CHAPTER IV. RESULTS

This chapter details the outcomes of the research effort. The study was divided into a pilot study and a formal online trip planning exercise. In the pilot study, the research methods were inspected and the instrumentation was refined. In the formal exercise, the structure of travel information search on the Internet was explored using three phases of analysis. In Phase I, a descriptive analysis of the trip planning exercise including a description of the whole information space and the subject's profiles were detailed. Concept mapping was used to explore the structure of travel information search in Phase I. Using semantic network analysis, Phase II derived the travelers' mental models and the semantic model of the travel information space and compared them in order to explore the overall congruence between the two models. In Phase III, correlation analysis was used to compare each individual information searcher's mental models with the semantic model of the travel information space; their congruence was further correlated with satisfaction of trip planning on the Internet and a travel information searcher's individual characteristics.

4.1 Results of Pilot Study

The pilot study was conducted during May, 2002. Five students in the Department of Leisure Studies at University of Illinois at Urbana-Champaign (UIUC) participated in the study. The subjects were asked to plan a trip to San Diego, California using a computer located in the National Lab for Tourism and eCommerce (NLTeC) at UIUC. The subjects were first asked to complete a questionnaire regarding their travel experience, computer and Internet use experience and their experience of using the

Internet as a travel information source. After competing this task, they were interviewed regarding their general travel preferences and their background knowledge regarding San Diego. They were then instructed to search information on the Internet in order to plan a trip to San Diego for a weekend trip, and they were instructed to fill out a travel plan form which includes every hour of their trip to San Diego. In addition, they were asked to complete another questionnaire regarding their satisfaction of trip planning. Last, the participants were interviewed regarding the problems they encountered during the information search.

The results of the pilot study indicated that there were some problems in the instrumentation and research procedures. The following are some specific insights regarding the research procedure.

1. Trip planning exercises were intensive mental activities and usually lasted from 30 to 50 minutes. After about 50 minutes, the subject usually felt mentally exhausted and the trip planning would terminate. Therefore, the scope of trip planning task needs to be limited in order to obtain meaningful results. The final trip planning on the Internet was limited to one hour and the trip planning task was restricted to activities in a two-day weekend vacation including one night of accommodation;
2. Trip planning on the Internet was an enjoyable experience, as reflected from the feedback from the subjects. Therefore, high level of satisfaction measure and low levels of variance were expected from the post-exercise survey. In the post-

exercise interview, the subjects were asked to explain their levels of satisfaction in order to capture the different underlying reasons for their satisfaction levels;

3. The subjects would plan the trips in different levels of details. Therefore, the details of the trip planning should be carefully designed and kept identical for each subject. The description of the planning task is explained in Appendix C;
4. Some items in both pre-exercise and post-exercise survey appeared confusing to the subjects. They were changed and/or deleted from the original version (See Appendices A, B, and D for revised version of the respective surveys);
5. In order to more closely simulate a real world scenario, each subject was informed that one participant among them would be randomly selected to win two round-trip airline tickets to San Diego, California. The subjects would be asked to plan a trip to San Diego assuming that they already won the two airline tickets;
6. A digital camcorder was used to record the subject's large scale information behavior (for example, writing down attraction information on a piece of paper or printing out web pages) and facial expressions, which will be complimentary to the data gathered through other sources;
7. An interview regarding the information search was conducted after the exercise in order to clarify the uncertainties on the subjects' information processing in case the subject' verbalization is not sufficiently clear and understandable;
8. The trip planning form used in the pilot study, in which the subjects were asked to complete their hour-to-hour activities in two days, was abandoned in the formal exercise. The subjects have different trip planning styles and some

subjects are not willing to plan the trip in much detail. Instead, the subjects were asked to write a short paragraph regarding their travel plans (where they are going to stay for one night, where they are going to visit for two days and a short explanation for their choices). Their description of the travel plan can be easily analyzed using semantic network analysis and compared with other types of semantic models;

9. In the post-exercise interview, the subjects would be interviewed regarding their uses of other types of information sources besides searching information on the Internet and their use of the Internet in the real world situation. The results could provide broader and holistic view on the use of the Internet as a travel information source in the entire travel information environment surrounding them.

4.2 General Information on the Trip planning Exercise

The formal exercise was conducted throughout August, 2002 in the National Laboratory for Tourism and eCommerce (NLTeC), UIUC. Fifteen subjects were recruited from various newsgroups in the university including 10 undergraduate students, 1 graduate student, and 4 researchers. Their ages ranged from 19 to 45 years old with an average age of 25; 6 subjects were men and 9 were women. Four of the subjects have been to San Diego at least once. The 15 subjects took an average time of 36 minutes to plan the trip to San Diego, with a minimum time of 20 minutes and a maximum time of 55 minutes. The subjects visited from 7 to 26 web sites with an average of 15 web sites (see Appendix E for a complete list of all the web sites the subjects visited). They

visited from 60 to 312 web pages, averaging 124 web pages per subject. About half of the subjects (7 of 15) used a printer to print out and organize travel related information. Twelve subjects used a piece of paper to copy and organize information and 2 subjects used WordPad or Microsoft Word to copy and organize information.

On average, each subject took 4.2 trips in the last year and they took 2.6 trips on an average year. They considered themselves relatively experienced travelers (a mean score 5.3 on a 7-point Likert Scale when asked "Do you consider yourself an experienced traveler?"). Most of the subjects had been to California. Twelve subjects among the fifteen had previously used the Internet to check out destination, hotel and airline information. Fourteen of the fifteen subjects have done at least one of the following three types of activities: book airline tickets, reserve hotel rooms or rent a car online. The subjects have used computers from 4 to 28 years with an average computer use history of 14.3 years. They used the World Wide Web from 4 to 17 years with an average of 7.5 years and they used emails from 4 o 17 years with an average of 7.7 years (see Appendix F for a complete list of the results of the questionnaire on the subject's individual characteristics).

4.3 Phase I. Examining the Process of Travel Information Search on the Internet

The goal of Phase I analysis was to understand the structure of travel information search on the Internet. In total four sets of data were captured during the trip planning exercise: (1) clickstream data was obtained in order to capture the subjects' navigation behavior (Table 4-1); (2) movie files of computer screen activities along with the subject's verbalization (Figure 4-1); (3) artifacts from the trip planning experiments including

printed web pages, sheets of paper used for copying and organizing information, and Word files or WordPad files used for copying and organizing information (Figure 4-2); and, (4) movie files of large information behavior and facial expression recorded through the digital camcorder (Figure 4-3). Table 4-1 shows an example of the raw clickstream data which contains time stamp, action performed by the computer user, computer programs used, next visited link, and time spent on that action. Figure 4-1 shows a screen shot of the movie file recorded from online screen capturing software. This software captured everything on the computer screen during one computer using session, including web pages visited, the movement of the mouse, and subject's verbalization captured through a microphone. Figure 4-2 shows one type of artifacts from the trip planning exercise: a Microsoft Word document used by the subject to copy and organize information obtained while searching the Internet and contains addresses and contact information regarding attractions and accommodations. Figure 4-3 is one screen shot of the movie file captured using digital camcorder. In the movie facial expressions and large information organization behavior were captured. Furthermore, subjects' verbalization through the online screen capturing software was transcribed and matched with clickstream data from Internet monitoring software. The artifacts and the movies from the digital camcorder were also matched with two other data sets in order to gain a better understanding of the information search and organization process. A detailed information search protocol was generated including information behavior (Click Link, Type In, Print, or Write on paper), time spent on that behavior, starting time, the subject's verbalization and the coder's interpretation (see

08/06/2002	20:15:40	00:01	TITLE	Program Manager	c:\winnt\explorer.exe	-
08/06/2002	20:15:41	00:04	TITLE	http://www.microsoft.com/isapi/redir.dll?prd=ie&pver=5.5&ar=msnhome - Microsoft Internet Explorer	c:\program files\internet explorer\iexplore.exe	-
08/06/2002	20:15:42	-	Link	http://www.microsoft.com/isapi/redir.dll?prd=ie&pver=5.5&ar=msnhome - Microsoft Internet Explorer	c:\program files\internet explorer\iexplore.exe	http://www.msn.com/
08/06/2002	20:15:45	-	Keystrokes	http://www.microsoft.com/isapi/redir.dll?prd=ie&pver=5.5&ar=msnhome - Microsoft Internet Explorer	c:\program files\internet explorer\iexplore.exe	w
08/06/2002	20:15:45	02:04	SUBTITLE	Welcome to MSN.com - Microsoft Internet Explorer	c:\program files\internet explorer\iexplore.exe	-
08/06/2002	20:15:45	-	Keystrokes	Welcome to MSN.com - Microsoft Internet Explorer	c:\program files\internet explorer\iexplore.exe	ww.sandiego.com
08/06/2002	20:15:54	-	Link	Welcome to MSN.com - Microsoft Internet Explorer	c:\program files\internet explorer\iexplore.exe	http://www.sandiego.com/
08/06/2002	20:16:14	-	Link	Welcome to MSN.com - Microsoft Internet Explorer	c:\program files\internet explorer\iexplore.exe	http://www.radisson.com/sandiegoca
08/06/2002	20:16:25	-	Link	Welcome to MSN.com - Microsoft Internet Explorer	c:\program files\internet explorer\iexplore.exe	http://www.radisson.com/hoteldirectory/hotellocationshome.jsp?hotelFeatures2ID=624&selection=location&hotelCode=SAN%20DIEG
08/06/2002	20:16:56	-	Link	Welcome to MSN.com - Microsoft Internet Explorer	c:\program files\internet explorer\iexplore.exe	http://www.seaworld.com/

Table 4-1. A Part of Clickstream data from Internet Monitoring Software for Subject #13

Figure 4-1. File Recorded from Online Capturing Software

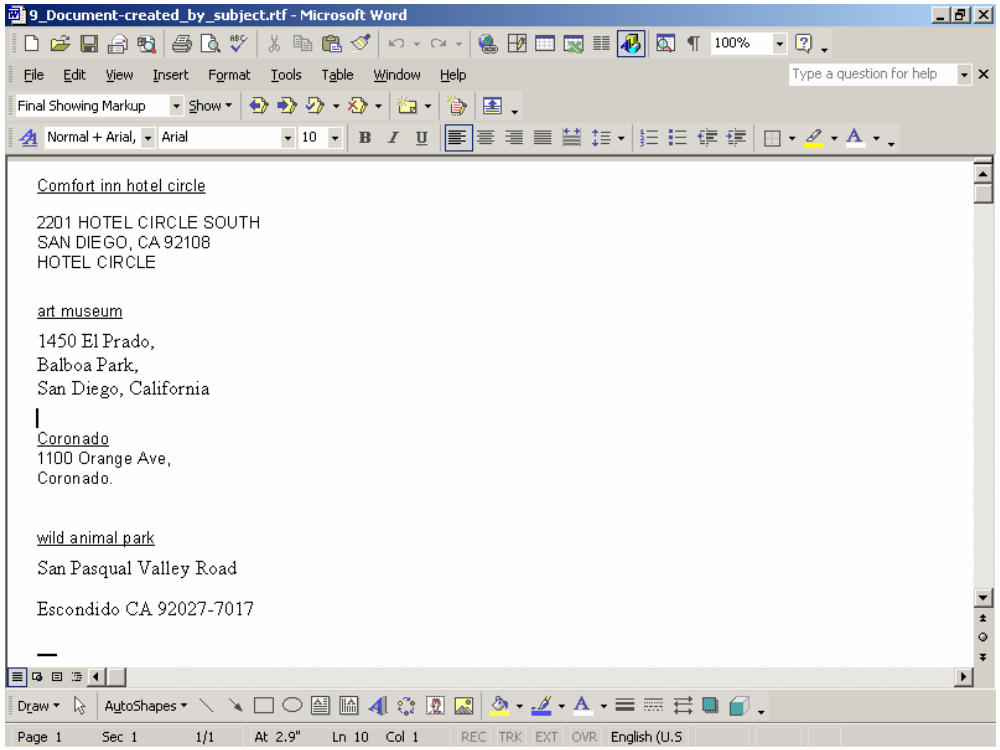

Figure 4-2. Software Used for Organizing Information

Figure 4-3. File Recorded from Digital Camcorder

Table 4-2 for a part of the protocol from one of the subjects). The following analyses were based on this general information search protocol data.

4.3.2 Mapping the Travel Information search on the Internet

Graphs have been used as a representation of web search behavior (Hodkinson, Kiel, & McColl-Kennedy, 2000). The nodes can represent the prominent semantic concepts in the subject's mind and therefore the most relevant semantic information the searchers are looking for and the semantic focus of the information searcher at a certain time point (Chi, Pirolli, Chen & Pitkow, 2001). They also represent a web page an information searcher navigated. Figure 4-4 shows a directed graph of the travel information search of Subject #13. A circle is used to represent the clicked anchor or a typed search term, indicating the dominant semantic concepts in the travel information

searcher's mind. At the same time, it also represents a web page resulting from the clicked link or the result from searched terms (See Table 4-2 for the clicked link anchors in the clickstreams in **Bold** texts). A square box represents the end results of one search. The subject may also write down information if the information is relevant, abandon the web page because it's irrelevant, or encounter a broken link. The subject may revisit and further explore a sub-graph because of relevant information cues. Undirected graphs can be used to represent the navigational path of the subjects and emphasizes the semantic structure of web sites, ignoring the repetitive behavior of the subjects (See Appendix H for the resulting graphs for each of 15 subjects).

As can be seen, sub-decisions during navigation process can be identified. The subject may find interested information and decide to check out one of the alternatives; s/he may get lost and decide to go back to the homepage to start the search again; or the subject encounters a broken link and get frustrated. These decisions demarcate different "episodes" in the navigation process which can be verified from the subject's verbalization. For example, Subject #13 said: "… Now hotel has been taken care of… next I'm going to see what to do over there…" A summary hierarchical graph can be generated from the clickstreams performed by the subjects in order to explore the detailed structure of the trip planning process (See Figure 4-5). The bottom level boxes contain original clickstream data, which is represented by the bold text in Table 4-2. A sequence of clicks comprises an episode, which involves evaluation of one alternative in a sub-decision. In Figure 4-5, the first two episodes focus on Radisson

Time Spent	Time	Behavior	Verbalization	Interpretation
0:00:03	20:15:42	Click Link	Go back to use Internet Explorer,	Start IE.
0:00:09	20:15:45	Type In	I'll go to google… SanDiego.com to see what I find here. See if there's anything interesting from the web site I can find out. Since I haven't been to the city.	Type in guess address: **sandiego.com**
0:00:20	20:15:54	Click Link	All right, so we're here. Just looking around to see what's on here. Hotels, all right. Radisson Hotel San Diego. That catches my eye.	Click on **Radisson Hotel San Diego**.
0:00:11	20:16:14	Click Link	Stayed at Radisson before. I like them. Looks pretty nice. 89 to 90. I can deal with that for a day.	Scan information and click on **location**.
0:00:31	20:16:25	Click Link	All right, location. Says, the heart of San Diego. Minutes from the major attractions. That's a good thing. Oh, Sea World. Oh, I probably make time to do that. I want to go to Sea World. I like fish. Working at marine labs, so. That'll be something I'm very interested in doing. All right. It's minutes, it says, to the attractions.	Click on **Sea World** link.
0:00:06	20:16:56	Click Link	I probably just take a taxi and not worry about renting a car for a day. It's a little bit too much.	Click on **California** on Sea World page.
0:00:10	20:17:02	Click Link	Too much work. All right. I'm at Sea World. So let's see. Park info.	Click on Park Info.
0:00:12	20:17:12	Click Link	Let's see. Wow, they have varieties, too. Hours of operation. Wow. It's expensive.	Check out detailed information and click on **Operation Schedule.**
0:00:03	20:17:24	Click Link	…	Choose month.
0:00:22	20:17:27	Click Link	Choose the month of October. I'm gonna be there 19th and 20th. So it's open until 7:30PM, open at 10AM. Sounds like a good way to spend my Saturday once I get there. Check in my hotel and go to the Sea World. Really great idea. That's what I wanna do Saturday.	Look at the **schedule** and decide to go there on Saturday.
0:00:04	20:17:49	Type In	Let me check hard rock café dot com.	Type in address of **Hard Rock café**.

Table 4-2. Sample Information Search Protocol

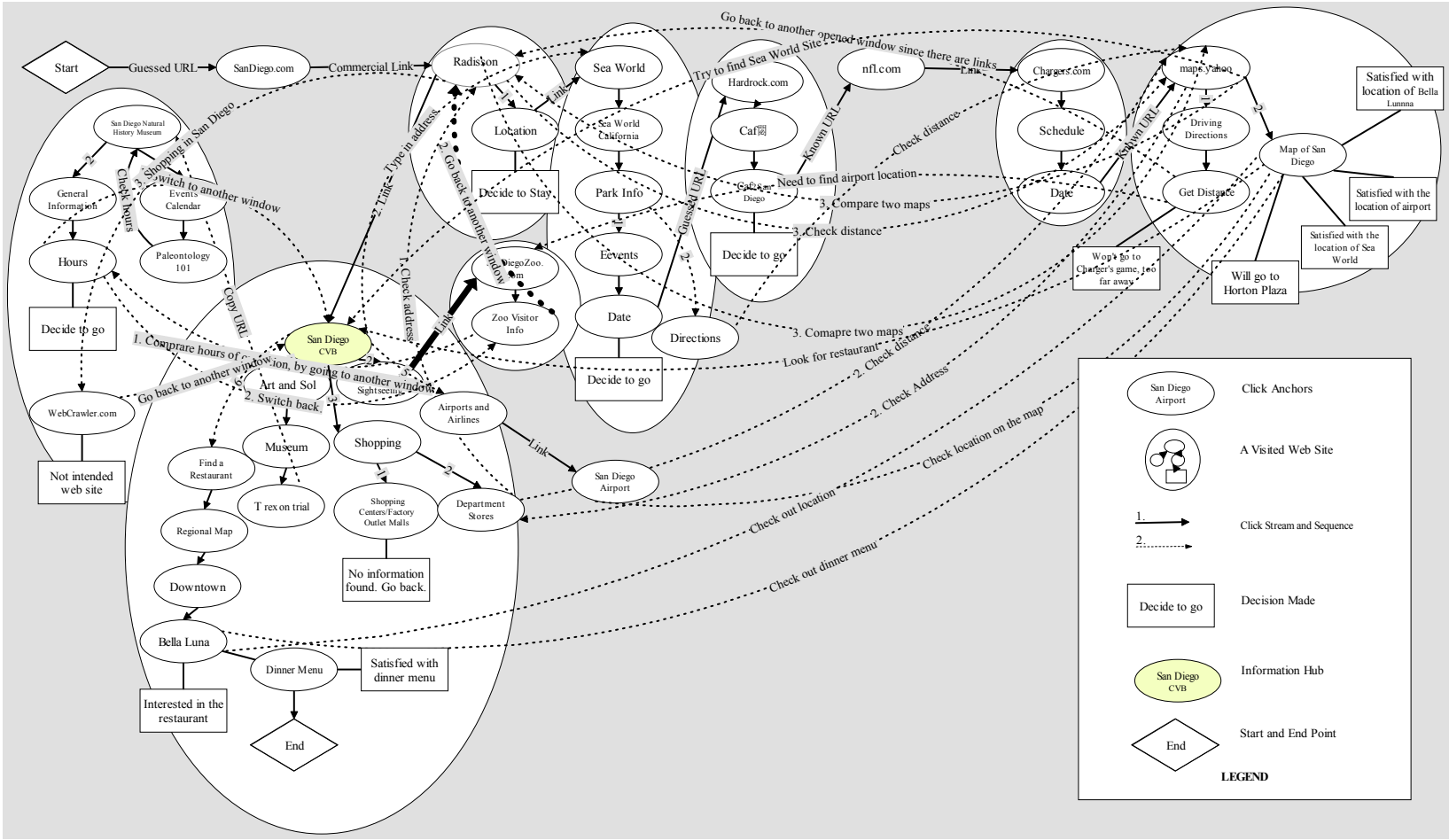

Figure 4-4. Clickstream Semantic Map of Subject #13

Hotel and Sea World. These episodes can further be aggregated into different chapters, representing one decision/sub-problem in the subject's travel plan, e.g. hotels, attractions, or transportations. For Subject #13, we can see s/he moved from hotel information, to attraction information and shopping information, and finally dinning information. However, a special "episode", which is indicated by the box with a thick border in Figure 4-5, can not be fit into any single type of travel plan. It is an information evaluation process in which the travel information searcher checked the address of the hotel and the attractions and located their relevant locations on the map. The travel information searcher re-evaluated the alternatives on accommodations, attractions, and dining places along the geographical frame to make her/his final decisions. The result indicates that the subjects' trip planning behavior follows a linear fashion in general but with some additional re-evaluation episodes which are the exceptions of linear search.

4.3.3 Assessing Themes of Travel Information Search on the Internet

Several common themes/findings emerged in the generation of travel information search protocols and graphs. These findings reflected the characteristics and the nature of travel information search on the Internet, not only include the structure and characteristics of online travel information search itself, but also the role of the Internet in the entire travel information environment. The following describes these findings:

1. The Internet (more specifically the web browser) was never used alone in the trip planning exercise. Most of subjects (12 of 15) used a sheet

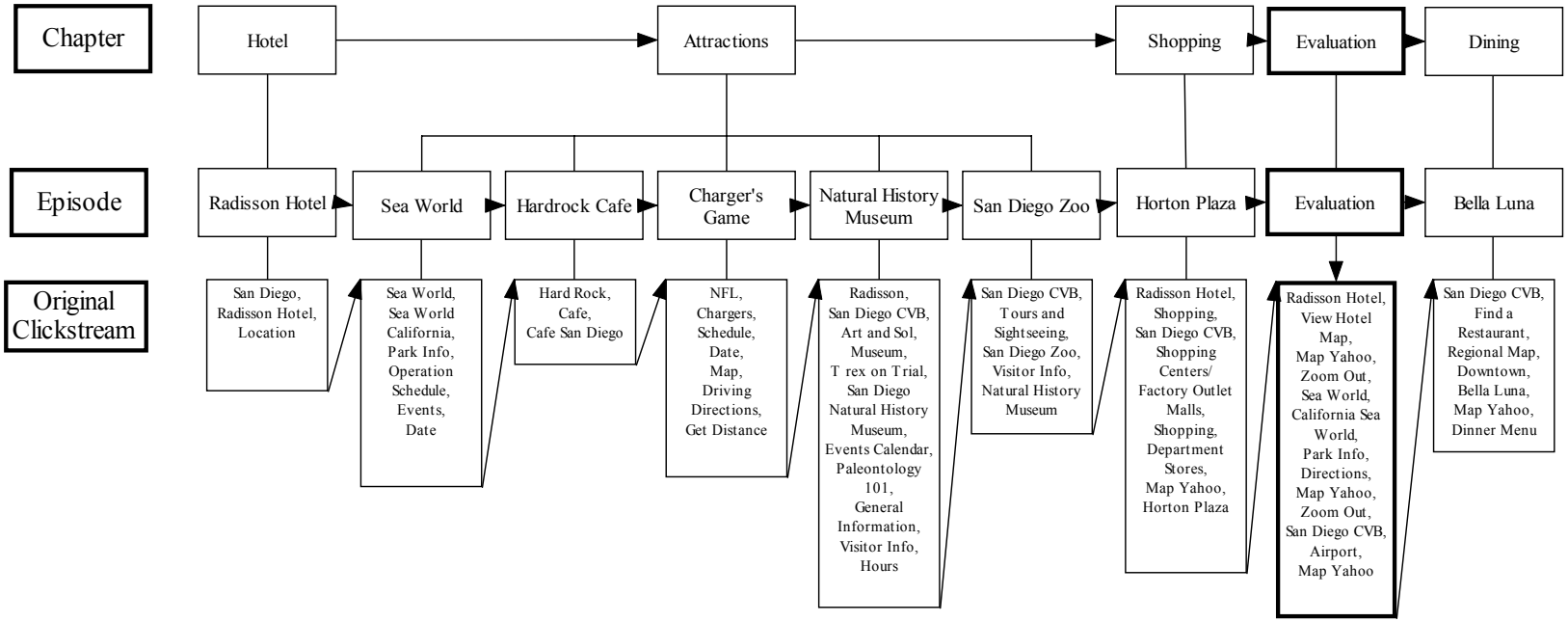

Figure 4-5. A Hierarchy of Travel Information search for Subject #13

of paper and a pen to copy down information from web pages and organize information. Two subjects used Microsoft Word or WordPad to copy information from web pages and organize information. About half of the subjects (7 of 15) used a printer to obtain hardcopies of the information from web pages for trip planning. These findings indicate the complexity of trip planning and the information searchers' limited information processing capability.

2. Information processing during travel information search on the Internet is more diverse than anticipated. Navigation and reading processes can not capture the whole spectrum of information activities the subjects were performing online. Instead, the subjects' behavior can be categorized into search (type in search terms in search box), navigation (click link), and information organization (print, and write write down information on a sheet of paper). Navigation can be further divided into goal-oriented navigation and browsing (Marchionini & Shneiderman, 1988). Contrary to previous expectations, reading pages are not necessarily the pages where the subjects stay for a longer period of time. Some browsing pages are extremely complex and contain large amounts of options, thereby requiring the subjects to spend a long time reading information and make a decision on which link to click in order to find more detailed information. Furthermore, it was almost impossible to distinguish a navigational page from a reading page since most web

pages contain a menu bar which represents the navigation structure of the web site.

3. More than half of subjects (8 in 15) started with hotel choices (the first chapter), and then switched to transportation or different activities and attractions. It indicated that in a higher level there are maybe more commonalities across all the subjects.
4. It was found that geographic information and time information were important constraints affecting choice of hotels, means of transportation, and activities and they are not provided across different tourism web sites. The subjects always had trouble locating various hotels and attractions whereby they were forced to switch between the hotel/attraction web sites and online map web sites (e.g. MapQuest) to find geographical locations.
5. All the subjects were looking for "information portals" or "hub" pages which are "authoritative" web pages with many links to famous attractions or accommodations. These "information hubs" are also those web pages the subjects returned to as they started a new episode of problem-solving. The San Diego Convention and Visitors Bureau web site (http://www.sandiego.org/), the official web site for the municipal government of San Diego (http://www.sannet.gov/), and a commercial information portal for San Diego (http://www.sandiego.cc/) were chosen (typed in the web address bar or clicked in the navigation process) by most of the subjects as the starting points for exploration of San Diego.

This finding is consistent with information foraging theory (Pirolli & Card, 1999) in that information searchers tend to search information in clusters in order to minimize inter-cluster information seeking cost.

The results of Phase I analysis showed that the Internet was never used alone but with other types of information organization tools and decision aids. Phase I analysis also confirmed that the travel information search can be broken up into different episodes in which each episode targets at a specific problem. Each episode represents an evaluation of an alternative of attraction, accommodation or mean of transportation. Several episodes constitute a chapter, which constitute a solution set for a sub-problem. A sub-problem is one aspect of the travel plan, which can be an accommodation, attraction or activity. When travelers search information on the Internet, they tend to look for information hubs, which is authoritative web pages containing links to a cluster of web pages on a specific destination in order to minimize inter-cluster information seeking cost. However, geographical information and time frame constraints are not well-represented on the Web for San Diego. Accordingly, the subjects kept switching back and forth between online map web sites and destination web sites in order to locate attractions or accommodations.

4.4 Phase II. Comparison of Semantic Models of Travel Information Searchers and Travel Information Space

The goal of the second phase of analysis was to compute and compare semantic models of the subjects with the semantic model of the travel information space. The interviews

were transcribed and aggregated across all the subjects in order to generate a semantic model of travelers through semantic network analysis. The full text of San Diego-related travel web sites were downloaded and analyzed using semantic network analysis (Doerfel, 1998). The number of common keywords was counted and QAP analysis was conducted using the two matrices of common. It is argued that the common and different concepts in the two semantic models reflects the different languages and views of the San Diego tourism industry and the subjects' mental models of San Diego as a tourism destination.

4.4.1 Examining Tourism Information Space of San Diego

Those web sites the subjects visited constitute the travel information space regarding one specific destination. The trip planning exercise showed that this space is huge and much diversified. In total 145 unique web sites were visited by the subjects, which include not only general search engines (e.g. http://www.google.com, http://www.yahoo.com/), general tourism reservation web sites (e.g. http://www.expedia.com/, http://www.orbitz.com/, http://www.hotwire.com/), local tourism destination marketing organizations (DMO) (e.g. http://www.sandiego.org/), web sites of local attractions (e.g. http://www.cafesevilla.com/, http://www.balboapark.org/), but also information portals provided by different commercial parties (e.g. http://www.sandiego.cc/, http://www.a-zsandiego.com/) and general recreational and educational web sites (e.g. http://www.nfl.com/, http://www.nps.gov/, http://www.sdsu.edu/) (see Appendix E for a complete list of web sites visited by the subjects). Msn.com was visited by 13 of 15 subjects but it was the

default page when the subjects launched the Internet Explorer; therefore, it was not counted as the most visited web site. Table 4-3 shows the 33 web sites which was visited by at least two subjects. Google was visited by 13 subjects which was the most frequently visited web site. San Diego Zoo was the most popular attraction web site which was visited by 8 subjects.

In order to obtain the semantic model of the travel information space, the first two levels of web pages of those San Diego related web sites were downloaded and analyzed using semantic network analysis. In the pilot study it was found out that downloading all the web pages on 145 web sites and analyzing them were not feasible. Instead, this research downloaded the first two levels of web pages of those web sites visited by at least two subjects (those web sites in Table 4-3). The reasons for using the first two levels of web pages instead of downloading all the web pages of 33 web sites are: (1) The first two levels of are the primary information and concepts provided by tourism marketers in order to capture the attention of travelers and they represent the most prominent characteristics of the destination; (2) Downloading every web page of those web sites will make the results biased toward those web sites with large quantity of web pages. For example, San Diego Zoo has more than 5,000 web pages which contain many discussions of animals. The analysis results produced a large number of concepts related with animals but they are not necessarily the travel information providers' focus; and, (3) The quantity of all the web pages on the 33 web sites listed in Table 4-3 will occupy several gigabytes of disk space and the semantic network analysis software CATPACII could not handle this amount of data. However, the first

pages of these 33 web sites are not necessarily directly related with San Diego. For example, the first page of yahoo.com is a general web search engine. Therefore, the first directory of San Diego was used (http://local.yahoo.com/?location_state=CA&location_city=San+Diego&location_lat=32.715710&location_lon=-117.156479). Some commercial web sites, including http://www.trafficmp.com/ and http://www.mapquest.com/ also were not included in the analysis since it has no direct connection with San Diego and it does not contain a section on San Diego. Appendix G shows the actual starting web pages where the two levels web pages were downloaded.

Table 4-3. Tourism Information Space for San Diego

Web Site	Number of Subjects Visited	Web Site	Number of Subjects Visited
www.google.com	13	www.sannet.gov	2
www.sandiegozoo.org	8	www.sandiego-online.com	2
www.sandiego.org	7	www.sandiego.cc	2
www.sdcommute.com	5	www.revup.biz	2
www.mapquest.com	5	www.reservetravel.com	2
www.trafficmp.com	4	www.portofsandiego.org	2
www.seaworld.com	3	www.orbitz.com	2
www.sandiego.com	3	www.netster.com	2
www.expedia.com	3	www.infosandiego.com	2
www.blueescape.com	3	www.hotwire.com	2
www.a-zsandiegoattractions.com	3	www.fodors.com	2
www.yahoo.com	2	www.citysearch.com	2
www.thebigbay.com	2	www.cafesevilla.com	2
www.sdsu.edu	2	www.balboapark.org	2
www.sdro.com	2	www.arestravel.com	2
www.sdnhm.org	2	www.4adventure.com	2
www.sdinsider.com	2		

4.4.2 Mapping of the Semantic Model of Tourism Information Space of San Diego

The downloaded web pages were reformatted into text files using HTML2TXT software, aggregated into one text file and analyzed using semantic network analysis (HDSE, 2003). The file included 6,539 words after deleting the stop words (e.g. “and”, “or”, “is” and etc.), including total 2,121 unique words. Figure 4-6 shows the top 160 keywords and their frequencies (the maximum number of keywords can be analyzed in CATPACII is 160). Since it is not feasible to analyze all the 160 words, the top 25 keywords in the file were chosen for the analysis. In tourism information space, the top 25 keywords appeared more than 25 times, and the rest of keywords appeared less than 25 times (Figure 4-6). Table 4-4 shows the top 25 most frequently used keywords generated from semantic network analysis, including its frequencies in the file and the percentage in the total number of keywords.

Figure 4-7 shows the scaling representation of the top 25 keywords generated from semantic network analysis on the travel information space. The top 25 keywords comprise 17.8% of all the keywords used. As can be seen in Figure 4-7, there are four clusters in the semantic model of the travel information space in the four quadrants. In Quadrant I of Figure 4-7, we can see that San Diego is a “city” in “California” and the general “information” regarding “San Diego” is specifically provided on “hotel”, “park”, and “map”; in Quadrant II, the marketing focus is on water activities. San Diego has “Sea World”, “bay” area, “harbor” and “cruise”. They have the “best” “discount” “price” on “tickets”; in Quadrant III, “art” attractions are promoted; in Quadrant IV, the marketing focus of San Diego is on “shopping” “center”, “event”, “museum”,

"restaurant", "center", "service", and "free" events. In conclusion, these results appear to reflect a marketing focus of online tourist information of San Diego. The four clusters reflected the popular marketing foci of travel information providers of San Diego, including water attractions, arts, general sightseeing, shopping and dining.

Table 4-5 shows the centrality measure of each concept and the overall group centrality for the semantic model of the travel information space (using Freeman Betweenness Centrality, Freeman, 1977). Group centrality measures the extent of difference among the centrality measures of the keywords in the semantic network. High group centrality value indicates that certain concepts have higher centrality measure than others. Network analysis software UCINET was used to calculate these measures (Borgatti, Everett & Freeman, 1999). From Table 4-5 we can see concepts "Price" and "Best" have the highest centrality value. This finding suggests that tourism information providers' marketing focus is related to price and promotion of quality. Furthermore, famous attractions were represented in the travel information space of San Diego as represented by keywords "Sea World", "Attraction", "Center" and "Park".

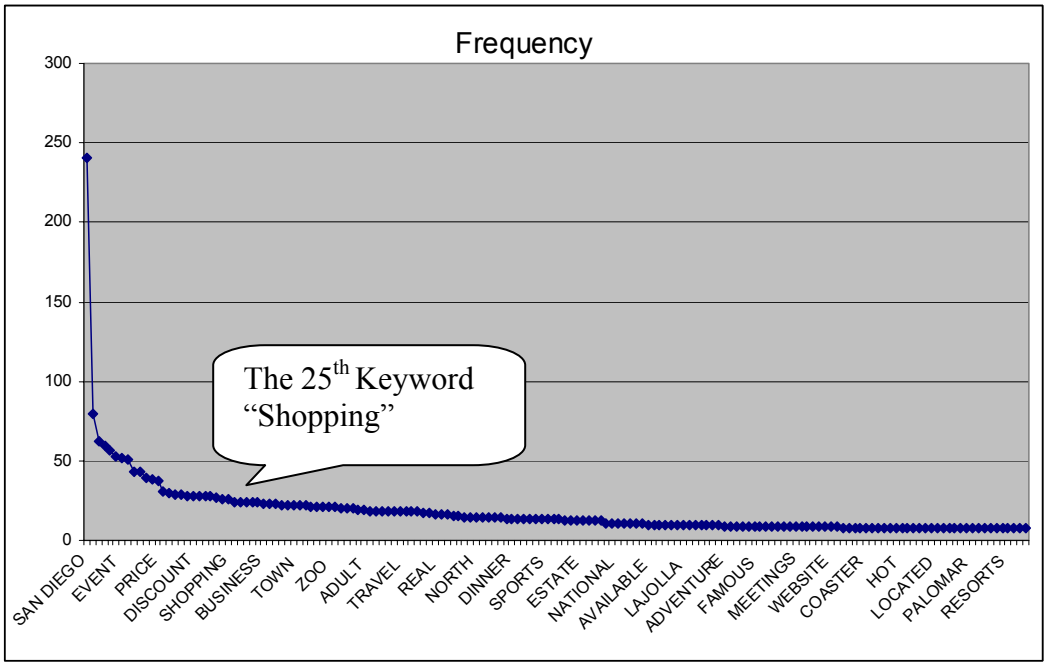

Figure 4-6. Top 160 Keywords and Their Frequencies in Tourism Information Space

Table 4-4. Top 25 Keywords in the Semantic model of Online Tourism Information

Words	Frequency	Percentage	Words	Frequency	Percentage
San Diego	241	8.2	Free	31	1.1
Ticket	80	2.7	Best	30	1.0
Information	62	2.1	Attraction	29	1.0
Hotel	59	2.0	Cruise	29	1.0
Center	57	1.9	Art	28	1.0
Tour	53	1.8	Discount	28	1.0
Event	52	1.8	Map	28	1.0
City	51	1.7	Museum	28	1.0
Park	43	1.5	Sea World	28	1.0
Restaurant	43	1.5	Harbor	27	0.9
Service	39	1.3	Bay	26	0.9
California	38	1.3	Shopping	26	0.9
Price	37	1.3			

Figure 4-7. Semantic model of Tourism Information Space

Table 4-5. Betweenness Centrality of Semantic model of Tourism Information Space

Words	Centrality	Words	Centrality
Price	19.33	Harbor	0.33
Best	7.00	Tour	0.33
Sea World	1.00	Free	0.20
Attraction	0.95	Shopping	0.20
Center	0.95	Art	0.00
City	0.95	Bay	0.00
Event	0.95	Cruise	0.00
Information	0.95	Discount	0.00
Park	0.95	Map	0.00
Restaurant	0.95	Museum	0.00
San Diego	0.95	Service	0.00
California	0.50	Ticket	0.00
Hotel	0.50		

Network Centralization Index = 6.74%

4.4.3 Mapping of the Travel Information Searchers' Semantic Mental Model

The pre-exercise interview was semi-structured and its goal was to elicit the subject's background knowledge of San Diego and their travel preferences related to their understanding of the Internet as a travel information source. The text of the transcribed responses was aggregated together and analyzed following the same procedure described above for generating semantic model of the travel information space (Doerfel, 1998). The aggregated file included 1,601 keywords of which 593 were unique keywords. Figure 4-8 shows the histogram of top 160 keywords and their respective frequencies. Similar to the choice of the top 25 keywords in the travel information space, the top 25 keywords were chosen for the analysis and comparison purpose since the top 25 keywords appeared more than 10 times and the rest of keywords appeared less than or equal to 10 times. Table 4-6 and Figure 4-9 (two dimensional scaling representation) shows the top 25 top keywords. The top 25 keywords comprise 30.6 % of all the keywords in the file.

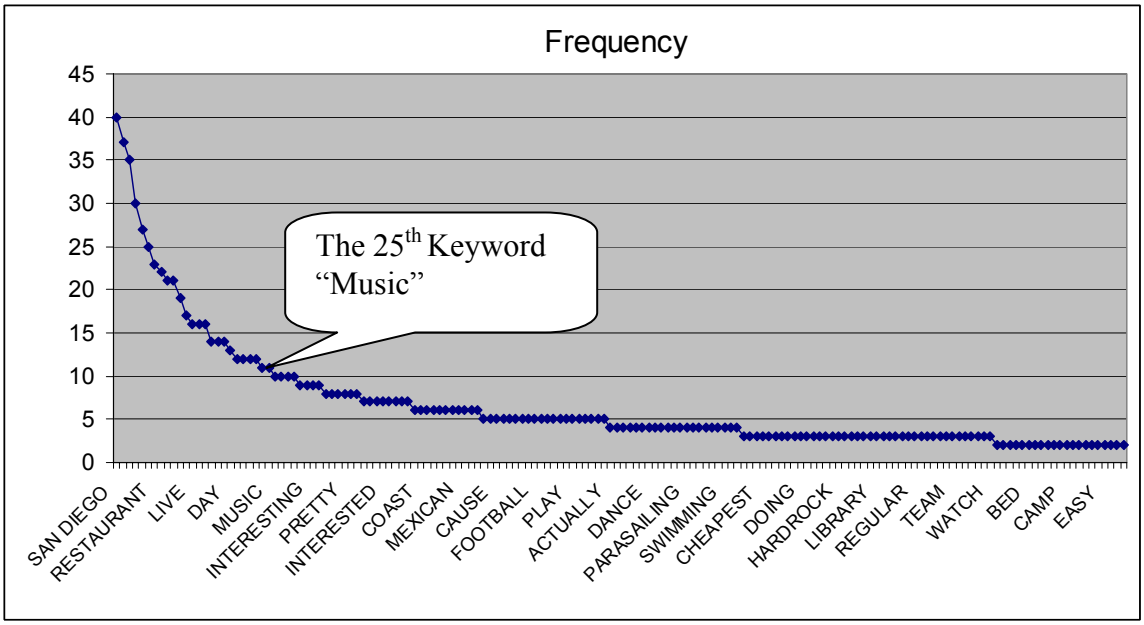

Figure 4-8. Top 160 Keywords and Their Frequencies in Travelers' Mental Model

Table 4-6. Top 25 Keywords in Traveler's Semantic Mental Model

Words	Frequency	Percentage	Words	Frequency	Percentage
San Diego	40	3.7	Stay	16	1.5
Place	37	3.4	Walk	16	1.5
Beach	35	3.2	Look	14	1.3
Museum	30	2.8	Tourist	14	1.3
Food	27	2.5	Zoo	14	1.3
California	25	2.3	Day	13	1.2
Restaurant	23	2.1	Big	12	1.1
Hotel	22	2.0	Different	12	1.1
Information	21	1.9	Good	12	1.1
Lot	21	1.9	Visit	12	1.1
City	19	1.8	Attraction	11	1.0
People	17	1.6	Music	11	1.0
Live	16	1.5			

Figure 4-9. Semantic Mental Model of the Travel Information Searchers

Similar to the semantic model of the travel information space, the subjects' semantic mental model also has four clusters of keywords/concepts. In Quadrant I of Figure 4-9, in general "San Diego" is related with "people", "beach", "city"; the subjects would like to "walk" around "city" and "beach" during the "day", and see "tourist" places and "live" shows. In Quadrant II, the subjects are looking for a good "hotel" or a "place" to "stay", and good "restaurants" for "food". As can be seen in Quadrant III, the subjects are looking for "information" regarding "good" and "different" "museums". In Quadrant IV, the subjects would like to visit "big" "attractions", like the "zoo" and "music" events. In general, we can see the travelers have a certain level of background knowledge regarding San Diego, for example, it's in "California" and has many "beaches". They are looking for information for "hotel", "restaurant", "attractions', and they will "walk" around the "city" in the "day" and they will go to "big" "attractions", like the "zoo". Their criteria for choosing activities and attractions are "big", "good" and sometimes "different".

Furthermore, the centrality measures of concepts in the traveler's semantic mental model were obtained (Table 4-7). Different from the semantic model of the travel information space, the travelers have "City", "Big", "Music", and "Hotel" as the most central concepts. The travelers can easily relate these concepts with other types of concepts/keywords. Besides hotel and dining, the travelers' interests of San Diego are centered on big attractions around the city and music events.

Table 4-7. Betweenness Centrality of Traveler's Semantic Mental Model

Words	Centrality	Words	Centrality
City	90.00	Restaurant	4.35
Big	76.00	Place	0.38
Music	62.00	Day	0.14
Hotel	44.00	People	0.14
Look	34.02	Attraction	0.00
Food	22.00	Different	0.00
Information	22.00	Good	0.00
Visit	22.00	Live	0.00
California	21.03	Museum	0.00
Lot	21.03	Stay	0.00
San Diego	21.03	Tourist	0.00
Walk	16.53	Zoo	0.00
Beach	8.35		

Network Centralization Index = 26.95%

4.4.4 Comparing Travelers' Semantic Mental Model with the Semantic model of Tourism Information Space

Direct comparison between these two semantic models can reveal different languages between tourism marketers of San Diego and the travel information searchers. The common and different keywords indicate their congruence and QAP analysis on the matrices of common keywords can also reveal their relationship. Comparing these two semantic models, there are only 8 common concepts in the top 25 most frequently appeared keywords (See Figure 4-10). The common keywords are mostly general type of concepts such as "San Diego", "California", "information", "city", "attraction", "hotel", "museum", and "restaurant". If travelers look for these types of information, they are more likely to locate their interested information. However, the travelers' semantic model includes more "experiential" and subjective concepts, such as "big", and more action-based verbs, such as "walk" and "look". On the other hand, the

semantic model of the travel information space is, as expected, more marketing-oriented, emphasizing pricing and products/services, such as “free”, “discount”, “price”, “cruise” and “services”. The choices of attractions are also different. For example, “zoo”, “beach”, and “music” were in the top 25 concepts of the subjects’ semantic mental model whereby the travel information space includes “Sea World” as the prominent attraction.

Apparently in the travel information space, there are more diverse keywords, which may be contributed to all types of tourism information providers online. Travelers have more limited vocabulary since they tend to use similar keywords to express their interests and background knowledge. On average each unique keyword appeared 3.1 times in the travel information space, and in traveler’s semantic mental model, each keyword only appeared 2.7 times. The network centrality index for the semantic model of the travel information space is 6.74% and the same measure for the subjects’ semantic mental model is 26.95% (Table 4-8). It indicates that in the subjects’ semantic mental model there is much difference in centrality measures of individual concepts, e.g. some concepts have central positions in the model while others don’t, indicating that the top concepts in traveler’s mind are more prominent compared with other concepts, whereas the concepts in the semantic model of the travel information space are more evenly distributed. Generally speaking, the travel information space contained much more diverse keywords than subjects’ semantic mental model, but the top keywords/concepts were more prominent in traveler’s semantic mental model than in the semantic model of the travel information space.

QAP analysis (Krackhardt, 1987, 1988) revealed that no significant relationship exists between the two matrices from the two semantic models with 8 common keywords (see Table 4-9).

Table 4-8. Comparison of Measures of Two Semantic Models

	Semantic Model of Travel Information Space	The Subjects' Semantic Mental Model
Total Number of Keywords	6,359	1,601
Total Number of Unique Keywords	2,121	593
Average Times for Each Unique Keywords	3.1	2.7
Group Centrality	6.74%	26.95%
Percentage of Top 30 Keywords	17.6%	30.8%

Table 4-9. QAP Analysis of the Two Models

	Value	Significance	Average	Standard Deviation	P(Large)	P(Small)
Pearson Correlation:	0.305	0.123	0.005	0.247	0.123	0.876
Simple Matching:	0.000	1.000	0.000	0.000	1.000	1.000
Jaccard Coefficient:	1.000	1.000	1.000	0.020	1.000	1.000
Goodman-Kruskal Gamma:		0.000	0.000	0.000	0.000	0.000

4.5 Phase III. Assessing Satisfaction of Travel Information Search on the Internet

Different from the first two phases of analysis, variables in the individual level were obtained and analyzed for each subject in an individual level. The goal of the third phase of the analysis was to examine the relationship between three sets of variables for each subject: (1) the level of congruence between each subject's semantic mental model and the semantic model of the travel information space; (2) traveler's individual characteristics (travel experience, computer and Internet use experience, and the experience of using the Internet as travel information source); and, (3) satisfaction of travel information search. Correlation analysis was conducted to test the relationships

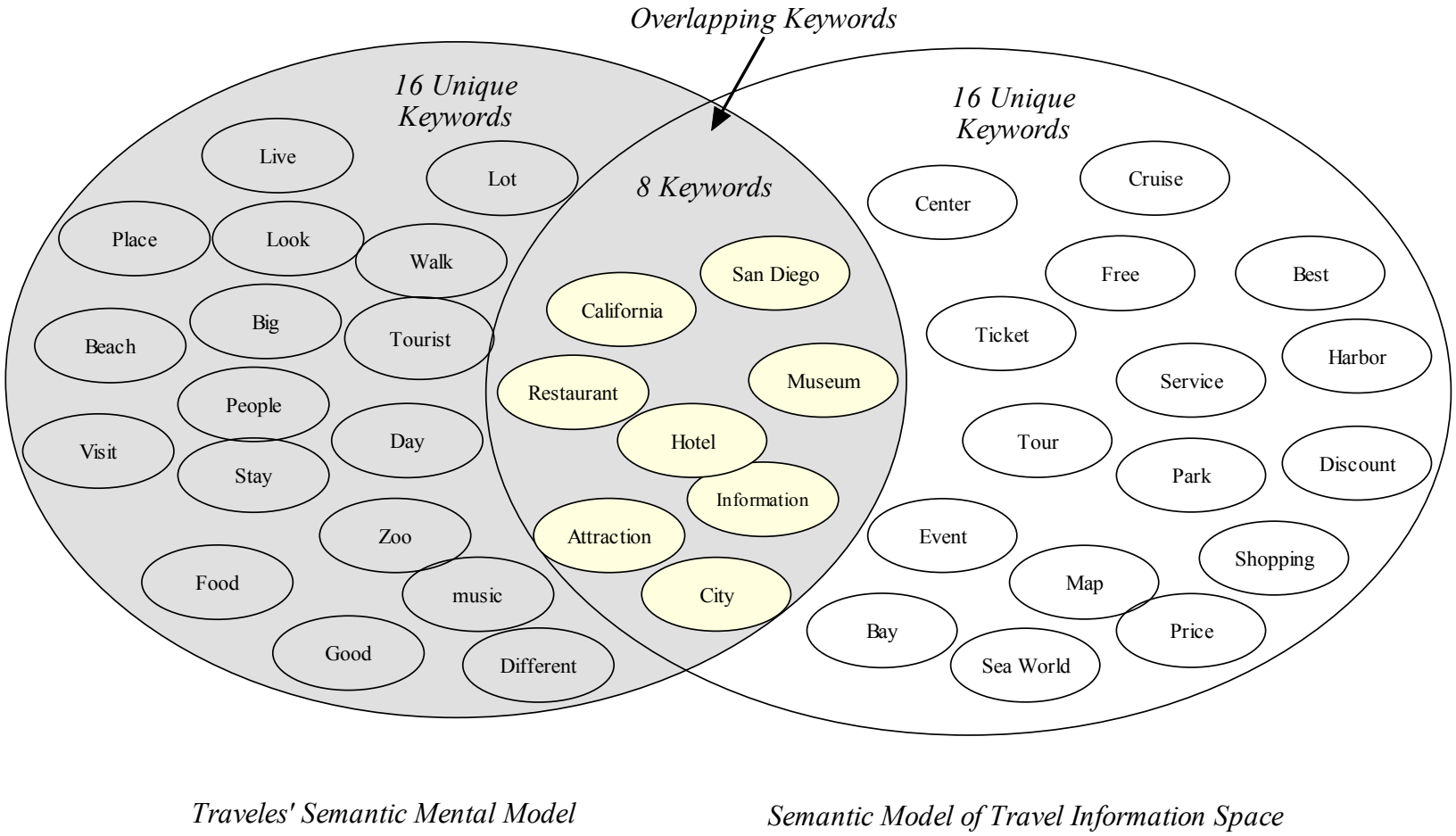

Figure 4-10. Common and Different Concepts of Two Semantic Models

between these variables (see Figure 4-11 for a graphical illustration of data collection and analysis methods used).

4.5.1 A Travel Information Searcher's Individual Characteristics and Satisfaction

As shown in Figure 4-11 Analysis A and B, the travel information searchers' individual characteristics (including their travel experience, computer and Internet use experience, and their experience of using the Internet as a travel information source) and their satisfaction toward travel information search were captured using a pre-exercise questionnaire and a post-exercise questionnaire respectively. The subjects' travel experience was measured by general questions regarding: the number of trips in an average year; the trips taken in last year; self-evaluated travel experience; destination-specific travel experience. The questions regarding the subjects' computer and Internet use experience measure their actual computer and Internet access, computer, Internet, and email use history, amount of email and Internet access, and self-evaluated computer and Internet use experience. The experience of using the Internet as a travel information source was measured by the questions regarding whether or not the travel information has used the Internet to check out destination information, hotel information, airline information, book airline tickets, reserve hotel rooms, and rent cars online were used to measure their experience of using the Internet as travel information source. See Table 4-10 for an overview of subjects' individual characteristics (See Appendix A for the original survey).

4.5.2 Assessing Congruence of Two Semantic Models

As shown in Figure 4-11 Analysis C, in order to measure the congruence between each subject's semantic mental model with the semantic model of the travel information space, the two semantic models need to be acquired. The subjects' semantic mental models were captured through pre-exercise interview regarding their background knowledge of San Diego as a travel destination, their travel preferences, and their understanding of the Internet as a travel information source. The transcripts of their interview were analyzed through map analysis. The semantic model of the travel information space was obtained by downloading the first two layers of web pages of frequently accessed San Diego travel web sites and analyzed through semantic network analysis using CATPACII (Woefel & Stoyanoff, 1993). The semantic model of the travel information space was already obtained in Phase II of the analysis.

Obtaining Each Subject's Semantic Mental Model

The individual travel information searcher's semantic mental model was obtained through map analysis based on the transcript of each subject's interview (Carley & Palmquist, 1992). Map analysis extracts keywords from each individual's verbalization and map into an undirected graph (see Figure 4-12 for an example). When two concepts appear in the same sentence, a link is formed in the graph. The results of map analysis for all the subjects showed that each subject has his/her own idiosyncratic mental model and their mental models range from simple to extremely complex. Comparing the semantic mental map of subject #13 with her clickstream semantic map (Figure 4-4),

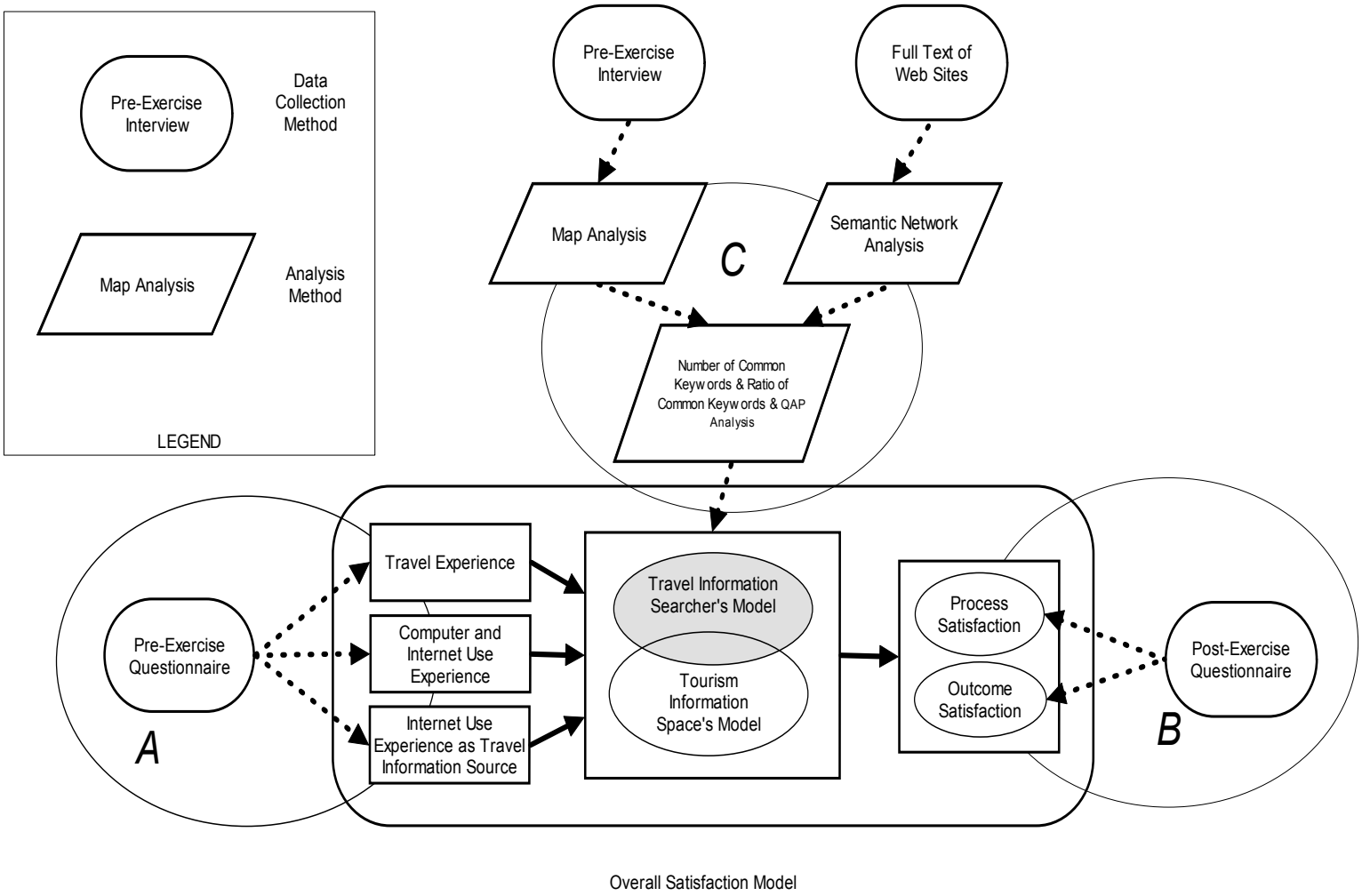

Figure 4-11. Data Collection and Analysis Methods for Examining Satisfaction Model

many prominent concepts appeared on both maps (indicated by nodes with dark background in Figure 4-12). It confirmed the original expectation that people search for information according to their semantic mental models. However, since mental models are always incomplete and fragmented (Kearsley, 2001), some concepts were not present at the navigation graph. See Appendix J for the semantic mental models of each of the 15 subjects.

Measuring the Congruence between the Two Semantic Models

The semantic model of the travel information space was represented by the top 25 keywords generated from the full text of most frequently accessed tourism web sites. The semantic mental model of each subject was generated from map analysis on the transcripts of their interview. Thus, the first measure of congruence between these two semantic models is the number of the common keywords in both models. The second measure of congruence is the ratio of the number of overlapping keywords to the number of total keywords in each subject's semantic mental model which represents the degree of overlapping between two semantic models. Furthermore, the matrix of common keywords in the semantic mental model can be retrieved using a part of the matrix including only the overlapping keywords. The small matrix can be correlated with the common keyword matrix of the travel information space using QAP analysis. Therefore, the third measurement of the congruence is the QAP correlation between the common keywords of two semantic models. Table 4-11 shows the number of common concepts, ratio of common concepts, and correlation results of QAP analysis. QAP analysis showed 6 of the 15 subjects were significantly correlated at .05 significant level.

Table 4-10. Subjects' Individual Characteristics

Survey Item	***Average***
Number of Pleasure Trips Last 12 Months	4.3
Number of Pleasure Trips In a Typical Year	2.6
Self-evaluated Travel Experience	5.3
Been to San Diego or not	4 Yes, 11 No
Family or Relatives Been to San Diego	2 three times, 1 one time, 12 None
Been to California or Not	1 four times, 1 three times, 10 one time, 3 None
Checked Out Destination Information Online or Not	15 Yes
Checked Out Hotel Information Online or Not	13 Yes, 2 No
Checked Out Airline Information Online or Not	14 Yes, 1 No
Booked Airline Tickets Online or Not	10 Yes, 5 No
Reserved Hotel Online or Not	7 Yes, 8 No
Rented Car Online or Not	7 Yes, 8 No
Own Computer at Home or Not	14 Yes, 1 No
Home Computer Online or Not	14 Yes, 1 No
Office Has Computer or Not	14 Yes, 1 No Office
Office Computer Online or Not	14 Yes, 1 No
Years Using Computer	14.3 Years
Years Using the Web	7.5 Years
Years Using Email	6.6 Years
Days per Week Checking Email	6.6 Days
Number of Emails per Day	17.3 Emails
Days per Week Accessing Web	6.5
Hours on the Web One Time	3.1
Self Evaluated Computer Experience	6.5
Self Evaluated Internet Experience	6.4
Gender	9 Females, 6 Males
Age	24.9
Marital Status	13 Married, 2 Not Married
Number of Children at Home	3 One's, 12 zero's

4.5.3 Correlation Analysis on Satisfaction of Travel Information Search

Correlation analysis was conducted on the inter-relationship between three sets of variables: (1) the congruence between each subject's semantic mental model and the semantic model of the travel information space; (2) each subject's satisfaction with travel information search on the Internet; and, (3) each subject's individual characteristics, including her/his travel experience, computer and Internet use experience, and the experience of using the Internet as a travel information source. QAP

correlation has no significant relation with either satisfaction of travel information search or traveler's individual characteristics. Therefore the number of common concepts and the ratio of the common concepts to the total concepts in the traveler's semantic mental model were used to represent the congruence of two semantic models. Table 4-12 shows the correlation model of two measures of congruence with satisfaction of travel information search. The results showed that at .10 level, congruence of semantic models in terms of common concepts is negatively correlated with satisfaction of travel information search (represented by outcome satisfaction), which contradict the our initial expectation.

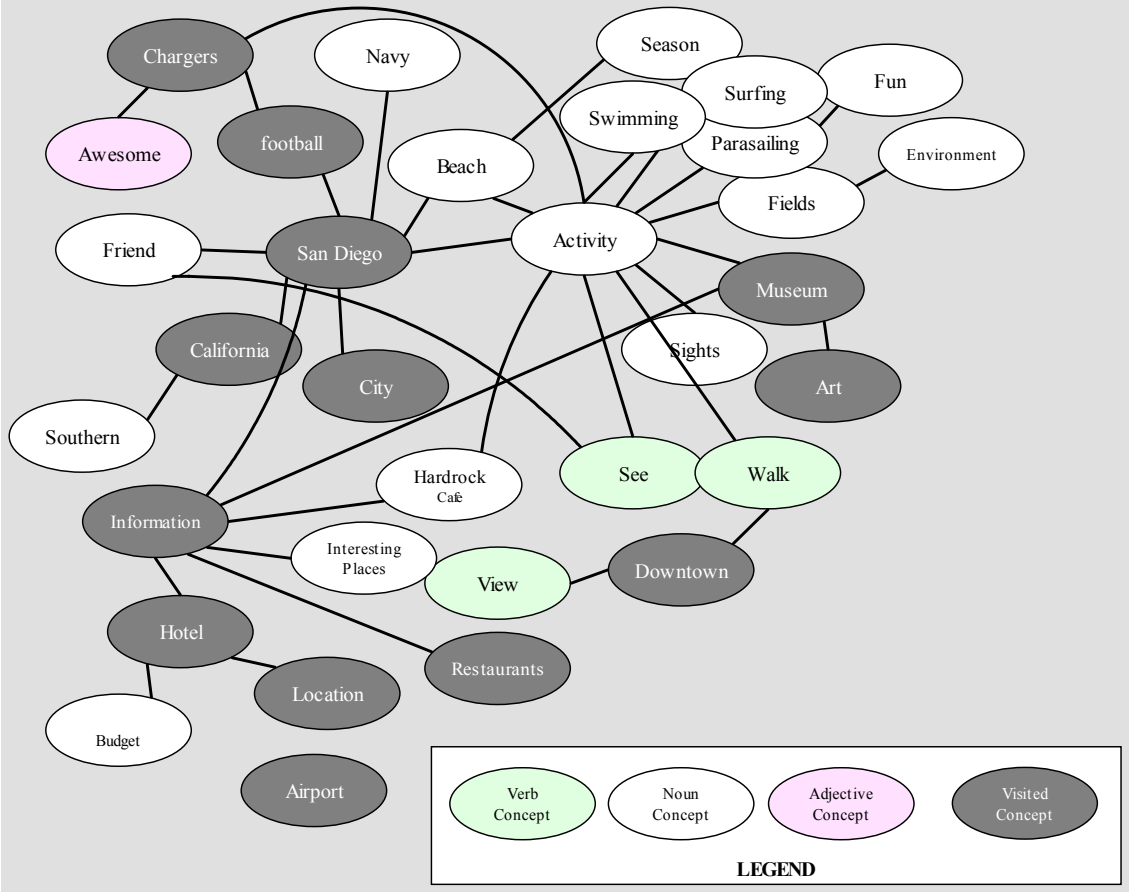

Figure 4-12. Semantic Mental Model of Subject #13

Table 4-11. Correlation of QAP Analysis on Common Keywords

Subject #	Number of Common Concepts	Ratio of Common Concepts	Person Correlation	Significance
1	4	0.20	0.385	0.34
2	3	0.20	0.798	0.33
3	5	0.17	0.278	0.12
4	7	0.41	0.277	0.01
5	6	0.10	0.297	0.26
6	6	0.33	0.479	0.00
7	8	0.16	0.305	0.03
8	6	0.17	0.104	0.50
9	6	0.24	0.359	0.11
10	10	0.14	0.326	0.03
11	9	0.26	0.176	0.26
12	7	0.11	0.236	0.16
13	8	0.24	0.347	0.01
14	4	0.18	0.444	0.05
15	6	0.09	0.131	0.41
Average	6.3	0.20	0.329	--

Table 4-12. Correlation between Congruence of Semantic Models and Satisfaction

	Ratio of Common Concept	Number of Common Concept
Process Satisfaction	-0.12	-0.35
Outcome Satisfaction	0.35	-0.46*
Overall Satisfaction	0.24	-0.46*

* Significant at .10 level.

Correlating Congruence of Two Models with Individual Characteristics

Correlation analysis was conducted on all the individual variables on travel experience, computer and Internet use experience and the experience of using the Internet as a travel information source with the congruence of two semantic models. Table 4-13 shows the significant correlations between the two measures of congruence with each subject's

individual characteristics. The results showed that whether or not travelers have been to San Diego is significantly correlated with number of common concepts since if a subject has been to the destination so they have more to talk about. However, the experience of using the Internet to book airline tickets or rent a car is negatively correlated with the congruence of two semantic models, indicating that the more the travelers are using the Internet for functional travel information search, the less congruence will be their semantic mental model with the semantic of the travel information space.

Table 4-13. Correlation between Congruence of Semantic Models and Traveler's Individual Characteristics

	Ratio of Common Concepts	Number of Common Concepts
Been to San Diego or Not	-0.35	0.59**
Check Airline Online	-0.67**	-0.09
Rent Car Online	-0.20	-0.61**

** Significant at 0.05 level.

Correlating the Subject's Individual Characteristics with Satisfaction of Search

Table 4-14 shows the correlation between each subject's individual characteristics (including travel experience, computer and Internet use experience, and the experience of using the Internet as a travel information source) with the satisfaction of travel information search on the Internet. Travel experience of San Diego is negatively correlated with outcome satisfaction; self-evaluated travel experience contributes positively to outcome satisfaction; the experience of using the Internet is negatively correlated with outcome and overall satisfaction. It indicates that the more experienced a subject feels about her/his travel, the more satisfied will s/he be with the travel

information search results. The more time s/he spent on the Internet, the less satisfaction will s/he will be with the outcomes of travel information search.

Table 4-14. Correlation between Satisfaction and Traveler's Individual Characteristics

	Process Satisfaction	Outcome Satisfaction	Overall Satisfaction
Been to San Diego or Not	-0.47*	-0.45*	-0.49*
Self-evaluated Travel Experience	0.34	0.53**	0.52**
Time Spent on WWW	-0.37	-0.57**	-0.55**

* Significant at 0.10 level; ** Significant at .05 level.

Overall Results

Figure 4-13 shows the final results from the correlation analysis. Travel experience does affect the congruence between semantic mental model of traveler and the semantic mental model of the travel information space positively and it can also negatively affect final satisfaction of travel information search. It indicates that semantic models can capture the travel experience of information searchers. The experience of using the Internet as a travel information source affects the congruence of two models negatively; the more the subjects used the Internet to check airline tickets and rent car online, the less congruent will be their semantic mental models with the semantic model of the travel information space. It showed that semantic mental model can not capture their functional use of the Internet as a travel information source. The computer and Internet use experience are negatively correlated with outcome satisfaction of travel information search, which was not captured in the congruence of the two semantic models. The congruence of two semantic models is correlated with outcome satisfaction and overall satisfaction negatively, which is the opposite of our initial expectation, which indicates that the more similar are the subjects' semantic mental model with the semantic model

of the travel information space, the less satisfied will they be toward their travel information search. Please note that because of the limited sample size (15 subjects), the results showed here is not conclusive. Power of statistical analysis is not large enough for generalization.

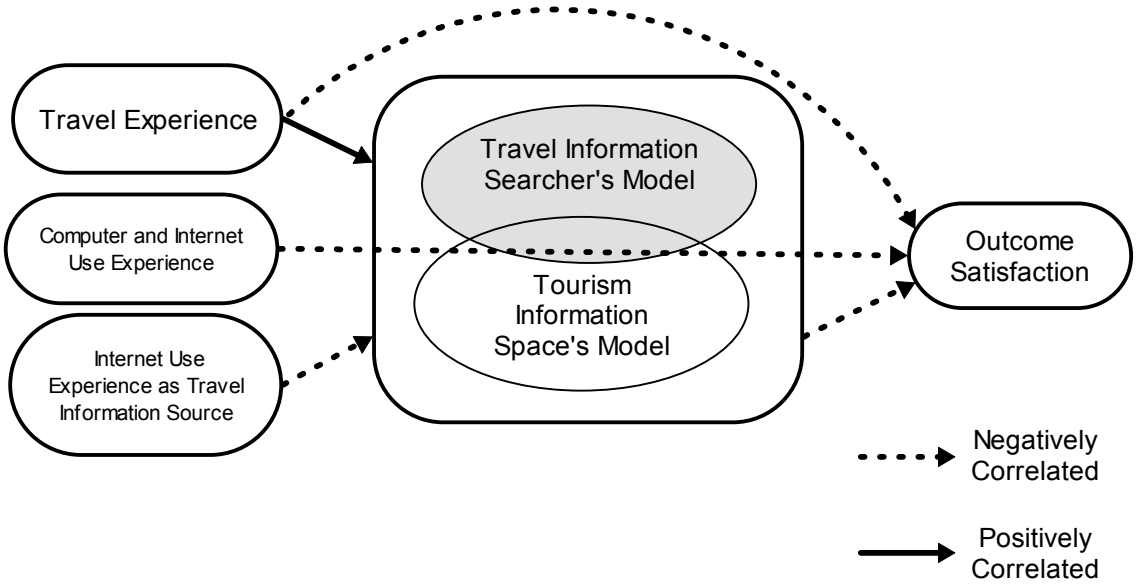

Figure 4-13. Final Results of Correlation Analysis on Satisfaction

4.5.4 Explaining Satisfaction of Travel Information Search

The research results appear to contradict the initial expectations. Congruence of the two semantic models may negatively impact satisfaction of travel information search. To further examine the underlying elements which may determine unsatisfactory travel information search, interviews were conducted after the trip planning exercise in which the subjects were asked "why are you satisfied/not very satisfied with the travel information search?" Results showed that fulfillment of both functional needs and

hedonic needs is indispensable for higher levels of satisfaction. For example, in Subject #14's words:

[trip planning] *"…is easy. Everything is settled, set. I have an itinerary set up already for me. I know when I am leaving, where my care will be at, which hotel I will be staying at, everything is reserved. I probably call a couple days before. I probably call the hotel and probably call the car rental company just to double-check, but basically everything is set up. Everything is paid for. Why I am satisfied? Not worried."*

Similarly, when asked about unsatisfactory travel information search, Subject #15 said:

"Well, I was satisfied, but not completely satisfied, because my ideal planning time for a vacation will probably be like a couple of hours. Maybe between 2 to 5 hours' research. Maybe looking at as much details as I can before making a decision. I feel probably this was as good as I could have done in the time set. I would like to have more time."

These examples indicated that the most important aspect of satisfaction comes from functional needs as trip planning. Subject #14 was satisfied because s/he planned her/his detailed trip; Subject #15 was unsatisfied because of the time constraint and s/he was not able plan the trip in much detail. However, subject #13 explained why she was satisfied with the trip planning:

"…Well, surprises all are found out. I don't know much about San Diego, so it's really exciting to see the Sea World there. Like, wow… I had no idea it was there. I really want to go. That would just be fascinating. You know. I heard about a lot about the zoo. I just really didn't think about it till I saw the site. I'm like, oh, yeah, that's good. You know they have a Hardrock café, so that's really made my day…"

The previous discussion showed that novel information content which is beyond the travel information searcher's mental model and her/his expectation, which leads to learning process and further excitement. The excitement had little to do with the final outcome of the trip plan but was created in the process of travel information search. The subjects were not able to distinguish the excitement stimulated by novel information in the search process with the final outcome; instead, s/he mingle them together.

From the researcher's experience with the subjects when conducting the research, it was discovered that the subjects were very adaptive to the travel information space. Sometimes they were lost in the travel information space as they encountered broken links. However they took these problems as granted and didn't even realize these are problems, and therefore, very little frustration emerged during information search or at least lasted to the end of the information search session. The average high level of satisfaction (6.1 of 7 on Process Satisfaction and 5.8 of 7 on Outcome Satisfaction) indicted that novel information content can stimulate excitement in the search process and can boost the travel information searchers' process satisfaction. The structure of

information which determines the efficiency of a searcher's navigation, is relatively insignificant.

4.6 Summary

This chapter described the detailed results of this research. The trip planning on the Internet has a hierarchical structure whereby it comprises of different layers of chapters and episodes. Several episodes constitute one broad aspect of travel plan or a sub-problem in a travel plan, including accommodations, attractions, activities, or transportations. When the travel information searchers search information on the Internet and plan their trips, their behavior is more than mere navigation and reading, but include a whole spectrum of information behavior, including searching, navigating, and organizing information. Travelers have a high level of satisfaction with the trip planning on the Internet. Travelers' semantic mental models are intrinsically different from the semantic model of the travel information space in that online tourism information is excessively focusing on marketing-oriented information while traveler's semantic models are more subjective and experiential. The results of correlation analysis showed that the congruence between semantic mental models of travelers with semantic model of the travel information space is negatively correlated with outcome satisfaction of travel information search. It indicated that travelers are looking for novel, exciting and enjoyable information (hedonic needs) besides finding facts about attractions and accommodations for trip planning purpose (functional needs).