of information needs would influence how people customize visual features of desktop environment. It is also important because it make us better understand how people would customize desktop applications to adapt to structures of information needs. The third research question tested how the use of customization tools and different structures of information needs would influence people’s search behavior. It worth the time since it can make us get a better understanding of how people’s search behavior would change depending on the use of customization options and different structures of information needs and thus design better
tools to match different information needs in order to optimize people’s search behavior. The fourth research question tested how the mix of internal and external information indexing strategies would change based on customization options and different structures of information needs. It is worthwhile because a better understanding of how the shift of internal and external information indexing will be influenced by both the customization tools and structures of information needs will greatly enhance our ability to predict when and how different customization tools should be designed for users who have different needs. Additionally, people’s information indexing strategies were implicit. We needed to analyze more explicit measures such search performance, customization behavior, and search behavior to reveal the strategies, and a better understanding of the strategies would in turn lead to a better understanding of people’s customization behavior and search behavior, and finally lead to a better design of customization tools in order to help people improve search performance.

3.2.2 Hypotheses

To test the above research questions, we propose four sets of hypotheses. Hypothesis set 1 corresponds to the first research question; hypothesis set 2 corresponds to the second research question, and so on. I list the four hypothesis sets as below.

Hypothesis set 1:
(1) The use of customization tools leads to finding and re-finding more accurate information in shorter time.
(2) The dominance structure of information needs leads to finding and re-finding more accurate information in shorter time.
(3) The dispersion structure of information needs leads to finding and re-finding more accurate information in shorter time.
(4) The use of customization tools and the dominance structure of information needs interactively lead to finding and re-finding more accurate information in shorter time.
(5) The use of customization tools and the dispersion structure of information needs interactively lead to finding and re-finding more accurate information in shorter time.
(6) The dominance and dispersion structures of information needs interactively lead to finding and re-finding more accurate information in shorter time.

The hypothesis set 1 tested how the performance of finding and re-finding information would change based on whether the customization tools were available and whether the dominance and/or dispersion structures of information needs existed. Since search performance is the most important measure for a search task, this set of hypotheses is worthwhile. By testing them, we can predict what tools should be designed for people to adapt to different structures of information needs and hence improve search performance.

**Hypothesis set 2:**
(1) The dominance structure of information needs leads to more customization behavior.
(2) The dispersion structure of information needs leads to more customization behavior.
(3) The dominance and dispersion structures of information needs interactively lead to more customization behavior.

The hypothesis set 2 tested how dominance and dispersion structures of information needs individually or interactively influence people’s customization behavior. By testing these hypotheses, we can better understand how people would utilize the tools to adapt to different structures of information needs.

**Hypothesis set 3:**
(1) The use of customization tools leads to more search behavior.
(2) The dominance structure of information needs leads to more search behavior.
(3) The dispersion structure of information needs leads to more search behavior.
(4) The use of customization tools and the dominance structure of information needs interactively lead to more search behavior.
(5) The use of customization tools and the dispersion structure of information needs interactively lead to more search behavior.
(6) The dominance and dispersion structures of information needs interactively lead to more search behavior.
The hypothesis set 3 tested how customization tools, dominance and dispersion structures of information needs individually or interactively influence people’s search behavior. By testing these hypotheses, we can better design the tools for people to adapt to different structures of information needs and thereby optimize their search behavior such that they can make less effort to find and re-find more useful information.

**Hypothesis set 4:**
(1) The use of customization tools leads to more diverse information indexing strategies.
(2) The dominance structure of information needs leads to more diverse internal and external information indexing strategies.
(3) The dispersion structure of information needs leads to more diverse internal and external information indexing strategies.
(4) The use of customization tools and the dominance structure of information needs interactively lead to more diverse internal and external information indexing strategies.
(5) The use of customization tools and the dispersion structure of information needs interactively lead to more diverse internal and external information indexing strategies.
(6) The dominance and dispersion structures of information needs interactively lead to more diverse internal and external information indexing strategies.

The hypothesis set 4 tested how customization tools, dominance and dispersion structures of information needs individually or interactively influence people’s information indexing strategies. A better understanding of how people will develop a mix of internal and external strategies of information indexing to adapt to different structures of information needs will make us design better tools for people to enhance their search performance, and therefore it is important to test these hypotheses.

I will present the method and results regarding to the second set of research questions and hypotheses in Chapter 5 (the second experiment).
3.3 THE THIRD SET OF RESEARCH QUESTIONS AND HYPOTHESES

Although many studies were performed regarding collaborative search, few studies were reported about providing customization tools and instant messaging software (e.g. MSN) for participants to perform a collaborative desktop search task. However, this kind of research worth the time because it simulates real life very well especially for some small business (e.g. small travel agency with several agents who have travel itineraries folders filed on their computer desktops and need to share the folders with their colleagues in order to serve their customers).

Therefore, we wanted to do some research in this area. Specially, we were concerned that if MSN was provided for the participants to either type text message to each other or directly talk to each other to share information, what the difference between the two communication methods was. In addition, we were concerned that if in the collaborative search task, different group members needed to ask their partners for different amounts of help, what the difference between group members was.

3.3.1 Research Questions

We would like to propose our research questions and hypotheses as below.

(1) How do different communication methods and different amounts of help needed influence search performance?

(2) How do different communication methods and different amounts of help needed influence customization behavior?

(3) How do different communication methods and different amounts of help needed influence communication behavior?

The first research question tested how the performance of finding and re-finding information would change when different communication methods and different amounts of help needed by different group members existed. We believed it was important because it would make us better understand the relationships between communication methods and search performance and between the amounts of help needed and search performance. The second research question tested how people’s customization behavior would differ when different communication methods...
and different amounts of help needed were given. It would be interesting if we could find that people tended to use customization options to adapt to their communication methods and their amount of help needed. The third research question tested how communication behavior would change when communication methods and amounts of help needed differ. Based on previous findings we believed that people’s communication behavior would change to adapt to different communication methods and different amounts of help needed, but how and to what extent this change would happen still kept unclear. It motivated us to perform an experiment to directly investigate it.

3.3.2 Hypotheses

To test the above research questions, we propose three hypothesis sets. Hypothesis set 1 corresponds to the first research question, and hypothesis set 2 corresponds to the second research question, and so on. I list the three hypothesis sets as below.

**Hypothesis set 1:**

1. Audio condition leads to finding and re-finding more accurate information in shorter time.
2. Less help needed leads to finding and re-finding more accurate information in shorter time.
3. Audio condition and less help needed interactively lead to finding and re-finding more accurate information in shorter time.

The hypothesis set 1 tested how the performance of finding and re-finding information would change based on different communication methods and different amounts of help needed by different group members. By testing them, we could design novel communication methods for different group members in a collaborative search task to improve their search performance.

**Hypothesis set 2:**

1. Audio condition leads to more customization behavior.
2. Less help needed leads to more customization behavior.
3. Audio condition and less help needed interactively lead to more customization behavior.
The hypothesis set 2 tested how different communication methods and different amounts of help needed would individually or interactively influence the amount of customization behavior. By testing these hypotheses, we could have a better understanding of how to design better customization tools to satisfy people’s different demands.

**Hypothesis set 3:**

(1) Audio condition leads to more communication behavior.

(2) Less help needed leads to more communication behavior.

(3) Audio condition and less help needed interactively lead to more communication behavior.

The hypothesis set 3 tested how different communication methods and different amounts of help needed by different group members individually or interactively influence the amount of communication behavior. By testing these hypotheses, we could have a better understanding of designing novel communication methods for different group members in a collaborative search task to facilitate communication between them.

I will present the method and results regarding to the third set of research questions and hypotheses in Chapter 6 (the third experiment).
CHAPTER 4
THE FIRST EXPERIMENT

To test the hypotheses of the first set of research questions talked in Chapter 3, we conducted the first experiment. In this experiment, participants were required to perform one individual desktop search task with a desktop icon size edit tool. We wanted to see how the cost and benefit of customization might greatly influence people’s search performance, willingness to customize and information indexing strategies.

4.1 METHOD

4.1.1 Experiment Design

In this experiment, two kinds of icon edit tools: *track bar* (low cost, abbreviated as Lo) and *increase/decrease buttons* (high cost, abbreviated as Hi), and two kinds of organization of icons: *random* and *organized* (abbreviated as Org) were used. Therefore, there were four conditions: Lo-Random, Lo-Org, Hi-Random, and Hi-Org.

The size of each icon could be changed from 1x1 pixel to 60x60 pixels, and originally each icon had the size of 20x20 pixels. The track bar and increase/decrease buttons were used to change the size of the icons. We considered the track bar a low cost tool and increase/decrease buttons a high cost tool because when participants used the track bar to adjust the size of a particular icon, they only needed to drag their cursors on the scale (which usually took less than 1 second for this simple point-and-drag action); but for the increase/decrease buttons tool, participants had to click the increase (or decrease) buttons multiple times (each time increase (or decrease) the size by 1 pixel). Additionally, in both organized and random conditions, the locations of icons always stayed the same. However, in the organized condition, the adjacent icons shared similar contents and this could be location cues for searching; but in random condition, the icons with similar contents were not grouped together, thus no location cues could be used. Comparing these two
conditions, we could find customization which left size cues could benefit the participants more in random than in organized condition.

We expected that in the high-cost conditions, participants were less willing to change the size of the icons compared to the low-cost conditions. Because there was a higher benefit for customizing the icon sizes in the random conditions, we expected that participants would be more likely to change the icon sizes in the random condition than in the organized condition.

4.1.2 Participants

Forty native English speakers were recruited from a university community. They were randomly assigned to one condition in a 2x2 between participants design.

4.1.3 Interface, Task and Procedure

All participants were given the same set of 13 search tasks, one in practice session and 12 in experiment session. Each task, no matter in practice session or experiment session, had a maximum of 10 minutes for participants to answer the question. In each task, each participant was instructed to imagine that he or she was a doctor of abnormal psychological diseases, and was asked to give out the diagnosis results based on the symptom shown in the question area and the reference information shown in each icon file. Participants were required to use any information available in the icon files, and they were told to not only emphasize accuracy and search time, but also the edit and search strategies they used. We chose this task because we did not want the background knowledge of the participants to be a confounding variable, and during recruitment we ensured that they did not have a psychology background.

Search tasks (see Figure 4.1 for an example) were designed such that participants need to make multiple accesses to icons to find, re-find, and integrate information. For instance, several abnormal psychological diseases have similar symptoms, thus participants have to access and re-access icon files to get the useful information.
The left-hand side of the interface (see Figure 4.2) simulated a computer desktop environment with 36 file icons (white squares in the green panel). Each search task was shown in the top left panel, and right below this panel was the answer box. Based on different conditions, an icon size edit tool (track bar or increase/decrease buttons) was shown in the edit tool area that was between the answer box and the area of 36 file icons.

The right-hand side of the interface was the reference information area. Each time one icon was clicked, it would be highlighted as grey color and at the same time a loading page would be shown in the right-hand side of the interface. The loading page was to simulate the time cost in accessing the information from a database, and it lasted for 1.5 seconds. The loading page would then be replaced by the reference information corresponding to the highlighted icon. Participants could then read the reference information, select another icon, and repeat until they could answer the question. At any time, participants could use the icon edit tool to change the icon size, making it bigger or smaller. Finally, participants could input the answer to the search task into the answer box, and then click submit button to finish the current task and go on to the next task. The whole experiment took about 2 hours.

4.2 RESULTS

4.2.1 Search Performance

Participants’ responses and their completion time to each question were recorded. The answers were graded on a 0-4 grading scale and the total scores were computed. There was no between-group difference in the total scores. The completion time was averaged across all questions for each participant and also compared across groups. Again, we did not find significant between-group difference. Related to the first hypothesis of the first set of research questions – i.e. lower cost and higher benefit of customization will lead to finding and re-finding more accurate information in shorter time – the results of total scores and average time did not prove this hypothesis. It is perhaps because of the low number of questions. Therefore, we would design more questions in later studies, hoping to see significant difference in search performance.
However, we did find significant differences in participants’ customization behavior and their strategies of information indexing.

### 4.2.2 Icon Size Editing Behavior

This section corresponds to the second hypothesis of the first set of research questions, i.e. lower cost and higher benefit of customization will lead to higher willingness to customize. In this section, we analyzed average number of icon size edits and final icon size and got the following results.

**Average number of icon size edits.** We looked at how often participants made icon edits when they were performing the tasks. The number of times participants edited the size of each icon was recorded, and the average number of edits was computed for each participant. Two-way Analysis of Variance (ANOVA) with icon edit cost and icon organization as between-subject variables yielded a significant interaction between the two independent variables ($F_{1,1418} = 20.47, p<.001$), suggesting that the effect of the ease of editing icon size varied based on whether the icons were organized or not (see Figure 4.3).

Figure 4.3 showed the average number of icon edits for each group. Consistent with the theory of soft constraints, results showed that cost had a significant effect on how many times participants edited icons in the random condition. In the random condition, participants did not have consistent spatial cues to guide them to find the correct icon, and thus it would be more beneficial if they could change the sizes of icons to help them index information in the environment. However, the size edit tool was only highly used when it was easy to use (low-cost), whereas increase cost of changing icon sizes dampened participants’ willingness of using of this feature. No significant effect of cost was found on number of edits in the organized condition. Apparently, participants in this condition were able to index information based on its location. Thus their use of the customization tools had a low benefit, and there were therefore fewer icon edits in the organized condition.
Final icon size. We further divided the 36 icons into two groups by identifying a number of icons as more relevant for participants to answer the questions and putting them into the useful group, and the rest into the less useful group. The content of each icon was used for at least one question throughout the experiment. However, for some icons, their contents were useful for answering more than three questions. Those icons were called useful icons. In contrast, the less useful icons were used to answer one question only throughout the experiment.

ANOVA showed different usage patterns for the two icon groups. A significant two-way interaction between cost and organization was found for the useful icons only ($F_{1,1100} = 20.135, p < .01$). Consistent with the results on icon edits, participants in Lo-Random condition utilized the feature of size edit the most by making the useful icons larger compared to participants in other conditions. It was not surprising that random icon location and low edit cost made participants make more use of the icon size edit tools by changing them into larger sizes (see Figure 4.4).

In summary, the manipulations of the two independent variables did have an effect on how often participants made changes to icon sizes and on the final sizes of the icons. Randomization of icon location increased the participants’ need to make use of customization cues other than location to facilitate searching while cost in editing icon size increased the barrier of utilizing the feature. A combination of the two effects made the Lo-Random condition provided the largest motivation for participants to customize the sizes of icons. The above results proved the second hypothesis of the first set of research questions, i.e. lower cost and higher benefit of customization do lead to higher willingness to customize. It implied that we could design novel customization tools with low cost and high benefit to encourage people to customize.

4.2.3 Icon Access Patterns

To test the third hypothesis of the first set of research questions, we then investigated participants’ icon access patterns. There were interesting patterns from the analysis of icon access transitions that informed the use of search strategies. The transition table shows how the frequency of one icon access to the next icon access. For instance, in the transition table of the
Lo-Random condition (see Table 4.1), the number 16 which is in the cell of the fifth row and the sixth column means there were 16 instances in which the participants accessed icon6 right after accessing icon5. In this experiment, four transition tables which correspond to the four experimental conditions (Lo-Random, Lo-Org, Hi-Random, and Hi-Org) were obtained when access frequencies were aggregated for all participants in each condition. A couple of interesting things in the transition tables were observed.

**Forward sequential accesses.** The upper right diagonal in each transition table corresponds to the forward sequential accesses (see Table 4.1). In all of the four conditions we found an obvious pattern of these diagonals, that is, cells on these diagonals have relatively larger numbers (marked in yellow in Table 4.1) which are equal to or greater than 10. ANOVA with between-subject variables showed a significant main effect of icon edit cost on forward sequential accesses \((F_{1, 136}=7.245, p<.01)\) as well as a significant main effect of icon organization \((F_{1, 136}=66.861, p<.001)\). However, the interaction between the icon edit cost and the icon organization was not significant.

Forward sequential accesses took a higher percentage of total accesses in the organized conditions than in the random conditions and in the high cost conditions than in the low cost conditions (see Figure 4.5 left). Consistent with previous results, participants in organized conditions knew the icons were grouped together according to their contents, thus they clicked sequentially to retrieve the most exact answer. But for random conditions, participants had less location cues; therefore they had to make more random attempts that led to the fact that forward sequential accesses took a lower percentage of total accesses. Compared with low cost conditions, in high cost conditions, participants made fewer total accesses to icons, hence the forward sequential accesses had higher percentage of the total accesses.

**Backward sequential accesses.** The lower left diagonal in each transition table stands for the backward sequential accesses (see Table 4.1). In random conditions, we found relatively larger numbers (equal to or greater than 10) in most of these diagonal cells. But in organized conditions, we had no similar findings. ANOVA with between-subject variables showed a
significant main effect of icon organization on backward sequential accesses ($F_{1, 136}=35.859$, $p<.001$). The icon edit cost has a marginal effect ($p=.094$). However, the interaction between cost and organization is not significant.

We found that backward sequential accesses had a higher percentage of total accesses in the random conditions than in the organized conditions (see Figure 4.5 right). This was probably because in random condition participants had less location cues then in organized condition, thus they tended to be less likely to sequentially search for information; but when they did, they were equally likely to search it either direction (i.e. either the icon on the left or on the right). However, in the organized condition, the locations of icons stayed the same, and an intuitive strategy was to start from the top-left corner and sequentially clicked on each icon to find the relevant content. Results suggested that in the organized condition, information indexing tended to rely more on the location cues, which guided them to search sequentially; however, in the random condition, participants would more likely start their search by clicking on a specific icon by, for example, recognizing their sizes, then searched for icons nearby. While the exact strategy used could vary across individuals, apparently the differences in the benefit of customization had led to a difference in the information indexing strategies. Basically, the above results proved the third hypothesis of the first set of research questions, i.e. lower cost and higher benefit of customization will lead to more diverse information indexing strategies. It is meaningful for us to design better customization tools to help people freely develop their own information indexing strategies.

4.3 CONCLUSION AND DISCUSSION

Consistent with the theory of soft constraints, participants in random conditions tended to make more icon edits when the cost of editing was low. Different usage patterns for the useful and less useful icons were found. For the useful icons, participants in Lo-Random condition utilized the feature of size editing the most by making the useful icons larger compared to participants in other conditions. This finding could also be explained by cost-benefit tradeoff considering that in this condition the cost of editing is low and the benefit of editing was high. A couple of interesting things were observed in the icon access transition tables that informed the use of
search strategies. Due to different reasons, forward sequential accesses took a higher percentage of total accesses in the organized conditions than in the random conditions and in the high cost conditions than in the low cost conditions; backward sequential accesses took a higher percentage of total accesses in the random conditions than in the organized conditions. Although failed to prove the first hypothesis of the first set of research questions, the results did prove the second and third hypotheses of the first set of research questions.

Future studies should further decrease the cost of icon edits and increase the benefit in order to encourage the participants to make more edits on the icon sizes. Derived from the finding of Mackay (1991) that users would more likely customize when they discovered that they would do something repeatedly, in future studies we plan to design more questions to increase the frequency of using each icon, hoping to make participants easily realize that each icon will be accessed again in the future to increase their motivation to customize. By increasing the number of questions (and their difficulty), we also hope to find out how increase in customization may lead to differences in information indexing strategies, which in turn may help people to improve their performance in finding and re-finding information.
CHAPTER 5
THE SECOND EXPERIMENT

To test the hypotheses of the second set of research questions mentioned in Chapter 3, we conducted the second experiment. Participants were required to perform one individual desktop search task with a set of desktop icon edit tools. We hoped to see how customization tools and different structures of information needs interact to influence people’s search performance, customization behavior, search behavior and information indexing strategies.

5.1 METHOD

5.1.1 Task and Interface

In this experiment, each participant was required to search for information in a computer desktop environment in order to answer 144 questions in 75 minutes. An example of the desktop interface was shown in Figure 5.1.

Each question was shown in the question area located in the upper left area of the screen. An example of the question was “Suppose you will travel to Seattle, and the cost of the travel is $300. What is the name of the operator?” Right above the question area was a clock showing how much time was left for the current question, and below the question area was an answer box and a Begin/Submit/Next button. The icon area containing thirty-six icons was shown in the lower left of the screen. Each icon corresponded to one travel plan. Located between the answer box and the icon area was an area that showed a tool bar (including 6 color buttons, 6 size buttons, and a sort button) in some experimental conditions. The top right area of the screen was the content area displaying the details of the travel plan.

During the experiment, every time a subject clicked on one icon, after 2 seconds’ loading time, he/she would see the corresponding travel plan. The travel plan (i.e. information) included 3 non-
5.1.2 Experimental Design

In this experiment, a 2 (tools) × 2 (dominance) × 2 (dispersion) between-subjects factorial design was used. The first factor (tools) was about whether the subjects were provided with the customization tools or not. In the with-dominance condition, most of the questions (129 out of 144) had the same pair of given dimensions (e.g., location=Seattle and cost=$300). In the without-dominance condition, the questions contained random pairs of given dimensions. In the with-dispersion condition, answers to most of the questions could be found in a smaller subset of icons than the without-dispersion condition. Specifically, in the with-dispersion condition, there were 6 most useful icons, each of which contained answers to 11 questions; 12 medium useful icons, each of which contained answers to 5 questions; and 18 least useful icons, each of which contained answers to one question. In the without-dispersion condition, there were 6 most useful icons, each of which contained answers to 5 questions; 24 medium useful icons, each of which 4 questions; and 6 least useful icons, each of which contained answers to 3 questions.

5.1.3 Participants

Ninety-six college students (Mean age=20.6, S.D.=1.8; 46.9% female) recruited from a university community were randomly assigned to one of the 8 experimental conditions.

5.1.4 Procedure

For each question, the subjects needed to find and re-find information by clicking on the icons on the screen and getting the requested information from the corresponding travel plans on the right side of the screen, typed the answer into the answer box, and then clicked the Submit button.
After confirming the submission in a pop-out window, subjects could click the Next button to see the next question. When a subject correctly answered one question, he/she would gain one point.

In the with-tool condition, a set of customization tools were provided. Participants could use one of the 6 color buttons and/or one of the 6 size buttons to change the color and/or the size of any icon. They could also use the Sort button to sort all icons according to their colors and sizes (all icons with same color would be in the same row, with the smallest size one on the left most of the row, the next smallest icon (with the same color) to the right of it, and so on). Participants could use the customization tools anytime they wanted to during the experiment.

5.2 RESULTS

5.2.1 Performance

This section tested the hypothesis set 1 of the second set of research questions. We wanted to investigate how customization tools, dominance and dispersion structures of information needs influence the performance of finding and re-finding information. In this section, we analyzed total score and average time, and got the following results.

Total score. ANOVA (Analysis of Variance) on total score showed significant main effects of tools (F(1,88)=118.98, p<.001) and dominance (F(1,88)=129.47, p<.001). The 2-way interaction effects of tools with dominance (F(1,88)=8.19, p<.01) and tools with dispersion (F(1,88)=8.82, p<.005) were also significant (see Figure 5.2). However, the main effect of dispersion, the 2-way interaction between dominance and dispersion, and the three-way interaction tools × dominance × dispersion was not significant.

As shown on the left side of Figure 5.2, tools and dominance interacted to affect total score. Tools helped subjects to obtain higher total scores both in the without-dominance (p<.001) and with-dominance (p<.001) conditions, but the difference was larger in the with-dominance condition. Dominance also helped subjects to obtain higher total scores both in the without-tools
(\(p<.001\)) and with-tools \((p<.001)\) conditions, but the difference was larger in the with-tools condition than in the without-tools condition.

Most questions in the with-dominance condition were designed to use the same pair of given information dimensions to look for the unknown information, and this would guide subjects to utilize the same given dimensions as indices to the travel plans. When this consistent structure in information needs were supported by the customization tools, the improvement was much larger than when the tools were not provided, suggesting that the tools were more useful when participants could offload and represent the consistent structures to the external environment by the customization tools.

As shown on the right side of Figure 5.2, tools and dispersion also interacted to influence performance. While tools helped subjects to obtain higher total scores both in the without-dispersion \((p<.001)\) and with-dispersion \((p<.005)\) conditions, the difference was larger in the without-dispersion than in the with-dispersion condition. Interestingly, dispersion helped subjects to obtain higher scores in the without-tools condition \((p<.05)\), but not in the with-tools condition \((p=.49)\).

In the with-dispersion condition, because questions were concentrated on a small subset of icons, the subjects could create memory indices to these most useful icons to retrieve answers to a high number of questions even without the help of customization tools (more analysis on this later). However, this was more difficult in the without-dispersion condition, and thus performance was worse. Interestingly, while tools helped subjects in the without-dispersion condition, performance was not significantly better with-dispersion compared to the without-dispersion condition when customization tools were provided. The results suggested that the effects of dispersion and customization tools were not additive, presumably because the tools did not help subjects to offload the information structures created by the dispersion of information distribution. We will come back to this point later when we analyzed the patterns of icon clicks.

**Average time spent per question.** ANOVA on the average time spent for answering one question showed significant main effects of tools \((F(1,88)=107.50, \ p<.001)\) and dominance
(F(1,88)=101.62, \( p<.001 \)) as well as the 2-way interaction between tools and dispersion (F(1,88)=7.08, \( p<.01 \)) (see Figure 5.3). The main effect of dispersion was found to approach significance (\( p=.08 \)), however, the 2-way interaction effects between tools and dominance, and between dominance and dispersion, and the 3-way interaction tools × dominance × dispersion were not significant.

As shown on the left side of Figure 5.3, the parallel lines showed that there was no interaction effect between tools and dominance, suggesting that, when measured by time, the tools helped subjects equally with or without the dominance structures. However, given that subjects in the with-dominance condition obtained much higher scores when augmented with the customization tools, the results confirmed that the better performance was not caused by participants spending more time on finding the correct icons, but rather, they achieved better performance (finding the right icons) with less time.

As shown on the right side of Figure 5.3, subjects spent less time with tools than those without tools in the without-dispersion (\( p<.001 \)) and with-dispersion (\( p<.001 \)) conditions, but the difference was larger in the without-dispersion than in the with-dispersion condition. Simple effect analysis also showed that subjects in the with-dispersion condition spent significantly less time in the without-tools condition (\( p<.05 \)), but not in the with-tools condition (\( p=.65 \)). Consistent with results on total scores, when the customization tools were available, dispersion of information distribution did not lead to significant differences in reaction time. However, without tools, dispersion did help subjects to finish each question faster, suggesting that subjects were able to obtain higher scores using less time. As we will show later in the analysis of icon clicks, subjects seemed to have learned to mentally index the most useful icons in the with-dispersion condition, which allowed them to click on these most useful icons to answer more questions correctly then the without-dispersion condition. This advantage, however, seemed to be redundant when the customization tools were available, which apparently allowed subjects to offload indexing of these icons to the external environment. Next, we will present our analysis on how subjects used the customization tools to help them to externalize indexing to the environment, after which, we will present results on icon clicks, which reflected how different
structures of information needs could influence the balance between internal and external indexing of information.

### 5.2.2 Customization Behavior

We found that 83% of the subjects in without-dispersion with-dominance condition (10 out of 12 subjects) and 100% of the subjects in with-dispersion with-dominance condition (12 out of 12 subjects) generated similar icon area patterns, in which they used colors to represent the dominant non-quantitative dimensions (e.g., city) and sizes to represent the dominant quantitative dimensions (e.g., cost). When none of the dimensions was dominant, subjects tended to use the colors tool to equally often index one of the non-quantitative dimensions and the sizes tool to equally often index one of the quantitative dimensions.

To test the hypothesis set 2 of the second set of research questions, we analyzed the raw data and obtained the interesting finding as follows.

**Average number of clicks on the customization tools.** We divided the number of clicks on the tools by the number of questions answered to calculate the average number of clicks on the tools. The patterns for clicks on the colors tool, the sizes tool, and the “sort” tool were similar, so we chose to report only the average number of clicks on the sizes tool, which measured how often the subjects customized the sizes of the desktop icons. ANOVA showed significant main effects of dominance (F(1,88)=16.10, p<.001) and dispersion (F(1,88)=5.04, p<.05). Moreover, the interaction between dominance and dispersion was also significant (F(1,88)=4.81, p<.05) (see Figure 5.4).

As shown in Figure 5.4, dominance and dispersion interacted in affecting the subjects’ behavior of clicking on the sizes tool. Dominance made the subjects clicked less on the sizes tool in the with-dispersion condition (p<.005), and the effect of dominance was marginally significant in the without-dispersion condition (p=.07), while dispersion made the subjects clicked more on the sizes tool in the without-dominance condition (p<.05), but not in the with-dominance condition (p=.92).
It was interesting to note that dominance led to fewer clicks on sizes tool. However, less customization did not necessarily mean a worse performance. The above analysis already showed that dominance structures helped subjects spend less average time getting higher total scores. We believed that dominance structures were easy to be perceived by the subjects. After learning this knowledge, the subjects could use the tools to offload indexing of the icons to the external environment. Upon obtaining complete city-cost (the pair of dominant dimensions) index, the next step for the subjects to do is just to utilize the index to answer questions. However, subjects in without-dominance conditions could not find dominant (or “indicative”) dimensions asked by the questions. Thus they kept customizing the icon sizes to offload their known indexing of the icons to the external environment. As a result, subjects made significantly more customization on icon sizes for each question in the rest time than those in the with-dominance condition. To further understand how the strategies differed in early vs. later trials, we conducted analysis based on behavior when they answered the first 10 vs. later questions.

**Clicks on the sizes tool in the first 10 questions.** We found 60% of the 48 subjects in with-tools conditions made 50% or more clicks on the sizes tool when they answered the first 10 questions. ANOVA on average number of clicks on sizes tool in the first 10 questions showed that none of dominance, dispersion, and their interaction was significant (see Figure 5.5).

As shown in Figure 5.5, dominance had no effect neither in the without-dispersion nor the with-dispersion condition, and dispersion had no effect no matter dominance structures existed or not. Both subjects in the with- and without-dominance conditions worked actively to offload indexing of the icons to the external environment in the very beginning of the experiment, and there were no significant difference regarding their customization on icon sizes. Results suggested that they were using the customization tools equally as they were adapting to the information structures during the early trials.

**Clicks on the sizes tool in the rest of the questions.** ANOVA on average number of clicks on the sizes tool in the rest of the questions showed a significant effect of dominance ($p<.05$), while dispersion ($p=.09$) and the interaction between dominance and dispersion approached significance ($p=.08$) (see Figure 5.6).
As shown in Figure 5.6, dominance was significant in the with-dispersion condition ($p<.05$) but not in the without-dispersion condition ($p=.52$), while dispersion was marginally significant in the without-dominance condition ($p=.08$), but not in the with-dominance condition ($p=.84$).

To further analyze the data, we also conducted t-tests to compare the total number of clicks on the sizes tool in the rest of the questions in different conditions, and no significant difference was found. In the meantime, we found that when tools were available, dominance led to more questions answered after the first 10 questions ($F(1,44)=19.48$, $p<.001$) and the interaction between dominance and dispersion was not significant. Results suggested that when tools were available, after the first 10 questions, the subjects in the with-dominance condition made similar total number of clicks on the sizes tool in relevantly more questions, and therefore they had lower average number of clicks on the sizes tool for each of the rest questions. It matched the results showed in Figure 5.6. The subjects in the with-dominance with-dispersion condition offloaded much indexing to the external environment in the first 10 questions, then they did not need to customize much for each of the rest questions, but rather, just utilize the indices to find the right answers, while subjects in the without-dominance with-dispersion condition had to keep attempting a better way to externalize known information to the environment throughout the experiment.

5.2.3 Search Behavior

In this section, we tested the hypothesis set 3 of the second set of research questions, and obtained the following results.

**Average number of clicks on the icons.** We calculated the average number of clicks on the icons by dividing the number of clicks on the icons by the number of questions answered. ANOVA showed significant main effect of tools ($F(1,88)=209.01$, $p<.001$) and dominance ($F(1,88)=119.83$, $p<.001$), as well as the 2-way interaction between tools and dispersion ($F(1,88)=4.55$, $p<.05$) (see Figure 5.7). The main effect of dispersion approached significance ($p=.09$), however, no other effect was significant.
As shown on the left side of Figure 5.7, the parallel lines indicated that there was no significant interaction between tools and dominance. Simple effect analysis showed that the tools led to less clicks on the icons both in the without-dominance ($p<.001$) and with-dominance ($p<.001$) conditions, and dominance led to less clicks on the icons both in the without-tools ($p<.001$) and with-tools ($p<.001$) conditions. Results showed that subjects were better at picking the right icons with the tools and when there were dominance structures in the questions.

As shown on the right side of Figure 5.7, tools led to less clicks on the icons both in the without-dispersion ($p<.001$) and with-dispersion ($p<.001$) conditions, but the difference is larger in the without-dispersion condition than in the with-dispersion condition, making the 2-way interaction significant. However, dispersion had no significant effect both in the without-tools ($p=.11$) and with-tools ($p=.82$) conditions. Consistent with previous results, dispersion led to fewer clicks to finish the trials when the tools were not available. In addition, we found again that there was no significant difference with-dispersion compared to the without-dispersion condition when customization tools were provided, suggesting that the effects of dispersion and customization tools were not additive. The tools probably did not help subjects to offload the information structures created by the dispersion of information distribution.

**Correct clicks on the most useful icons.** To better understand search behavior, we calculated the percentages of clicks on the most useful icons when answers to the questions could be found in these icons (i.e., correct choice of icons). ANOVA on this measure showed significant main effects of tools ($F(1,88)=191.28$, $p<.001$), dominance ($F(1,88)=109.90$, $p<.001$) and dispersion ($F(1,88)=6.56$, $p<.05$). The 2-way interaction tools with dominance ($F(1,88)=58.72$, $p<.001$), and dominance with dispersion ($F(1,88)=9.88$, $p<.005$) were also significant (see Figure 5.8). However, the 2-way interaction between tools and dispersion did not reach significance, and the 3-way interaction of tools $\times$ dominance $\times$ dispersion was marginally significant ($p=.085$).

As shown on the left side of Figure 5.8, tools led to higher percentages of clicks on the most useful icons both in the without-dominance ($p<.001$) and with-dominance ($p<.001$) conditions, but the difference was larger in the with-dominance condition than in the without-dominance condition. Dominance also lead to more clicks on the most useful icons both in without-tools
(p<.001) and with-tools (p<.001) conditions, but the difference was larger in the with-tools condition than in the without-tools condition. Results showed that with the customization tools, subjects were much more likely to index the most useful icons (the most useful icons were correct icons in the current situation) than when no tools were available, and this effect was much stronger when there were dominance structures in the questions. It suggested that the effects of tools and dominance were additive. The external and internal indexing of information worked together for the subjects to better understand the desktop environment. In addition, the tendency also showed that when tools were not available, dominance structures led to more indexing effort to the most useful icons. It reflected that when without tools, subjects would rely more on the structures of information needs to mentally index information.

As shown on the right side of Figure 5.8, there was a significant interaction between dominance and dispersion. Dominance led to more clicks on the most useful icons both in the without-dispersion (p<.01) and with-dispersion (p<.001) conditions, but the difference was larger in the with-dispersion condition than in the without-dispersion condition. Although there was a significant main effect of dispersion, simple effect analysis showed that, however, the effect of dispersion was not significant in the without-dominance (p=.69) and with-dominance (p=.12) conditions. The results showed that the interaction effect was mainly caused by the significantly higher number of clicks on the most useful icons when there were both dominance and dispersion in the information needs, suggesting that both structures were useful for subjects to internally and externally index information.

Wrong clicks on the most useful icons. We also calculated the percentages of clicks on the most useful icons when the answers to the questions could be found in the least useful icons (i.e., wrong choice of icons). Assuming that subjects were sensitive to the information structures and would prioritize their indexing effort to the most useful icons, this measure reflected the extent to which subjects failed to index the least useful icons. ANOVA showed significant main effects of tools (F(1,88)=31.89, p<.001) and dispersion (F(1,88)=8.61, p<.005). The 2-way interaction effects of tools × dominance (F(1,88)=9.14, p<.005) and tools × dispersion (F(1,88)=4.06, p<.05) were also significant (see Figure 5.9). The 3-way interaction effect tools× dominance ×
dispersion was significant \( F(1,88)=4.88, p<.05 \) too. However, the main effect of dominance and the 2-way interaction effect between dominance and dispersion were not significant.

As shown on the left side of Figure 5.9, tools led to fewer wrong clicks both in the without-dominance \( (p<.05) \) and with-dominance \( (p<.001) \) conditions, but the difference was larger in the with-dominance condition. Dominance structures led to fewer wrong clicks when tools were provided (marginally significant, \( p=.08 \)) but more wrong clicks when there were no tools \( (p<.05) \). The trend clearly showed that with customization tools, subjects were much less likely to index the most useful icons (the most useful icons were wrong icons in the current situation) than when there were no tools. It reflects that the subjects successfully used the tools to externalize useful information to the environment. Moreover, this correct indexing was reinforced when dominance structures existed. Again, it implied that tools and dominance interacted in affecting internal and external information indexing, and the effects of them were additive. In addition, dominance structures were again found to lead to more internally indexing effort of the most useful icons when no tools existed.

As shown on the right side of Figure 5.9, tools led to fewer wrong clicks both in the without-dispersion \( (p<.001) \) and with-dispersion \( (p<.005) \) conditions, but the difference was larger in the without-dispersion than in the with-dispersion condition. Simple effect analysis also showed that dispersion had significant effect in the with-tools condition \( (p<.01) \), but not in the without-tools condition \( (p=.39) \). Interestingly, the subjects in the with-dispersion condition clicks significantly more wrong clicks compared to the without-dispersion condition when the tools were available. Dispersion of information distribution allowed the subjects to click on the most useful icons to answer more questions correctly. After the subjects learned this, they would tend to mentally index more information regarding the most useful icons. This tendency was still strong even though the tools existed. As a result, it led to more wrong clicks. Again, it suggested that the effects of tools and dispersion were not additive. The information structures created by the dispersion of information distribution could not be offloaded to the external environment under the help of tools, but rather, it led to more internal indexing of information and may finally disturb the subjects’ external indexing of known information by the customization tools.
5.2.4 Icon Access Patterns

In this section, we tested the hypothesis set 4 of the second set of research questions and reported the following results.

We found interesting patterns from the analysis of icon access transitions that indicated the use of internal and external information strategies. The transition table shows the frequency of one icon access to the next icon access. For example, in the transition table of participant 1 in condition 1 (see Table 5.1), the number 5 in the cell of the fifth row and the eleventh column means there were 5 instances in which the participants accessed icon11 right after accessing icon5. In this experiment 96 transition tables corresponding to 96 participants were obtained. Some interesting findings in the transition tables were observed.

As shown in Table 5.1, the cells on four diagonals are marked to represent horizontal forward sequential accesses, horizontal backward sequential accesses, vertical forward sequential accesses, and vertical backward sequential accesses. For each of the 96 transition tables, we added up all numbers in these marked cells to calculate the total number of sequential accesses. Finally, we found that sequential accesses account for a pretty high percentage of total transition accesses (Mean=67%, S.D.=14%). ANOVA on sequential accesses showed significant main effect of tools (F(1,88)=4.28, p<.05) and dominance (F(1,88)=4.09, p<.05). The 2-way interaction dominance with dispersion was marginally significant (p=.08). However, none of other effects were significant.

As shown in Figure 5.10, tools led to higher percentage that sequential accesses account for total transition accesses in the with-dominance condition (p<.005), but not in the without-dominance condition (p=.72). Dominance led to lower percentage that sequential accesses account for total transition accesses in the without-tools condition (p<.05), but not in the with-tools condition (p=.68). Results showed that when dominance structures exist, if tools are available, participants would utilize the tools to adapt to the dominance structures. As a result, they could develop an external indexing of information to offload information to the external environment and after that they could fully use the external indexing to perform a sequential search in order to find useful
information efficiently. However, when there were no dominance structures in the questions, even though tools were provided, the usefulness of the tools was greatly weakened. Participant could not develop an appropriate external indexing of information to adapt to dominance structures just like what they did when dominance structures exist, and thus it led to the insignificant results. Results also showed that in the without-tools condition, dominance structures led to lower percentage that sequential accesses account for total transition accesses. It implied that when tools are not available, participants tend to develop more diverse internal indexing of information. As a result, they perform more random search (i.e. access one icon, and then jump to another icon which is not adjacent to the previous one) instead of sequential search. The above results are interesting and meaningful because they made us better understand how people would make use of customization tools to develop their own internal and external indexing of information when dominance structures of information needs exist or not. It will help us design better tools for participants to adapt to different structures of information needs.

5.2.5 Post Test

One-click test score. After the main session of the experiment, subjects performed a one-click test that directly measured how well they had encoded indices to the icons. During the test, subjects would answer 18 questions as before, but they did not need to input the answer into answer box. Instead, they needed to decide which icon contained the correct answer and clicked on that icon. If the icon they clicked contained the answer, they would get one point. ANOVA on the one-click test score showed significant effects of tools ($F(1,88)=38.53, p<.001$), dominance ($F(1,88)=6.97, p<.05$) and the interaction between dominance and dispersion ($F(1,88)=4.89, p<.05$) (see Figure 5.11). No other effect was significant.

As shown on the left side of Figure 5.11, subjects with tools obtained higher one-click test scores than those without tools both in the without-dominance ($p<.001$) and with-dominance ($p<.001$) conditions, confirming that they were better at indexing information with the customization tools. Interestingly, indexing was significantly better when there were dominant structures in the without-tools condition ($p<.05$), but not in the with-tools condition ($p=.11$). In the with-tools condition, the subjects could use tools to offload the indexing of icons to the external
environment, and thus the effect of dominance was not obvious. However, when tools were not available, dominance structures encouraged subjects to make more mental indexing of the icons. Another interesting point was that there were only two questions out of the 18 questions were designed asked about dominant dimensions, but looking at the results, subjects in the with-tools with-dominance condition even got scores close to 4 points. It suggested the existence of both internal and external indexing of information regarding the icons.

As shown on the right side of Figure 5.11, dominance made the subjects obtain higher one-click test scores in the with-dispersion condition ($p<.001$), but not in the without-dispersion condition ($p=.83$), while dispersion had no significant effect on one-click test score both in the without-dominance ($p=.17$) and with-dominance ($p=.22$) conditions. The 2-way interaction was mainly caused by the subjects’ good performance in the with-dominance with-dispersion condition. It suggested that dominance and dispersion interacted in affecting internal and external indexing of information.

**Score for two particular questions.** In one-click test, two out of the 18 questions were asked about the two dominant dimensions. We then directly analyzed these two questions to see whether dominance structures influenced indexing of information. ANOVA on the score of the two questions showed significant effects of tools ($F(1,88)=26.29$, $p<.001$), dominance ($F(1,88)=70.96$, $p<.001$), and the interaction between tools and dominance ($F(1,88)=13.42$, $p<.001$) (see Figure 5.12). No other effect was significant.

As shown in Figure 5.12, dominance structures did affect the score of the two questions both when tools were available ($p<.001$) and unavailable ($p<.005$). In addition, tools led to better performance in the with-dominance condition ($p<.001$), but not in without-dominance condition ($p=.40$). We found that the subjects in the with-tools with-dominance condition reached the scores of exactly 2 points (full mark), suggesting that subjects fully made use of tools to index dominance structures. Once more, it implied that the effects of tools and dominance were additive.
Results showed that the customization tools were in general useful for indexing information, but their effects were stronger when subjects could offload information indexing which represented the dominance and dispersion structures using customization cues provided by the tools. We found that all subjects were using the customization tools extensively as they were adapting to the information structures during the early trials, but they did not customize much for each of the rest questions especially in the dominance conditions, as the structures remain stable throughout the experiment. Interestingly, subjects in the without-dominance with-dispersion condition showed the largest number of use of the customization tools in the experiment, presumably because participants perceived structures in the questions (that they tended to point to the small subset of icons), but because the indexing dimensions were changing throughout the experiment, they had to keep adapting to the “structures” by indexing different dimensions throughout the experiment. This showed how a mismatch of tools and information structures could lead to poor adaptation to the environment.

In general, better performance was achieved when the customization tools were available. However, performance was much better in the dominance conditions than other conditions: tools did not seem to improve performance in the dispersion conditions, as reflected by the similar performance between the with- and without-dispersion conditions. This pattern of results suggested that the match between the tools and information structures was critical for performance. When tools were not available, both structures helped participants to achieve better performance, suggesting that mental indexing was effective in helping subjects to re-find information, although the limited memory capacity had apparently put a cap on performance.

Further analysis on distribution of icon clicks confirmed that indexing was the best in the dominance and dispersion conditions when tools were available, as reflected by the higher number of correct clicks and lower number of wrong clicks on the most useful icons. This again confirmed that customization tools were most effective when they could be utilized to offload indexing of information structures to the external environment. However, even without the tools, structures of information needs (both dominance and dispersion) were still useful for information
indexing. Dominance and dispersion structures interacted to lead to higher internal and external indexing of information.

Results from the post-tests of information indexing showed that even in the with-tools condition, subjects had developed mental indexing of information. In particular, in the dominance condition, subjects could correctly click on the correct icons even when the questions were not using the dominant dimensions, suggesting that they were able to develop a mix of external and internal indexing strategies to help them find the correct icons. This could also explain why there was an interaction between dominance and tools in the final performance, in which the tools helped performance significantly more in the dominance condition. Similarly, when the dispersion structures existed, subjects were much better at picking out the correct icons even without the tools, suggesting that they had encoded, at least partially, which icons were most frequently requested. When the dispersion structures did not exist, indexing apparently was much worse, as subjects could not index all icons that were equally requested. Results suggested that people were sensitive to both information structures when they adaptively learn to select icons for internal indexing.

To summarize, we found that structures of information needs interacted with the use of customization tools to influence finding and re-finding of information. In fact, we found complex interactions between the three factors that we manipulated, which influenced not only how the subjects used the customization tools, but also how they searched and internally indexed icons. The current experiment provided a good understanding of both internal and external indexing of information, and we found the dynamic shift between them seemed to depend much on the different structures of information needs. In fact, our results seemed to suggest that, there is much benefit for customization tools to be designed with respect to the statistical structures of information needs, in addition to the more tangible information dimensions or features. Future research could address how other structures could influence indexing, and how different customization tools that are specifically designed for indexing these structures could be designed.
CHAPTER 6
THE THIRD EXPERIMENT

To test the hypotheses of the third set of research questions mentioned in Chapter 3, we conducted the third experiment. In this experiment, every two participants would work as a group to perform a collaborative desktop search task. The two group members were required to use MSN to communicate, typing text message (in the text condition) or speaking to each other (in the audio condition). However, each of the two group members would operate one individual computer, using a set of customization tools to help him/her find, re-find, index, and share desktop information. In addition, we designed the experiment such that the help one group member (denoted as “side B”) needed from the other group member (denoted as “side A”) was much more than the help “side A” needed from “side B”. By the above design, we hoped to find that people’s search performance and communication behavior would change based on different communication methods and different amounts of help needed by different group members.

6.1 METHOD

6.1.1 Task and Interface

In this experiment, every two participants (indicated as side A and side B) composed one group. While each of the two group members operated one individual computer to act as a travel agent, they were required to use MSN to communicate with each other to share information and answer simulated customers’ questions (i.e. provide the customers with travel suggestions). During the experiment, each participant – no matter side A or side B – was required to answer 144 questions in 80 minutes. However, the sequence of the 144 questions was designed to be different for side A and side B. In addition, for both side A and side B, the correct answers to the first 36 questions could be found on their own desktops, so they did not need their partner’s help for those questions. However, from question 37 to the end of the whole task, side A would meet 48 questions, the correct answers to which could only be found on his/her partner’s desktop and side
B would meet 96 such questions. By designing like this, we made different group members (i.e. side A and side B) need different amount of help during the experiment. We expected to see this kind of “different amount of help needed by different group members” in a collaborative search task would greatly influence group members’ search performance and customization behavior.

An example of the desktop interface was shown in Figure 6.1. Each search question was shown in the upper left area of the screen. An example of the question could be “Suppose the customer wants to travel to Chicago, and he/she hopes to spend exactly $100. Which plan should you suggest?” Icon area containing thirty-six icons was shown in the lower left area. Each icon corresponds to one travel plan. Located between the above two areas was an answer box, six help-rate radio buttons (They were used to rate how much help a participant got from his/her partner for the current question), a begin/submit button, six color buttons, six size buttons, and a sort button. The top right area of the screen was used to show this participant’s accumulative score. Right below the score area is the content area displaying the detailed content of the travel plan.

During the experiment, every time a subject clicked on one icon, after 2 seconds’ loading time, he/she would see the corresponding detailed travel plan shown on the content area. A typical example of travel plan (i.e. information) could be “Icon: hek, Location: Chicago, Purpose: relaxation, Cost ($): 100, Duration: 5 day(s), Details: …” This information was displayed in separate rows. The first row was a 3-letter icon title which could be used as the answer to a question if this icon’s corresponding content matched the requirements of customer in that question. The second row was the travel destination city. There were six cities involved in this experiment: New York, Los Angeles, Chicago, Houston, Seattle, and New Orleans. The third row was the purpose for the travel. Six purposes including going shopping, going to the museums, going to the zoos, listening to the operas, relaxation, and celebration were relevant in the experiment. The fourth row showed travel cost. Twelve different costs were designed: $100, $200, $300… $1100, and $1200. However, only six different costs were involved in each participant’s computer desktop. The fifth row was the travel duration (how long the customer would stay in the destination city). Similarly to the cost, the duration also had twelve different
levels: 1 day, 2 days, 3 days… 11 days, and 12 days. And each participant’s computer desktop only included 6 different levels. The last row was the details of the travel plan.

### 6.1.2 Experiment Design

We made a 2 (method) x 2 (side) between-subjects factorial design. The first factor was communication method. There are two communication methods: text and audio provided in this experiment. While in the text condition the two group members could use MSN to send text message to each other in order to share information, in the audio condition the two group members could use MSN (with headphones and speakers) to talk to each other, sharing information. The second factor was side, indicating side A or side B. Since in this experiment, side B needed much more help than side A, the factor side actually indicated different amount of help that different group members needed.

### 6.1.3 Participants

Forty-eight college students (Mean age=21.2, S.D.=1.7; 48.3% female) were recruited from a university community, composing 12 text groups and 12 audio groups, respectively. Each group, no matter text or audio, included two participants. The participants were required not to be color-blind and be between 18 and 35 years old. Finally we collected data from all of the 48 participants.

### 6.1.4 Procedure

In the experiment, for each question participants needed to find and re-find information by clicking on the icons on the screen and getting the requested information from the corresponding travel plans on the right side of the screen. If he/she believed the answer to the current question was not on his/her own desktop, he/she could use MSN to ask his/her partner for the answer. Upon obtaining the answer, he/she could input the answer into the answer box, clicked one help-rate radio button to indicate how much help got from his/her partner for the current question, and
then clicked the Submit button. After confirming the submission in a pop-out window, participants could click the Next button to see the next question.

A set of customization tools consisting of six color buttons (white, gray, yellow, red, light blue and navy blue), six size buttons (from 35x35 pixels to 110x110 pixels) and a sort button was provided in this experiment. When subjects adjusted the color and/or size of an icon, they needed to first click the icon, and then clicked the color and/or size buttons. When subjects clicked the sort button, the travel plan icons would be re-located: the icons with the same color would be grouped together (i.e. be put adjacently) and in one color group (the sequence of color groups was white group, gray group, yellow group, red group, light blue group and navy blue group), the sequence of icons would be from the smallest to the largest. Participants were encouraged to use customization tools to help them mark accessed icons for later use.

Customer acceptance rate was another feature of interest in this experiment. Every time a participant submitted an answer (i.e. provided a travel suggestion), he/she could see immediately a message box pop out to show whether the customer accepted the suggestion or not. For example, if the customer wanted to go to Chicago and the expected cost was exactly $600, then the travel suggestion corresponding to Chicago and $600 would be absolutely accepted by the customer (the customer behavior was simulated by a computer program). We called these travel suggestions absolutely correct answers or optimal answers. However, if the travel suggestion submitted indicated the correct travel destination – Chicago in this example – but the cost was not exactly $600, it would still have some chance to be accepted. Specifically, if the cost difference from $600 was no more than $200 (i.e. the cost was among $400, $500, $700 or $800), then the travel suggestion would be accepted by the customer with 70% acceptance rate. We therefore called these travel suggestions sub-optimal answers with 70% acceptance rate. If the cost difference was more than $200, then the travel suggestion would only have 30% probability to be accepted, and we thereby called these travel suggestions sub-optimal answers with 30% acceptance rate. However, if even the travel destination was wrongly provided, the travel suggestion would be absolutely rejected by the customer, so we call these travel suggestions absolutely wrong answers. Every time a travel suggestion was accepted by the customer, one point would be counted towards the participant’s total score.
6.2 RESULTS

This experiment described a collaborative desktop search task, we therefore would like to shift most of our attention from the analyses regarding individual customization behavior and search behavior to communication behavior although we conceded that this experiment still required participants to use the customization tools to search information. By analyzing the data, we hoped to test the hypotheses of the third set of research questions mentioned in Chapter 3, i.e. reveal how different communication methods and different amounts of help needed by different group members would influence group members’ search performance, customization behavior, and communication behavior. The results are shown as follows.

6.2.1 Performance

To test the hypothesis set 1 of the third set of research question, we analyzed the data and got the following results regarding total score and average time which indicated the performance of the search task.

**Total score.** ANOVA (Analysis of Variance) on total score yielded a significant main effect of communication method \((F(1,44)=14.80, \ p<.001)\), but not for side and the interaction between method and side (see Figure 6.2).

As shown in Figure 6.2, audio condition led to higher total scores than text condition did both for side A \((p<.05)\) and side B \((p<.005)\). Although the difference was larger for side B than for side A, it was not large enough to make the 2-way interaction between method and side significant. However, side A and side B had no significant difference in affecting total score both in the text \((p=.32)\) and audio conditions \((p=.88)\).

It was not surprising that audio condition help participants achieve higher total score since in the audio condition, participants could use MSN to talk to each other. Compared to communicating with the partner by typing text message, directly talking to the partner was much efficient.
**Average time spent per question.** ANOVA on the average time spent for answering one question also showed a significant main effect of communication method (F(1,44)=13.94, p<.005). However, the main effect of side and the interaction method with side are not significant (see Figure 6.3).

As shown in Figure 6.3, the two lines were almost parallel, suggesting that there was no interaction effect between communication method and side. When measured by average time, the audio condition seemed to equally help side A (p<.05) and side B (p<.01). Simple effect analysis also showed side A and side B had no significant difference in affecting average time both in the text (p=.62) and audio conditions (p=.84). It proved the insignificant main effect of side in ANOVA.

Given that participants in the audio condition obtained much higher scores, the results confirmed that the better performance was not caused by participants spending more time on finding useful information, but rather, they achieved better performance of finding and re-find information with less time.

To summarize, the above results showed that communication method did influence participants’ performance of finding and re-finding information by help participants obtain higher total scores in less average time, and therefore it proved the correctness of the first hypothesis of the hypothesis set 1. Though more communication devices are required (e.g. headphones and speakers) in the audio condition, the higher performance achieved looked worth the expenses. Additionally, the results did not show that the factor side – i.e. “different amount of help needed by different group members” – influence search performance in such a collaborative desktop search task, and thereby the second hypothesis of the hypothesis set 1 was proved to be false. Furthermore, the lack of 2-way interaction between method and side implied that the third hypothesis of the hypothesis set 1 was not true.
6.2.2 Customization Behavior

To test the hypothesis set 2 of the third set of research question, we analyzed the data and found that 94% of the participants (45 out of 48 participants) generated similar icon area patterns, in which they used colors to represent the dominant non-quantitative dimensions (i.e., city) and sizes to represent the dominant quantitative dimensions (i.e., cost). In addition, we also got the following findings.

**Average number of clicks on the sizes tool.** The patterns for clicks on the colors tool, the sizes tool, and the “sort” tool were similar, so we chose to only report the average number of clicks on the sizes tool, which measured how often the subjects customized the sizes of the desktop icons. We divided the number of clicks on the sizes tool by the number of questions answered to calculate the average number of clicks on the sizes tool. ANOVA showed a significant main effect of side (F(1,44)=5.10, \( p < .05 \)), but not for communication method and the interaction between method and side (see Figure 6.4).

As shown in Figure 6.4, participants on side B made larger number of clicks on the sizes tool than those on side A did in the text condition (\( p < .05 \)), but not in the audio condition (\( p = .52 \)). However, the text and audio conditions had no significant difference in average number of clicks on the sizes tool both for side A (\( p = .70 \)) and for side B (\( p = .10 \)). The significant main effect of side was mainly caused by the large average number of clicks on the sizes tool for side B in the text condition. In the text condition, it was less convenient and efficient to communicate, and therefore the participants with more help-needed questions would customize more than those with less help-needed questions to better offload useful information to the external environment in order to better help themselves and minimize the disturbance to their partners. The results suggested how people would develop an external indexing of information to adapt to the communication method and the difficulty of the task.
6.2.3 Communication Behavior

To test the hypothesis set 3 of the third set of research questions, i.e. to reveal how different communication methods and different amount of help needed by different group members influence communication behavior, we analyzed the data and obtained some interesting findings.

In our experiment, even though the answer to a question was not exactly right, it still had some chance to be accepted as long as this travel suggestion indicated the right travel destination. We called these answers sub-optimal answers (with 70% or 30% acceptance rate). Thus, all answers submitted could be separated into four sets: absolutely correct (or optimal) answers, sub-optimal answers with a 70% acceptance rate, sub-optimal answers with a 30% acceptance rate, and absolutely wrong answers.

For the participants both on side A and side B the optimal answers to quite a lot of questions were exclusively filed on their partners’ desktops, so the participants only could obtain these optimal answers by asking their partners for help. Therefore, the ratio of the number of optimal answers submitted to the number of questions answered to some extent reflected how often the participants actively communicated with their partners. At the same time, for a given question, even when the optimal answer was not available on the participants’ own desktops, they still could always find several sub-optimal answers (with 70% or 30% acceptance rate) filed on their desktops, and hence the ratio of the number of sub-optimal answers submitted to the number of questions answered to some extent reflected how often the participants gave up asking their participants for help, and used the sub-optimal answers on their own desktops instead. However, the absolutely wrong answers would always be rejected by the customers, so the ratio of the number of absolutely wrong answers submitted to the number of questions answered reflected how often the participants made mistakes. We analyzed the above sets of answers for all of the 48 participants, and found that optimal answers accounted for much higher percentage of all of the answers (Mean=88.64%, S.D.=12.89%) than sub-optimal answers with 70% acceptance rate did (Mean=8.06%, S.D.=12.53%), p<.001; sub-optimal answers with 70% acceptance rate accounted for higher percentage of all of the answers than absolutely wrong answers did (Mean=2.62%, S.D.=2.28%), p<.005; absolutely wrong answers accounted for much higher
percentage of all of the answers than sub-optimal answers with 30% acceptance rate did (Mean=0.68%, S.D.=0.96%), p<.001. These results showed that participants usually did not make many mistakes, and most of the answers they submitted were either optimal answers or sub-optimal answers with 70% acceptance rate. Interestingly, participants submitted very few sub-optimal answers with 30% acceptance rate. It implied that participants were not satisfied with those sub-optimal answers which only had a 30% chance to be accepted. When they found this set of answers, they would ignore them and continue to search for optimal answers or sub-optimal answers with 70% acceptance rate. More interestingly, we found that to what extent participants would give up continuing to search for optimal answers and just use sub-optimal answers with 70% acceptance rate instead would greatly influenced by the communication method (i.e. text or audio) participants used and the amount of help-needed questions they had (i.e. side A or side B). We presented the details as below.

**Ratio about optimal answers.** ANOVA on the ratio of the number of optimal answers submitted to the number of questions answered yielded significant main effects of communication method ($F(1,44)=4.64$, $p<.05$) and side ($F(1,44)=11.90$, $p<.005$). However, the interaction between communication methods and side was not significant (see Figure 6.5).

As shown in Figure 6.5, audio condition led to this ratio higher than text condition did for side B (marginally significant, $p=.07$), but not for side A ($p=.33$). Participants on side A had higher ratios than those on side B both in the text ($p<.01$) and audio conditions (approached significance, $p=.07$). Participants on side A had much less questions to which the optimal answers were filed on their partners’ desktops than participants on side B did. Therefore, the results of “participants on side A had higher such ratios than those on side B in both of text and audio conditions” implied that in the two-participant group, the one with much less help-needed questions (i.e. side A) tended to ask his/her partner for more help than his/her partner did in order to obtain more optimal answers, while the one with much more help-needed questions (i.e. side B) tended to ask less questions in order to avoid interrupting his/her partner all the time. It was not surprising that audio condition led to higher such ratio than text condition did for side B since it was more convenient and efficient for participants to exchange information in the audio condition than in the text condition. We believed that the lack of significant difference between
text and audio conditions on side A was due to a ceiling effect. Specifically, if we could filter out the ceiling effect for side A in the audio condition, its ratio should be much higher, then we would find audio condition also could led to higher such ratio than text condition did for side A. We then continued to analyze the ratio about sub-optimal answers with 70% acceptance rate. The results are presented as below.

**Ratio about sub-optimal answers with 70% acceptance rate.** ANOVA on the ratio of the number of sub-optimal answers with 70% acceptance rate submitted to the number of questions answered yielded significant main effects of communication method \( (F(1,44)=4.63, p<.05) \) and side \( (F(1,44)=10.86, p<.005) \), but the 2-way interaction method with side was not significant (see Figure 6.6).

As show in Figure 6.6, audio condition led this ratio lower than text condition did for side B (marginally significant, \( p=.07 \)), but not for side A (\( p=.26 \)). Participants on side A had lower ratios than those on side B both in the text (\( p<.05 \)) and audio conditions (approached significance, \( p=.07 \)). The results of “participants on side A had lower such ratios than those on side B in both of text and audio conditions” suggested that the participants with much less help-needed questions (i.e. side A) tended to avoid submitting sub-optimal answers which only had a chance of 70% to be accepted by asking their partners more questions. However, the participants with much more help-needed questions (i.e. side B) tended to ask for less help, instead, they used more sub-optimal answers with 70% acceptance rate. It was not amazing that audio condition led to lower such ratio for side B since audio condition led to more convenient and efficient communication than text condition did and thus it encouraged participants to pursue more optimal answers and less sub-optimal answers with 70% acceptance rate. We did not find significant difference between text condition and audio condition for side A. Again, it was probably due to a ceiling effect, i.e. in the audio condition, the participants on side A could probably obtain much lower such ratio if the ceiling effect was deducted.

Moreover, we also analyzed the ratio of the number of sub-optimal answers with 30% acceptance rate submitted to the number of questions answered and the ratio of the number of
absolutely wrong answers submitted to the number of questions answered, but we did not find any significant result among the main effects or 2-way interaction effects.

To summarize, the above results showed that the audio condition made participants submit more optimal answers and less sub-optimal answers with 70% acceptance rate than the text condition did. It implied that the audio condition made participants on both sides communicate more with each other to share information. This was probably because the audio condition made communication between participants more convenient and efficient. When the optimal answers were not too hard to obtain in the audio condition, the participants prefer to submit these optimal answers rather than use sup-optimal answers with 70% acceptance rate instead. However, when the optimal answers were much harder to obtain in the text condition, the participants had to submit more sup-optimal answers with 70% acceptance rate instead of optimal ones. The above results also showed that the participants on side B who had much more help-needed questions submitted less optimal answers and more sub-optimal answers with 70% acceptance rate than the participants on side A did. It suggested that compared to the participants on side A, those on side B would ask their partners for less help. It was probably because that the participants on side B faced twice the number of questions which necessitated their partners’ help, and in this case, the participants on side B had to submit more sub-optimal answers with 70% acceptance rate instead of optimal answers to avoid bothering their partners all the time to ask for help.

To further prove the above results and study the participants’ communication behavior, we transcribed the audios which recorded the participants’ communication behavior in the experiment, and directly compared the difference communication behavior in different conditions. The results were presented as below.

**Average number of icon titles requested per help-needed question.** In this experiment, from question 30 to question 144, both participants on side A and side B would meet some questions to which the optimal answers were exclusively filed on their partners’ desktops. To obtain the optimal answers to these questions, the participants had to ask their partners for help. Since participants needed to input the icon titles as the answers, it was not surprising that participants always gave the known conditions to their partners and requested these 3-letter icon titles. We
therefore analyzed the number of icon titles requested in different conditions, hoping to observe interesting findings in participants’ communication behavior. Since we designed the experiment such that from question 37, there were 4 help-needed questions in every 9 questions for side A, and there were 8 help-needed questions in every 9 questions for side B, we decide to use every 9 questions as a block to further analyze the data. In each block, we used the number of icon titles requested divided by the number of help-needed questions in that block to calculate the number of icon titles requested per help-needed question in that block. It indicated how many times the participants asked their partners for help for answering one help-needed question in that block (see Figure 6.7).

Add up the number of icon titles requested per help-needed question in all blocks and divide the sum by the number of blocks, we finally obtained the average number of icon titles requested per help-needed question. It indicated how many times the participants ask their partners for help for answering one help-needed question. ANOVA on the average number of icon titles requested per help-needed question showed significant main effects of communication method (F(1,44)=93.94, p<.001) and side (F(1,44)=72.34, p<.001), however, the 2-way interaction between method and side was not significant (see Figure 6.8).

As shown in Figure 6.8, the parallel lines showed that there was no interaction effect between communication method and side. Audio condition led to larger total number of icon titles requested than text condition both on side A (p<.001) and on side B (p<.001), and participants on side A had larger total number of icon title requested than side B both in the text (p<.001) and audio conditions (p<.001).

Audio condition made communication more convenient and efficient than text condition did, and hence participants in the audio condition tended to ask their participants for more help. Participants on side B needed to answer much more help-needed questions than those on side A did, so they tended to ask their partners for less help for each of those help-needed questions to avoid bothering their partners all the time. The results here confirmed the finding in the four ratios regarding optimal, sub-optimal, and absolutely wrong answers. Based on the above results,
we can conclude that the first and second hypothesis of the hypothesis set 3 was true; however, the third hypothesis of the hypothesis set 3 was false.

6.3 CONCLUSION AND DISCUSSION

Results showed that audio condition in general led to better performance than text condition did, and the better performance was not caused by participants spending more time on finding and re-finding useful information, but rather, they achieved better performance of finding and re-finding information with less time. However, we did not find the less help-needed questions would lead to better performance. Moreover, we did not find audio and less help-needed questions would interact to lead to a better performance. Results also showed that for all of the participants, those on side B in the text condition customize the sizes tool the most and this was probably because compared to the participants in other conditions, those on side B in the text condition had stronger willingness to develop a good external indexing of information to adapt to the communication method and the difficulty of the task in order to better help themselves and minimize the disturbance to their partners.

Comparing the ratio of the number of optimal answers submitted to the number of questions answered, the ratio of the number of sub-optimal answers with 70% acceptance rate submitted to the number of questions answered, and he ratio of the number of absolutely wrong answers submitted to the number of questions answered, we found that in general audio condition led to that participants submitted more optimal answers and less sub-optimal answers with 70% acceptance rate than text condition did, suggesting audio condition led to more communication between the group members. This was probably because in the audio condition communication between group members was more convenient and efficient than text condition.

When in the audio condition, it was not too hard to obtain optimal answers by asking the partners for help, the participants prefer to get help from their partners to submit more optimal answers rather than use sup-optimal answers with 70% acceptance rate instead. However, when in the text condition, it was much harder to obtain the optimal answers by asking the partners for help, the participants had to submit more sup-optimal answers with 70% acceptance rate instead of
optimal ones. We also found that the participants on side B who had much more help-needed questions submitted less optimal answers and more sub-optimal answers with 70% acceptance rate than the participants on side A did, suggesting that the participants on side B tended to ask their partners for less help than those on side A did. We believed that it was because the participants on side B had much more help-needed questions, they tended to submit more sub-optimal answers with 70% acceptance rate instead of optimal answers to avoid keeping interrupting their partners to ask for help.

The average number of icon titles requested was an indicative measure for how often the participants asked their partners for help. We thereby directly comparing the average number of icon titles requested in different conditions and found that basically the participants in the audio condition tended to ask their participants for more help. It was probably because that audio condition made communication more convenient and efficient than text condition did. We also found that the participants on side B who needed to answer much more help-needed questions than those on side A did tended to ask their partners for less help for each of those help-needed questions. We believed that by doing so, the participants on side B hoped to avoid bothering their partners all the time. Those results confirmed the findings in the four ratios regarding optimal, sub-optimal, and absolutely wrong answers.

To summarize, audio condition and less help-needed questions led to more questions asked for each of those help-needed questions, and therefore they led to a higher ratio of the number of optimal answers submitted to the number of questions answered and a lower ratio of the number of sub-optimal answers with 70% acceptance rate submitted to the number of questions answered, and as a result, audio condition and less help-needed questions led to a better performance of finding and re-finding information in such a collaborative desktop search task.

Future research could provide the participants with more convenient and efficient communication methods and less help-needed questions to encourage the participants to communicate more with each other, more interesting findings were expected.
CHAPTER 7
CONCLUSIONS

In this thesis, I proposed three sets of research questions and hypotheses and reported the results of three experiments. The first experiment was about an individual desktop search task with an icon size edit tool; the second experiment was still an individual desktop search task, but we provided the participants with more complicated tools – a set of customization tools which involved icon color edit tool, icon size edit tool and sort tool; the third experiment still used the customization tools which had already been used in the second experiment, but this time, a collaborative desktop search task was performed.

The major goal of the first experiment was to test how the cost and benefit of customization would influence search performance, willingness to customize, and information indexing strategies. We did not find significant difference in search performance to verify the first hypothesis probably due to the low number of questions designed in the experiment. However, we did find significant differences in participants’ willingness to customize and their strategies of information indexing which verified the second and third hypotheses. Consistent with the theory of soft constraints, we found that participants in random conditions tended to make more icon edits when the cost of editing was low. When we separate the 36 desktop icons into useful and less useful icons, we found different usage patterns. For the useful icons, participants in low cost and high benefit condition utilized the feature of size editing the most by making the useful icons larger compared to participants in other conditions. Besides, we found a couple of interesting things in the icon access transition tables that informed the use of search strategies. Specifically, forward sequential accesses took a higher percentage of total accesses in the organized conditions than in the random conditions and in the high cost conditions than in the low cost conditions; backward sequential accesses took a higher percentage of total accesses in the random conditions than in the organized conditions.
The second experiment was designed to test how customization tools, dominance and dispersion structures of information needs influence participants’ search performance, customization behavior, search behavior and information indexing strategies. We found that in general better performance was achieved when the customization tools were available. However, performance was much better in the dominance conditions than other conditions: tools did not seem to improve performance in the dispersion conditions. This pattern of results suggested that the match between the tools and information structures was critical for performance. When tools were not available, both structures helped participants to achieve better performance, suggesting that mental indexing was effective in helping participants to re-find information in spite of the limitation of memory capacity. We found that all participants were using the customization tools extensively as they were adapting to the information structures during the early trials, but they did not customize much for each of the rest questions especially in the dominance conditions, as the structures remain stable throughout the experiment. Interestingly, participants in the without-dominance with-dispersion condition showed the largest number of use of the customization tools in the experiment, presumably because participants perceived structures in the questions (that they tended to point to the small subset of icons), but because the indexing dimensions were changing throughout the experiment, they had to keep adapting to the “structures” by indexing different dimensions throughout the experiment. This showed how a mismatch of tools and information structures could lead to poor adaptation to the environment. In general, the customization tools were useful for indexing information, but their effects were stronger when participants could offload information indexing to represent the dominance and dispersion structures using customization cues provided by the tools. In addition, we found from the posttest of information indexing that even in the with-tools condition, subjects had developed mental indexing of information. In particular, in the dominance condition, the participants were able to develop a mix of external and internal indexing strategies to help them find the correct icons. Similarly, when the dispersion structures existed, participants could encode, at least partially, which cons were most frequently requested. Results suggested that people were sensitive to both information structures when they adaptively learn to select icons for internal indexing.

To perform the third experiment, we wanted to investigate how different communication methods and different amount of help needed by different group members would influence
search performance and communication behavior. We found that audio condition in general led to better performance by making participants communicate with each other more efficiently and share information with less time. However, we did not find less help-needed questions led to better performance. We also found that for the participants with less help-needed questions, audio condition led to a higher ratio of the number of optimal answers submitted to the number of questions answered and a lower ratio of the number of sub-optimal answers with 70% acceptance rate submitted to the number of questions answered. It was probably because audio condition made the communication more convenient and efficient, so participants prefer to ask their participants for more help in order to get more optimal answers rather than use sub-optimal answers instead. We did not have such finding for the participants with much more help-needed questions and we believed it was because of a ceiling effect. Interestingly, we found participants with much more help-needed questions got a lower ratio of the number of optimal answers submitted to the number of questions answered and a higher ratio of the number of sub-optimal answers with 70% acceptance rate submitted to the number of questions answered, suggesting they use more sub-optimal answers filed on their own desktop instead of optimal answers which only could be obtained by asking their partners for help. By doing so, these participants probably hoped to reduce disturbing their partners. Besides, we found from the analysis of the average number of icon titles requested that basically audio condition made participants ask their partners for more help for each of the help-needed questions. It was probably because comparing to text condition, audio condition was more convenient and efficient and it thereby encouraged more communication. We also found that the participants with less help-needed questions tended to ask their partners for more help for each of the help-needed questions to obtain more optimal answers, while participants with much more help-needed questions tended to ask for less help for each of the help-needed questions probably to avoid bothering their partners too often.

In sum, we found in the individual desktop search task with an icon size edit tool that low cost and high benefit of customization would increase participants’ willingness to customize and encourage participants to attempt more diverse information indexing strategies. We found in the individual desktop search task with a set of customization tools that customization tools were in general useful for indexing information, especially when dominance and dispersion structures of information needs exist. A match of tools and information structures could lead to good
adaptation to the environment. We found that participants utilized tools to offload information to the external environment and at the same time, they also used the internal (mental) indexing to help them find and re-find information. Therefore, the participants actually tended to develop a mix of internal and external indexing of information to help them search for information. Finally, we found in the collaborative desktop search task with a set of customization tools that more convenient and efficient communication method and less help-needed questions respectively led to more questions asked for each of those help-needed questions, and therefore they led to a higher ratio of the number of optimal answers submitted to the number of questions answered and a lower ratio of the number of sub-optimal answers with 70% acceptance rate submitted to the number of questions answered. As a result, audio condition led to a better performance of finding and re-finding information (less help-needed questions did not lead to such finding probably due to a ceiling effect).
CHAPTER 8
GENERAL DISCUSSION

“People are busy and customizing takes time, so they only customize when they deem it worth the trouble.” (Mackay, 1991) Although anecdotal evidence suggest that people do not customize their desktop for re-finding of information, but we believe there are other factors that may influence people’s willingness to customize, and one such important factor is the cost-benefit tradeoff involved in customization. We therefore designed the first experiment to directly manipulate the cost and benefit of customization, and finally we did find the variation of cost and benefit of customization influence people’s customization behavior and their information index strategies. Specifically, low cost and high benefits lead to more customization behavior and more diverse strategies of information. However, we did not find significant between-group difference in search performance which was informed by total scores and average time per question answered. According to previous research (Mackay, 1991) that “People are most likely to customize when they discover that they are doing something repeatedly”, we believed the lack of significant results was due to the low number of questions – there were only 13 questions – involved in the experiment. Participants probably considered that the benefit of customization could not outweigh its cost, so they did not fully employ the customization options, and finally it led to the insignificant results in search performance. Thus, we planned to design more questions in the next experiment to encourage more customization.

In addition to greatly increasing the number of questions – there were 144 questions involved – in the second experiment, we also designed dominance and dispersion structures of information needs. Through extensive field interviews, we found the dominance and dispersion structures of information needs are very common in real life, and we thereby were very interested in how these structures interact with the use of customization tools to influence finding and re-finding of information. From the analysis of the experimental data we found that the customization tools were in general useful for indexing information, but their effects were stronger when participants could make use of the tools to adapt to the dominance and dispersion structures of information...
needs. We found that when tools were available, better performance was achieved, suggesting that tools could help participants offload information indexing to external environment, however, when tools were not available, both dominance and dispersion structures helped participants achieve better performance, suggesting that internal (mental) indexing was effective in helping participants to find and re-find information in spite of the limitation of memory capacity. Interestingly, we found that instead of generating internal indexing and external indexing separately, participants actually developed a mix of internal and external information indexing strategies to help them find the useful information. The results let us better understand how people use tools to adapt to information environment in order to find and re-find useful information efficiently in a desktop environment, and it would thereby shed light on how we could design novel tools to help people more effectively index, find and re-find information.

Different from the first and second experiments, the third one is about a collaborative desktop search task. Customization tools were still available, and dominance and dispersion structures of information needs were still involved in the experiment, but this time we shift most of our attention to participants’ communication behavior during the experiment. Finally, we found that less convenient and efficient communication method (e.g. text condition) made participants ask their partners for less help and it thereby led to a lower proportion of optimal answers and higher proportion of sub-optimal answers. As a result, it decreased people’s search performance. At the same time, we also found that more help-needed questions made participants tend to ask for less help for each of the help-needed questions to minimize the disturbance to their partners, and as a result, it led to a worse search performance. When participants asked their partners for help, they usually needed to say something like “thank you” and “good job” to express their appreciation and encourage their partners to provide help later. Indeed, if we treated these words as the special “cost” of communication, the results implied that higher cost of communication (either because of poor communication method such as text condition or because of more appreciation and encouragement needed to express) in a task would make participants use more sub-optimal information instead of optimal information. This probably revealed people’s communication strategies in such a collaborative desktop search task.
There were some limitations in our studies. In the first experiment, there were only 13 questions in the task and all of the 40 participants in the four experimental conditions completed all of the 13 questions. Therefore, we could compare the total scores among the four conditions purely depended on the accuracy rate, not depended on the number of questions answered. Similarly, we could compare the average time for answering one question based on the completion of the whole task. We thereby believed that the results were more meaningful than those which were based on partially completed task. However, the number of questions – 13 – was too small to encourage participants to customize, and it led to the lack of between-group differences in customization behavior and search performance. Thus, we decided to design more questions in the following experiments.

In the second and third experiment, there were 36 travel plans, each of which was represented as a desktop icon. We designed 144 questions in the search task – averagely 4 questions repeatedly asked about one desktop icon. By doing so, we hoped to make participants realize the huge benefit of customization and encourage them to actively customize the visual features (e.g. color and size) of these desktop icons. From the analysis of the experimental data we found that the large number of questions did increase participants’ customization behavior and thus improve their search performance. However, we also found that many participants could not complete the whole task, especially for those participants in the condition that the customization tools were not available and the dominance and dispersion structures of information needs did not exist. It showed that 144 questions were probably too many to finish in the given time. In this case, we had to compare the total scores among the 8 conditions based on the number of questions answered rather than the accuracy rate and compare the average time for answering one question based on the partially done task. Comparing the results in the first experiments and those in the second and third experiments, we thought a more appropriate number of questions should be
used in the next experiment – not too large and not too small – to not only encourage participants to customize but also enable all of them or at least most of them to finish all of the questions.

Besides, in the third experiment, the program was designed such that if a participant finished his/her whole task (i.e. 144 questions), then his/her desktop interface would be locked. Therefore, the participants who finished the experiments first had to only use their memories – i.e. only use their internal indexing of information in place of the mix of internal and external indexing of information – to help their partners and they could not directly click on the icons to confirm their memory. Obviously, this design depreciated the value of the help from the participants who finished the experiments first and thereby decreased the search performance (measured by total score and average time per question answered) of their partners who finished the experiments second. In later experiments, we should revise the program to guarantee even though the participants have finished the whole tasks, their desktop interface will still be in the state of activation. By making this revision, more valuable help from the participants who finish the experiments first and better search performance of their partners will be expected.
Table 4.1. Transition table of the Lo-Random condition. The highlighted cells indicated forward and backward sequential accesses, which had the highest frequencies compared to others.
Table 5.1. An Example of Transition table (transition table of participant 1 in condition 1).
A housewife suffered attacks of dizziness that left her quite incapacitated. She would be overcome with feelings of extreme dizziness, accompanied by slight nausea, 4 or 5 nights a week. Inexplicably, the attacks almost always occurred at about 4 PM. She usually had to lie down on the couch and often did not feel better until 7 or 8 PM. After recovering, she generally spent the rest of the evening watching TV; more often than not, she would fall asleep in the living room, not going to bed in the bedroom until 2 or 3 AM. *The symptoms are apparently not intentionally produced*.

The patient had been pronounced physically fit by her internist, a neurologist, and an ear-nose-throat specialist on more than one occasion. Hypoglycemia had been rule out by glucose tolerance tests. *The list of physical symptoms cannot be accounted for by a general medical condition*.

When asked about her marriage, the patient describes her husband as a tyrant, frequently demanding and verbally abusive. *Apparently she has a stressful relationship with her husband*. She admitted that she dreaded his arrival home from work each day, knowing that he would comment that the house was a mess and the dinner, if prepared, not to his liking. Recently, since the onset of her attacks, he and the 4 kids would go out to eat. After that, he would watch TV and their conversation would be minimal.

**Figure 4.1. Example of health information search task**
Figure 4.2. Interface of search task, answer box, icon size edit tool, 36 file icons, and reference information.
Figure 4.3. Mean number of size edits

Figure 4.4. Mean final icon size
Figure 4.5. Forward Sequential Accesses (left) and Backward Sequential Accesses (right)
Figure 5.1. Desktop Interface
Figure 5.2. Total score.

Figure 5.3. Average time.
Figure 5.4. Average number of clicks on the sizes tool.

Figure 5.5. Average number of clicks on the sizes tool in the first 10 questions.

Figure 5.6. Average number of clicks on the sizes tool in the rest of the questions.
Figure 5.7. Average number of clicks on the icons.

Figure 5.8. Percentages of clicks on the most useful icons when the answers could be found in these icons.
Figure 5.9. Percentages of wrong clicks on the most useful icons when the answers could not be found in these icons.

Figure 5.10. Percentages that sequential accesses account for total transition accesses.
Figure 5.11. One-click test score.

Figure 5.12. Score for the two questions asked about the two dominant dimensions.
Figure 6.1. An example of the desktop interface in the third experiment.
Figure 6.2 Total score

Figure 6.3 Average time spent for answering one question
Figure 6.4 Average number of clicks on the sizes tool.
Figure 6.5 Ratio of the number of optimal answers submitted to the number of questions answered.

Figure 6.6 Ratio of the number of sub-optimal answers with 70% acceptance rate submitted to the number of questions answered.
Figure 6.7 Number of icon titles requested per help-needed question across blocks

Figure 6.8 average number of icon titles requested per help-needed question
REFERENCES

Alvarado, C., Teevan, J., Ackerman, M. S. and Karger, D. Surviving the Information Explosion: How People Find Their Electronic Information. *AI Memo 2003-006* (2003), Massachusetts Institute of Technology Artificial Intelligence Laboratory.


